



Role of a potential pioneer factor in the differentiation of plant stem cells

Eugenia Brun Hernández

► To cite this version:

Eugenia Brun Hernández. Role of a potential pioneer factor in the differentiation of plant stem cells. Vegetal Biology. Université Grenoble Alpes, 2018. English. NNT : 2018GREAV017 . tel-03210490

HAL Id: tel-03210490

<https://theses.hal.science/tel-03210490>

Submitted on 28 Apr 2021

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

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

THÈSE

Pour obtenir le grade de

**DOCTEUR DE LA COMMUNAUTÉ UNIVERSITÉ GRENOBLE
ALPES**

Spécialité : **Biologie Végétale**

Arrêté ministériel : 25 mai 2016

Présentée par

Eugenia BRUN HERNÁNDEZ

Thèse dirigé par **Renaud DUMAS** et codirigé par **Gilles VACHON**,
préparée au sein du **Laboratoire de Physiologie Cellulaire Végétale**
dans l'**École Doctorale Chimie et Sciences du Vivant**

Mécanisme d'action d'un facteur potentiellement pionnier dans la différenciation des cellules souches végétales

Thèse soutenue publiquement le **27 Avril 2018**,
devant le jury composé de :

Dr. Cristina Ferrándiz

Directeur de Recherche, CSIC, Rapportrice

Dr. Charlie SCUTT

Directeur de Recherche, CNRS, Rapporteur

Dr. Francisco Madueño

Directeur de Recherche, CSIC, Examineur

Dr. Marc JAMIN

Professeur, IBS, Président

Dr. Gilles VACHON

Chargé de Recherche, CNRS, Co-directeur de thèse

Dr. Renaud DUMAS

Directeur de Recherche, CNRS, Directeur de thèse



Abstract

LFY is a key transcription factor in plant development, and especially in flowering in angiosperms. It has an important role, first, in the establishment of floral meristems and later, in the specification of their floral organ identities. This activity implicates major chromatin rearrangements in cell nucleus. Target loci need to go from a closed to an opened state. In the last two decades, Pioneer Transcription Factors (PTFs) have been studied because they can bind their target sites on nucleosomal DNA, they are able to overcome the steric constraints of nucleosomes and establish a “competent state” in a particular region, so it can be further regulated by other partners (Iwafuchi-Doi & Zaret, 2014). Besides, genome-wide data analyses strongly suggest that its N-terminal oligomerization domain, confers LFY access to closed regions of chromatin (Sayou et al., 2016). In addition, LFY has been demonstrated to physically and genetically interact with two ATPases belonging to ATP-dependent chromatin remodeling complexes, SYD and BRM (Wu et al., 2012). Altogether, these data suggest that LFY presents common features with PTFs. My PhD work was to better understand LFY’s mode of action in relation to the mentioned ATPases as well as with chromatin.

In Chapter I, through *in vitro* experiments, LFY’s potential interaction with nucleosomes, was assessed. We performed reconstituted nucleosomes by identifying nucleosome-enriched regions in Arabidopsis genome, efficiently targeted by LFY. These regions were selected from ChIP-seq data of LFY in overexpressing lines and DNase-seq as well as MNase-seq data, which were useful to analyze chromatin landscape (Zhang, Zhang, & Jiang, 2015; Zhang, Zhang & Wu, 2012). A strong but non-specific binding of LFY from the gymnosperm *Ginkgo biloba* to nucleosomes was observed. However, LFY from *Arabidopsis thaliana*, showed a weak binding to nucleosomes.

In Chapter II, the aim was to map the minimal necessary interacting domains of LFY and of ATPases SYD and BRM. Using the HTRF technique, LFY’s C-terminal domain was shown to interact with BRM’s HSA domain. In addition, through an *in vivo* approach, we observed the loss of the 35S:LFY phenotype in the F1 plants from each of the three crosses: 35S:LFY *syd-5*, 35S:LFY *brm-1* and 35S:LFY *brm-3*. This suggested a strong genetic interaction, meaning

that when BRM or SYD are not functional, LFY's function is affected and no ectopic flowers are formed.

Résumé

LFY est un facteur de transcription clé dans le développement des plantes, et en particulier dans la floraison des angiospermes. Il a un rôle important, d'abord, dans l'établissement des méristèmes floraux et plus tard, dans la spécification de leurs identités d'organes floraux. Cette activité implique des réarrangements majeurs de la chromatine dans le noyau des cellules. Des loci cibles doivent passer d'un état fermé à un état ouvert. Au cours des deux dernières décennies, les Facteurs de Transcription Pionniers (PTF) ont été étudiés car ils peuvent lier leurs sites cibles à l'ADN nucléosomique, ils peuvent surmonter les contraintes stériques des nucléosomes et établir un état «compétent» dans une région particulière pour qu'il puisse être davantage régulé par d'autres partenaires (Iwafuchi-Doi & Zaret, 2014). Par ailleurs, il a été démontré que LFY interagit physiquement et génétiquement avec deux ATPases appartenant à des complexes de remodelage de la chromatine ATP-dépendants, SYD et BRM (Wu et al., 2012). En outre, des analyses de données à l'échelle du génome suggèrent fortement que son domaine d'oligomérisation N-terminal, confère à LFY un accès à des régions fermées de la chromatine (Sayou et al., 2016). Ces données suggèrent que LFY présente des caractéristiques communes avec les PTFs. Mon travail de thèse a été de déterminer si LFY peut interagir avec l'ADN nucléosomique et se lier aux ATPases mentionnées.

Dans le chapitre I, à travers des expériences *in vitro*, l'interaction potentielle de LFY avec les nucléosomes a été évaluée. Nous avons d'abord identifié des régions enrichies en nucléosomes dans le génome d'*Arabidopsis*, ciblées efficacement par LFY. Ces régions ont été sélectionnées à partir des données génomiques de ChIP-seq de LFY dans des lignées de surexpression ainsi que des données de DNase-seq et de MNase-seq, qui ont été utiles pour analyser le paysage chromatinien (Zhang, Zhang, & Jiang, 2015; Zhang, Zhang & Wu, 2012). Une liaison forte mais non-spécifique de LFY du gymnosperme *Ginkgo biloba* aux

nucléosomes a été observée. Cependant, LFY d'*Arabidopsis thaliana* a montré une faible liaison aux nucléosomes.

Dans le chapitre II, l'objectif était de cartographier les domaines d'interaction minimaux entre LFY et les ATPases SYD et BRM. En utilisant la technique HTRF, nous avons montré que le domaine C-terminal de LFY interagit avec le domaine HSA de BRM. De plus, grâce à une approche *in vivo*, nous avons observé la perte du phénotype 35S:LFY dans les plantes F1 de chacun des trois croisements: 35S:LFY *syd-5*, 35S:LFY *brm-1* et 35S:LFY *brm-3*. Cela suggère que la fonction de LFY requiert la présence de SYD et BRM.

Acknowledgments

The writing of this thesis has required the efforts and support of many people to whom I wish to thank.

I would like to thank first of all the members of my thesis jury, Marc Jamin, Cristina Ferrándiz, Charlie Scutt and Francisco Madueño, to have accepted to be part of my thesis jury and to devote time and energy to evaluate my work.

Merci également aux membres de mon comité de suivi de thèse pour leurs conseils et critiques constructives. Je souhaite remercier aussi à Norbert Rolland de m'avoir accueilli au sein du LPCV, ainsi qu'à Jérôme Garin pour son accueil au BIG.

Un énorme merci à Gilles, pour ton précieux soutien tout le long de ma thèse. Je ne suis pas la première et je ne serai pas la dernière étudiante qui part admirative de ta sérénité et ta bienveillance tout en gardant un esprit critique. Tout cela a rendu le travail avec toi très plaisant. Merci de m'avoir mise en confiance de te partager toutes mes questions et préoccupations. Sans aucun doute tu as rendu ces années passées au labo plus agréables. Merci Renaud pour ta formation initiale au laboratoire, ainsi que pour ton travail d'encadrement. Merci pour ta grande disponibilité et toutes tes attentions. François, merci de m'avoir reçu dans ton équipe. Merci de l'aide que tu m'as apporté en réunion ainsi que pour la préparation des présentations orales. J'ai beaucoup appris en travaillant avec toi autant sur le plant scientifique que relationnel. Finalement, merci à tous les trois de m'avoir apporté votre soutien pour finir ce projet.

Je veux remercier tout particulièrement à Manu, je ne rencontrerai pas deux comme toi. Merci pour les incalculables fois que tu m'as apporté ton aide, pour m'apprendre à faire des manipes, pour faire des manipes avec moi, pour répondre à mes questions, pour me rendre des services, pour m'écouter quand tout allait bien et quand tout n'allait pas bien, pour me faire rire une infinité de fois, pour toutes les sorties que l'on a fait ensemble. Raquelita, gracias por todo tu apoyo, por hacer mis días en el labo más agradables y por ser

una amiga antes que una compañera de trabajo. Catarina ça a été une énorme chance d'avoir partagé avec toi ces années. Merci pour tout le temps que tu as pris pour me donner des précieux conseils, je les ai énormément apprécié. Un très chaleureux merci à Marie. Merci m'avoir appris à travailler avec des plantes avec beaucoup de patience ainsi qu'avec la gentillesse et la douceur qui te caractérise. Merci beaucoup Chloé de m'avoir apporté une aide inattendue et énormément bénéfique dans la partie finale de ma thèse. J'apprécie beaucoup que tu m'aies transmis ta motivation. Merci à Sophie, Danielle et Tiffany, d'être des très gaies voisines de bureau. Merci pour votre excellent travail. À tous les membres de nos équipes, Gaby, Christelle, Véro, Agnès, Aditya, Adrien, Arnaud, Leonie, merci pour vos conseils et d'avoir partagé des moments conviviaux. Je veux remercier aussi à tous nos collaborateurs du projet GRAL notamment à Jean-Luc Pellequer et Jean-Marie Teulon.

Por último pero no menos importante, gracias a mi querido Nico por ser mi principal apoyo y a mi familia, por su amor incondicional.

Abbreviations

3-AT 3-Amino-1,2,4-triazole

A-HSA Adjacent-HSA domain

aa Amino acid

ADE2 Adenine 2, reporter gene

AFM Atomic Force Microscopy

AG *AGAMOUS*, gene

AG *AGAMOUS*, protein

AP1 *APETALA 1*, gene

ALA10 phospholipid-transporting ATPase 10

ALIS1 ALA-interacting sub unit 1

AP1 *APETALA 1*, protein

AP2 *APETALA 2*, gene

AP2 *APETALA 2*, protein

AP3 *APETALA3*, gene

AP3 *APETALA3*, protein

ARP Actin-Related Protein

Asf Anti-Silencing Factor 1

ATP Adenosine triphosphate

ATPase Enzyme using energy from DNA hydrolysis

ATX1 ARABIDOPSIS HOMOLOG OF TRITHORAX1

bHLH Basic helix-loop-helix domain

BiFC Bimolecular Fluorescence Complementation

bp base pairs

cDNA complementary DNA, made from the retro transcription of a messenger RNA

BRM BRAHMA, ATPase

BSA Bovine Serum Albumin

BSH Bushy

c-Myc myc roto-oncogene

ca carpel

CFR Coverage Fold Reduction

CHC Clathrin Heavy Chain

ChIP-Seq Chromatin immunoprecipitation followed by sequencing

CLF CURLY LEAF

CLV1 *CLAVATA 1* gene

Col-0 Columbia-0

CRC Chromatin-Remodeling Complex

DBD DNA-Binding Domain

DHS DNaseI Hypersensitivity

DNA desoxyribonucleic acid

DNase Desoxyribonuclease, enzyme having as a function DNA degradation

DNaseI-Seq DHS assays coupled to high-throughput sequencing

EDTA Ethylene Diamine Tetra Acetic Acid

EMSA Electrophoretic Mobility Shift Assay

EH End-helix Surface

EMF2 EMBRYONIC FLOWER 2

EMS Ethyl methanesulfonate

FLC FLOWERING LOCUS C protein

FM Flower Meristem

FoxA Forehead box transcription factor

FRAP Fluorescence Recovery After Photo Bleaching

GAF GAGA-Associated factor

GATA Proteins that binds to DNA sites with the consensus sequence 5'-[AT]GATA[AG]-3

GFP Green Fluorescent Protein

GR Glucocorticoid Hormone-activated TF

GST Glutathione-S-transferase

HAT Histone acetyltransferase

HDAC Histone deacetylase

HIS3 Histidine 3, reporter gene

HMG High-mobility Group domain

HMGB1 High-Mobility Group Protein B1

HNF3 Hepatocyte Nuclear Factor 3

HSA Helicase/SANT-Associated domain

HTH Helix-Turn-Helix, DNA binding motif

HTRF Heterologous Time-Resolved Fluorescence

IM Inflorescence Meristem

IPT Immunoglobulin-like fold, plexins, transcription factors domain

IPTG Isopropyl- β -D-1-thiogalactopyranoside

kDa Kilodalton

LFY *LEAFY* gene

LFY LEAFY protein

LFY-C LFY's C-terminal domain (DNA binding domain)

LFY-N LFY's N-terminal domain (oligomerisation binding domain)

ISWI Imitation Switch family of ATP-CRCs

Klf4 Krueppel-like factor 4

Ler-0 Landsberg erecta

LncRNAs Long Non-coding RNAs

ML Mid-loop Surface

MNase Micrococcal Nuclease

MNase-seq MNase digestion combined with high-throughput sequencing

mRNA messenger RNA

MADS-box MADS's homeotic genes domain (from the first identified genes: *MCM1*, *AGAMOUS*, *DEFICIENS*, *SRF*)

MS Murashige and Skoog

MW Molecular Weight

p35S Constitutive promoter before RNA of the virus 35S from the cauliflower mosaic virus (CaMV)

Nap1 Nucleosome Assembly Protein 1

NF-1 Nuclear Factor 1 transcription factor

Oct4 Octamer-binding transcription factor 4

Pax7 Paired box 7 transcription factor

PcG Polycomb Group proteins

PCR Polymerase Chain Reaction

pe petal

PHD domain Cys₄-His-Cys₃ motif in the Plant Homeodomain

PI PISTILLATA protein

PKL PICKLE

Pol II RNA polymerase II

POZ Pox virus and Zinc finger domain

PTF Pioneer Transcription Factor

PDB Protein Data Bank

PWM Position Weight Matrix, associating a weight to each nucleotide of the binding site of a transcription factor

qRT-PCR quantitative Reverse Transcription followed by PCR

PTM Post Translational Modification

PU.1 PU-box 1, an ETS-domain transcription factor encoded by the SPI1 gene

RAX1 REGULATOR OF AXILLARY MERISTEMS 1

RNA Ribonucleic acid

SAM Sterile Alpha Motif

SDS-PAGE Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

SAM Shoot Apical Meristem

SDS sodium dodecyl sulfate

se Sepal

SEC Size Exclusion Chromatography

SELEX Systematic Evolution of Ligands by Exponential enrichment

SEP3 *SEPALATA3*, *A. thaliana*'s gene

SEP3 SEPATALLA 3 protein

Sox2 (Sex determining region Y)-box 2

SPR Surface Plasmon Resonance

st stamen

SWI/SNF2 Switch Cucrose Non-Fermentable family of ATP-CRCs

SWN SWINGER

SYD SPLAYED, ATPase

TCEP Tris(2-carboxyethyl)phosphine

TBE Tris Borate EDTA buffer

TF Transcription Factor

UFO UNUSUAL FLORAL ORGANS, protein

WT Wild type

TFL1 *TERMINAL FLOWER 1 gene*

TrxG Trithorax Group proteins, which have a genetic function antagonistic to PcG proteins

TSA Thermal Stability Assay

UIF1 ULTRAPETALA INTERACTING FACTOR 1 protein

ULT1 *ULTRAPETALLA 1 gene*

WUS *WUSCHEL gene*

WUS WUSCHEL protein

Y2H Yeast-two Hybrid

YFP Yellow Fluorescent Protein

Index

INTRODUCTION	23
I. The life of flowering plants	23
1. Our model, <i>Arabidopsis thaliana</i>	24
2. Triggering reproduction	25
3. Flower development	25
II. Flowering is controlled by epigenetic regulation	29
1. Structure-function relationships of chromatin.....	29
2. Chromatin modifications.....	33
A. Histone-modifying enzymes	34
B. Chromatin remodeling complexes (CRCs).....	37
3. Transcription factors (TFs).....	42
A. Pioneer transcription factors (PTFs).....	43
III. <i>Leafy</i> as a potential pioneer transcription factor.....	47
1. A master function in a genetic context.....	48
2. A unique structure.....	49
3. Genome-wide insights of <i>LFY</i> 's pioneer role	53
4. <i>LFY</i> and Chromatin Remodeling Complexes	59
OBJECTIVES.....	63
CHAPTER I. <i>LFY</i> INTERACTION WITH NUCLEOSOMES	65
I. The study of pioneer transcription factors through EMSAs	65
1. Specific and non-specific nucleosome binding.....	66

2. Pioneer and non-pioneer transcription factors	70
II. Nucleosome reconstitution, our biophysical approach	73
1. The nucleosome as a family of structures	73
2. Selection of DNA sequences	74
3. Nucleosome reconstitution.....	78
A. Production and purification of histone octamers	78
B. Nucleosome assembly in vitro.....	81
C. AFM as a tool to verify nucleosome reconstitution.....	83
III. Interactions between LFY and nucleosome core particles	86
1. LFY's binding and specificity for nucleosomes.....	86
A. Comparative binding of AtLFY and GbLFY to free DNA.....	86
B. LFY binding to nucleosomes	89
C. LFY specificity for nucleosomes	93
D. AtLFY and GbLFY structural differences.....	97
2. Other TFs binding to nucleosomes	99
3. Effect of LFY's binding on nucleosomes' structure.....	101
4. Role of LFY's oligomerization	104
IV. Discussion	107
CHAPTER II. TWO LFY PARTNERS: CHROMATIN REMODELLING COMPLEXES SPLAYED AND BRAHMA	110
I. Genetic, biochemical and biophysical context for a complex interplay.....	111
1. Production of recombinant proteins for interaction assays	112
A. Construct design of SYD, BRM and LFY deletions	113
B. Protein production of SYD, BRM and LFY constructs.....	114

C. Protein purification of SYD, BRM and LFY constructs	116
2. Interaction assays.....	118
A. Pull down assay.....	118
B. HTRF assay	120
C. Y2H assay	122
II. Chromatin remodeling complexes and lfy function in planta	125
1. Phenotypic analysis of LFY over-expression in syd and brm mutant backgrounds	125
2. T-DNA Insertions.....	126
III. Discussion.....	130
MATERIALS AND METHODS	134
I. Materials	134
1. Plant material	134
2. Bacterial material	134
3. Yeast material	134
II. Methods	134
1. Plant biology	134
2. Molecular Biology and Protein Biochemistry	138
CONCLUSIONS AND PERSPECTIVES	150
REFERENCES	155

Index of Figures

Figure 1. Siimplified phylogenetic tree of the major groups of plants.	24
Figure 2. Plant organogenesis occurs after embryogenesis.....	26
Figure 3. A meristematic perspective of the floral transition and flower development.	27
Figure 4. Patterning of floral organs.	28
Figure 5. The different levels of chromatin condensation.	30
Figure 6. A gallery of electron micrographs of chromatin.	31
Figure 7. Structure of the nucleosome.	32
Figure 8. ATPase domains characterizing the major SWI/SNF2 family members in Arabidopsis.	38
Figure 9. Model of the nucleosome remodeling process.	39
Figure 10. Activator-mediated targeting model.	40
Figure 11. Domain architecture of Arabidopsis SYD and BRM.....	41
Figure 12. Hypothetical composition of BRAHMA (BRM) and SPLAYED (SYD) CRC complexes.	41
Figure 13. Common DNA-binding motifs.....	42
Figure 14. Model of Pioneer Transcription Factors mechanism.	44
Figure 15. Schematic domain representation of EBF1.	46
Figure 16. Schematic representations of the 3D DBDs of Oct4, Sox2 and Klf4 interacting with DNA.....	47
Figure 17. LFY contains two conserved domains.	50
Figure 18. Logo representing the 19-bp motif targeted by LFY's DBD.	51
Figure 19. Oligomerization of the N-terminal LFY crystal structure from the gymnosperm <i>Ginkgo biloba</i>	52

Figure 20. The LFY-SAM domain is essential for proper flower development in Arabidopsis. .	53
Figure 21. Impaired LFY oligomerization affects genome-wide LFY DNA-binding <i>in vivo</i>	54
Figure 22. Relation of chromatin accessibility to genome-wide LFY DNA-binding.....	56
Figure 23. Genome-wide LFY DNA-binding in relation to chromatin state.	57
Figure 24. Crystal arrangement of LFY DBD on DNA.	58
Figure 25. <i>brm</i> and <i>syd</i> mutants strongly enhance the floral homeotic defects observed in <i>Ify-5</i> mutant.	60
Figure 26. Interactions between LFY, SYD and BRM.	61
Figure 27. Antagonistic effects of CLF and BRM/SYD in <i>AP3</i> and <i>AG</i> regulatory regions.....	62
Figure 28. Schematic representation of a conventional EMSA.....	65
Figure 29. Relative GR affinity to 161-bp nucleosomal templates n3GoA1, n3Go1 and nTG...	67
Figure 30. EMSA assessing the interaction of c-Myc, GATA-4, FoxA1, HMGB1 and NF-1, against two nucleosomes.	68
Figure 31. EMSA assessing the specific and non-specific binding of FoxA1 to nucleosome NCP147.	69
Figure 32. EMSA showing the shift patterns of increasing amounts of recombinant TFs to free DNA and nucleosomes.	71
Figure 34. Model for LFY binding in a nucleosome core particle (NCP).	74
Figure 35. Identification of six DNA regions suitable for nucleosome formation.....	78
Figure 36. Protein sequence alignment of the four core histones using Multialin.	79
Figure 37. Results obtained following the global procedure for the co-expression and co-purification of histone octamers of Shim et al., (2012).....	80
Figure 38. Nucleosome reconstitution by salt dialysis.	82
Figure 39. Nucleosome reconstitution by salt dilution.	83
Figure 40. Nucleosome reconstituted with sequence 4 (482 bp).....	84

Figure 41. Schematic representation of nucleosomes made with DNA sequences of different lengths for EMSAs.	85
Figure 42. EMSAs with AtLFY-40 and different DNA probes, in the presence of increasing concentrations of fish DNA competitor.	88
Figure 43. EMSA of free DNA 5c and 5cm3 (as a negative control) with increasing amounts of GbLFY-40.....	88
Figure 44. EMSA of nucleosome 601 _{LFY} and 2 in the presence of increasing amounts of AtLFY-40 protein.	90
Figure 45. EMSA of nucleosome 5c with AtLFY-40, in the presence of increasing concentrations of fish DNA.	91
Figure 46. EMSA of the interaction between GbLFY-40 and nucleosome 2 and nucleosome 5c compared to nucleosome 601 (negative control).	93
Figure 47. Summary of the binding ability of the two LFY proteins AtLFY-40 and GbLFY-40 to nucleosome 601 _{LFY} , 2 and 5c.....	93
Figure 48. EMSA of nucleosome 5c and 5cm3 (negative control) with AtLFY-40, in the presence of increasing concentrations of fish DNA.	94
Figure 49. EMSA of nucleosome 5c and 5cm3 (negative control) with increasing amounts of GbLFY-40.....	95
Figure 50. EMSA of GbLFY-40 binding on free DNA 5c and 5cm3 with a high DNA to protein ratio.....	96
Figure 51. GbLFY and AtLFY oligomerization interface.	97
Figure 52. Logos constructed from a biophysical model, built on biochemical data and taking into account the three-dimensional structure of LFY DBD in complex with DNA.	98
Figure 53. EMSA of nucleosome 5c against four TFs.	100
Figure 54. EMSA with <i>CLV-1</i> and <i>WUS</i> probes, the respective specific DNA targets for RAX1 and UIF1 proteins.	101

Figure 55. EMSAs of Cy5-labeled DNA 5c and nucleosome 5c (containing Cy3-labeled histones) with increasing amounts of GbLFY-40.	104
Figure 56. EMSA of GbLFY-40 and GbLFY _{TERE} binding comparison.	106
Figure 57. Y2H and pull down assay showing interactions of SYD and BRM with LFY.	111
Figure 58. Schematic representation that summarizes pull down assays between LFY, SYD-N and BRM-N.	112
Figure 60. Constructs designed for LFY, BRM and SYD.	115
Figure 62. Solubilisation and purification of some of the BRM and SYD constructs produced.	117
Figure 63. Representation of GST-LFY and BRM constructs used for pull down assays.	118
Figure 64. Stain-free PAGE gel of the pull down assay between GST-LFY and BRM.	119
Figure 66. Representation of GST-LFY constructs starting from the C-terminal and His-BRM used for HTRF assays.	121
Figure 67. HTRF at six different concentrations of GST-LFY constructs against a fixed concentration of His-BRM ETH428.	121
Figure 68. Representation of the constructs of LFY and BRM proteins used for Y2H assays..	123
Figure 70. Phenotypes of 5-week old plants of the Columbia ecotype.	126
Figure 71. Western blot comparing LFY expression in WT, 35S:LFY, 35S:LFY <i>brm-3</i> and 35S:LFY <i>syd-5</i> plants.	127
Figure 72. Number of ectopic flowers in the parents and the plants from crosses between SALK lines.	129
Figure 73. UV-light exposure to seeds carrying the transgene 35S:LFY, producing GFP.	135
Figure 74. General steps for making dialysis buttons.	141
Figure 75. Simplified LEAFY phylogeny.	152

Index of Tables

Table 1. Overview of the six types of histone modification and their related functions.	34
Table 2. Histone modifications in Arabidopsis and their effects on transcription.	36
Table 3. Predicted/Validated Pioneer Transcription Factors (From Iwafuchi-Doi & Zaret, 2014).	45
Table 4. Primers used for genotyping <i>syd-5</i> , <i>brm-1</i> and <i>brm-3</i>	136
Table 5. Primers used for amplification of DNA fragments for nucleosome reconstitution...	140
Table 6. Oligonucleotides for use as non-specific DNA competitor in EMSA.	142
Table 7. Purified LFY proteins used in EMSA with their respective buffer.	142
Table 8. Primers used for site-directed mutagenesis and amplification of GST-LFY proteins.	144

INTRODUCTION

Developmental processes culminate in the formation of fascinating structures such as flowers. When environmental signals and genetic pathways converge, floral transition begins. This means that the fate of a specific group of cells is settled up to become a flower. A whole gene network tightly controls this phase, in which the floral regulator LEAFY (LFY) plays an essential role. At the cellular level, fate specification involves transcription from regions in the genome that were previously in a silent state. How are these regions targeted in order to be in an active state is an open question. In this introduction, I will explain the context in which my project fits. For this, I start describing flowering plants and the role that LFY plays in flower formation. Afterwards, I discuss about flower development and the epigenetic context in which this process takes place and the main actors involved in the re-organization of chromatin. Finally, I will present the hypothesis that LFY could act as a pioneer transcription factor and I explain the evidences that strongly suggest it.

I. THE LIFE OF FLOWERING PLANTS

Plants are an omnipresent otherness that has conquered practically every kind of ecosystem. The origin of angiosperms, or flowering plants, is one of the most important phenomena, together with the appearance of vascular and seed plants, in the last 480 million years of land plant evolution (Figure 1). In particular, angiosperms are one of the most recent major groups of land plants to have evolved. They diversified and spread very rapidly across the planet, with a fossil record extending back only to the early Cretaceous, more than 130 million years ago (Friedman & Floyd, 2001). In addition, angiosperms have developed important mutualistic relationships with many groups of organisms, such as pollination interactions with insects, birds, and small mammals (Rosas-Guerrero et al., 2014; Van der Niet & Johnson, 2012). Angiosperms are among the most species-rich group of organisms on the planet, and are by far the largest group of plants (Foster, 2016). Among the over 300,000 species, we can observe a great diversity of flowering patterns, yet there appears to be an underlying evolutionary conservation of flowering genes and common patterns of flowering regulation (Gilbert, 2000).

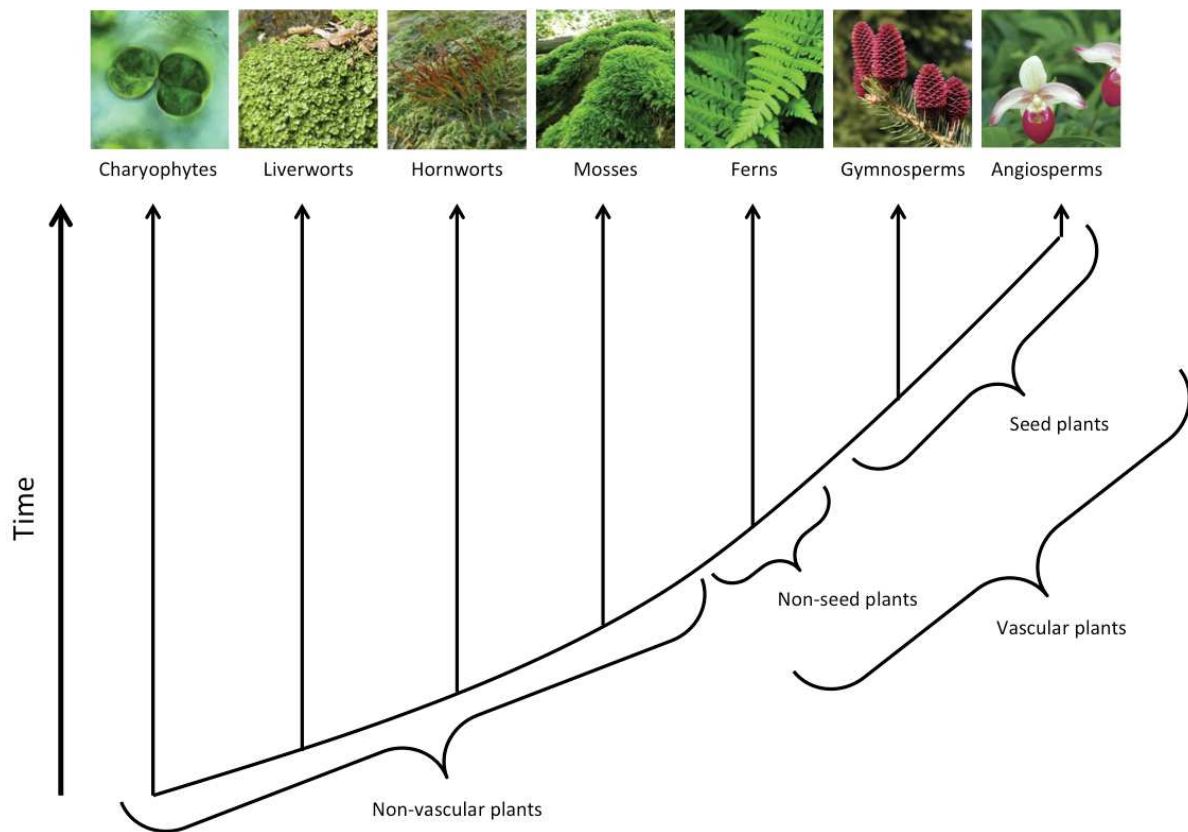


Figure 1. Simplified phylogenetic tree of the major groups of plants.

On top, images of plant groups, such as charophytes, that evolved and developed features (on the bottom), that allowed them to survive in the changing conditions on earth (Adapted from Oliver, Tuba, & Mishler, 2000).

1. OUR MODEL, *ARABIDOPSIS THALIANA*

During my thesis, I worked with *Arabidopsis Thaliana* (named from now on Arabidopsis). This small dicotyledonous species is a member of the Brassicaceae or mustard family. It has been for more than 40 years subject to profound genetic, biochemical and physiological study and it continues to be widely used as a model organism in plant biology.

Arabidopsis is an excellent model for several reasons. It has a small genome (125 Mb total) that has been sequenced in the year 2000, a rapid life cycle (about 8 weeks from germination to mature seed), a large number of mutant lines and genomic resources, many of which are available.

All of the following plant illustrations and figures will refer to Arabidopsis. However, it is important to mention that it is highly possible that the molecular mechanism I tried to shed some light on could be conserved among other eukaryotic organisms.

2. TRIGGERING REPRODUCTION

In Arabidopsis, the transition from vegetative to reproductive development is initiated according to different internal and external signals. The genetic pathways of photoperiod and vernalization induce flowering through the perception of long days and an extended exposure to cold temperatures, respectively. On the other hand, two metabolic pathways that are internally controlled are the gibberellic acid-dependent as well as the autonomous pathway (Farrona, Coupland, & Turck, 2008). Interestingly, these paths come together at the transcription level of a group of genes, called pathway integrators, such as *FLOWERING LOCUS T (FT)*, *CONSTANS (CO)* and *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 (SOC1)*. Since this is a highly active research topic, new findings have shown interactions between epigenetic mechanisms and key players in hormone signaling to coordinate flowering. For instance, as Campos-Rivero et al., (2017) stated that, “some plant hormones, such as gibberellins, jasmonic acid, abscisic acid and auxins, have important effects on chromatin compaction mediated by DNA methylation and histone posttranslational modifications, which hints at the role that epigenetic regulation may play in flowering through hormone action”.

3. FLOWER DEVELOPMENT

Once the mentioned pathways integrate the different signals that induce flowering, flower development is undertaken. In plants, unlike animals, the process of production and development of organs, called organogenesis, occurs continuously during the life of a plant. This is possible thanks to the preservation of plant-specific structures called meristems (Machida, Fukaki, & Araki, 2013).

- An asset of plants, a continual source of growth

The meristem consists of a small area of undifferentiated cells that give rise to almost all structures in the plant. It has the capacity to divide repeatedly and give rise to a number of tissues. Two meristems are established in the embryo, one at the root tip and one at the tip of the shoot. Meristematic cells are progressively set aside to initiate specific tissues or lateral organs (Figure 2).

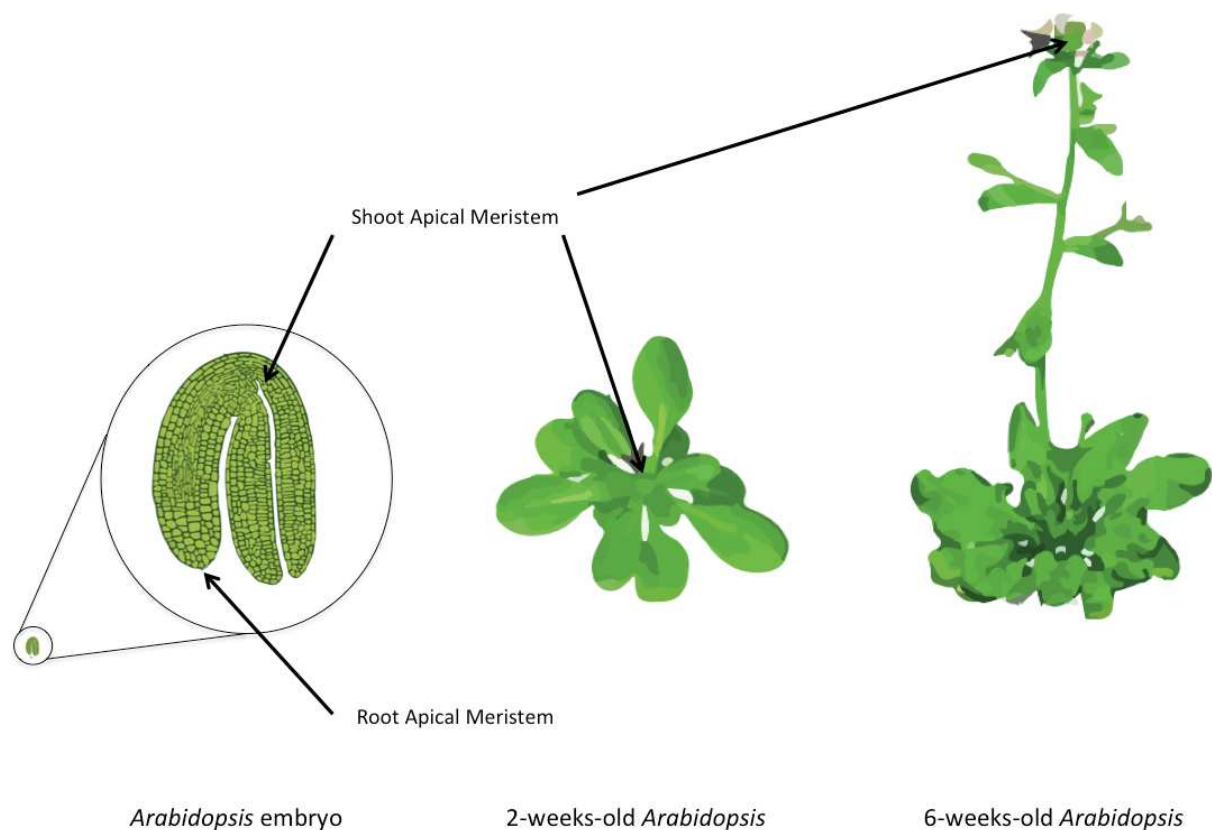


Figure 2. Plant organogenesis occurs after embryogenesis.

Plant organogenesis occurs after embryogenesis, and most developmental programs (vegetative development, floral transition, flower organogenesis) are induced after seed germination. (Engelhorn, Blanvillain, & Carles, 2014).

At the beginning of floral induction, the shoot apical meristem (SAM) switches from a vegetative to a reproductive structure (Figure 3). The SAM is then an inflorescence meristem (IM) that will produce flower meristems (FM) on its flanks. Inflorescence and floral identity are mutually exclusive statuses in the nascent meristem (Denay et al., 2017).

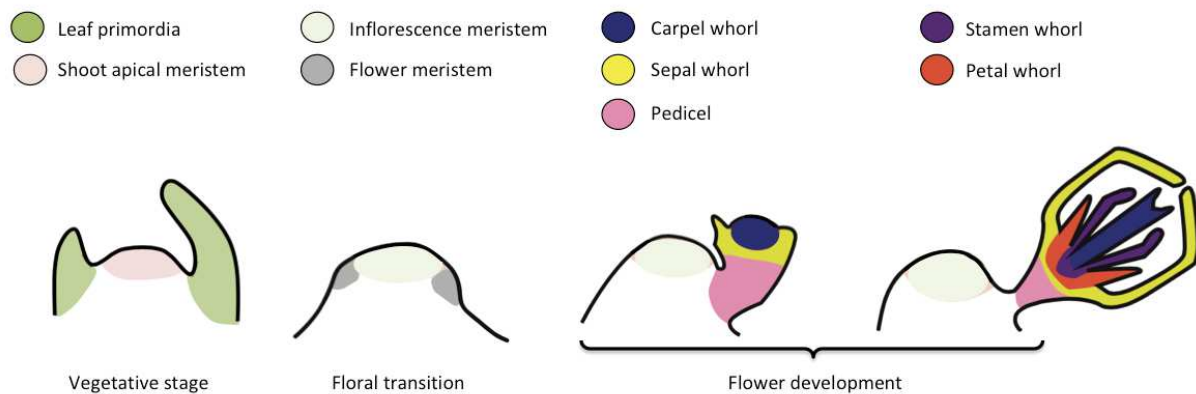


Figure 3. A meristematic perspective of the floral transition and flower development.

Longitudinal sections of flower meristems. The SAM corresponds to the central zone of the shoot and gives rise to aerial organs on the plant. This structure gives rise to lateral primordia that will form leaves during vegetative development or flowers, after floral transition. During this transition, the SAM turns into an IM, which produce in its turn, FMs. Each floral meristem will produce concentric whorls of organs that make a flower. (Adapted from Denay et al., 2017).

One of the most important actor in the flowering process is the LFY transcription factor. This protein, localized both in the nucleus and in the cytoplasm, is able to migrate from cell to cell through plasmodesmata (Lee et al., 2003). LFY arises among other transcription factors as a master regulator. It induces expression of the earliest of group of genes that confers floral identity to cells, called meristem-identity genes. It acts as a regulator of meristem emergence and growth first, and then in the control of flower and floral organ identity (Figure 4). For example, LFY directly upregulates the expression of the meristem-identify gene *APETALA1* (*AP1*), in both leaf and floral primordia. A *lfy* mutant blossoms later than a wild type plant (Blázquez et al., 2006), and on the opposite, LFY overexpression induces precocious flowering, showing that LFY alone is able to trigger flowering (Weigel, Detlef, & Nilsson, 1995).

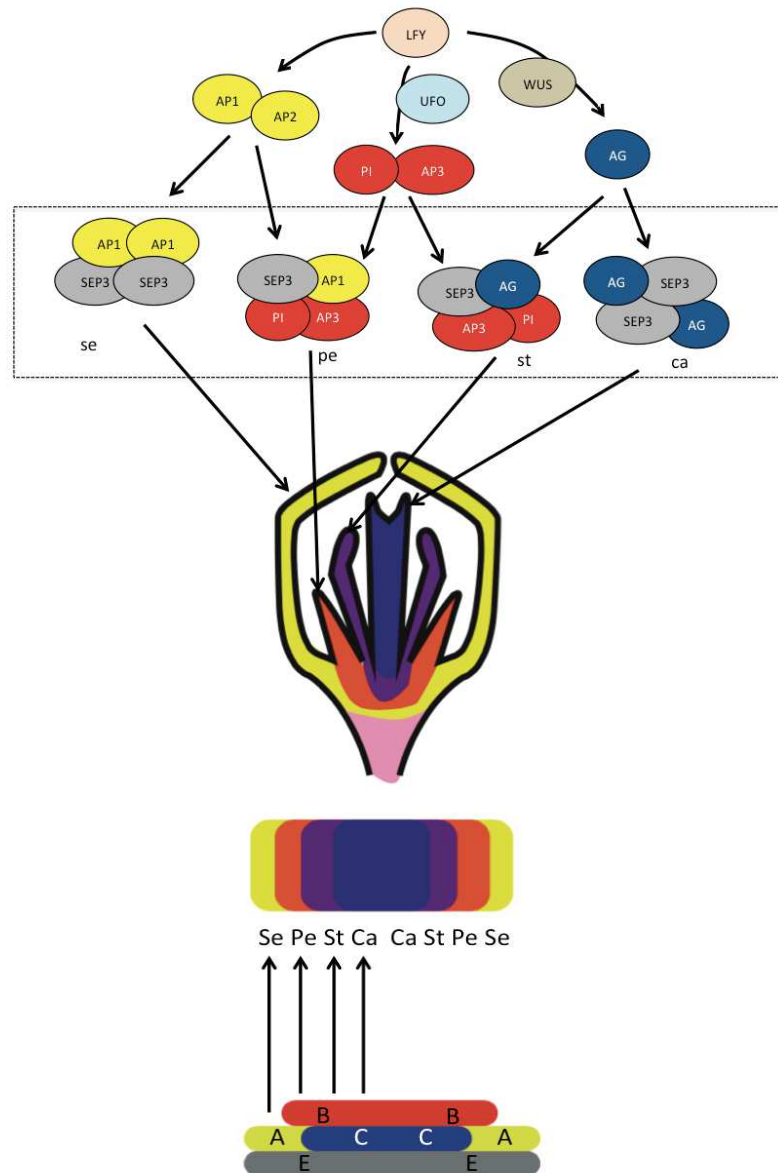


Figure 4. Patterning of floral organs.

APETALA1 (AP1), APETALA2 (AP2), PISTILLATA (PI) and SEPALLATA (SEP) proteins belong to the floral homeotic MADS-box family. (A set of proteins that has developmental functions in nearly all eukaryotes). The combined activity of these different proteins confer identity to floral organs and can be classified in four classes, A, B, C and E. Sepal (se) identity is conferred by class A activity, petal (pe) identity, by class A and B, stamen (st) by class B and C and carpel (ca) identity by class C. Class E activity is required for the specification of all organ types and is fulfilled by SEPALLATA proteins (Pelaz et al., 2000). (Adapted from Krizek & Fletcher, 2005; Silva et al., 2016). The finding that floral MADS domain proteins can associate in higher-order complexes has led to the formulation of the biochemical QUARTET MODEL, represented inside the dotted box, composed by a combination of AGAMOUS (AG), APETALA3 (AP3), SEPALLATA3 (SEP3) and PI (Theissen, 2001). LFY induces the expression of some genes belonging to Class A and B genes together with its co-regulator UNSUSUAL FLORAL ORGANS (UFO) and of class C genes together with its co-regulator WUSCHEL (WUS).

As described above, flower development involves precise activation of several sets of genes at the right place and time. Yet, chromatin compaction in the nucleus represents a physical obstacle for gene expression. Thus, gene expression implicates transformation of the chromatin landscape. “One of the present great challenges in molecular and developmental biology is to understand how chromatin is brought from a repressive to permissive state on specific loci and in a very specific cluster of cells” (Engelhorn et al., 2014). In the next section, I will describe in more detail of how the chromatin structure is organized in different levels of compaction, as well as its dynamic nature.

II. FLOWERING IS CONTROLLED BY EPIGENETIC REGULATION

Floral transition and flower development are key processes for the effective reproduction of plants. Floral transition in particular, is a time of integration of external as well as internal signals. The regulation of floral pathway integrators such as LFY, FLOWERING LOCUS T (FT), and SUPPRESSOR OF OVEREXPRESSION OF CO 1 (SOC1), which are proteins involved in both the photoperiod and vernalization pathways, has been linked with many chromatin-related processes (Farrona et al., 2008). Transcription of these floral integrator genes requires chromatin to be in an active state. Thus, floral transition is an excellent opportunity to study chromatin rearrangements. Similarly, flower development also requires chromatin regulators as illustrated by the requirement of chromatin remodelers for LFY-dependent processes (Wu et al., 2012). In the following sections, this epigenetic context will be more precisely described.

1. STRUCTURE-FUNCTION RELATIONSHIPS OF CHROMATIN

- Chromatin

The chromatin is a complex structure. Its main components are DNA and histone proteins. Its basic repeating unit, the nucleosome, is involved at several hierarchical levels of compaction (Figure 5). This compaction allows fitting the genome of a cell into its nucleus.

Without the interaction with histone proteins, DNA could not be tightly packaged, thus gene expression and DNA replication could not be controlled.

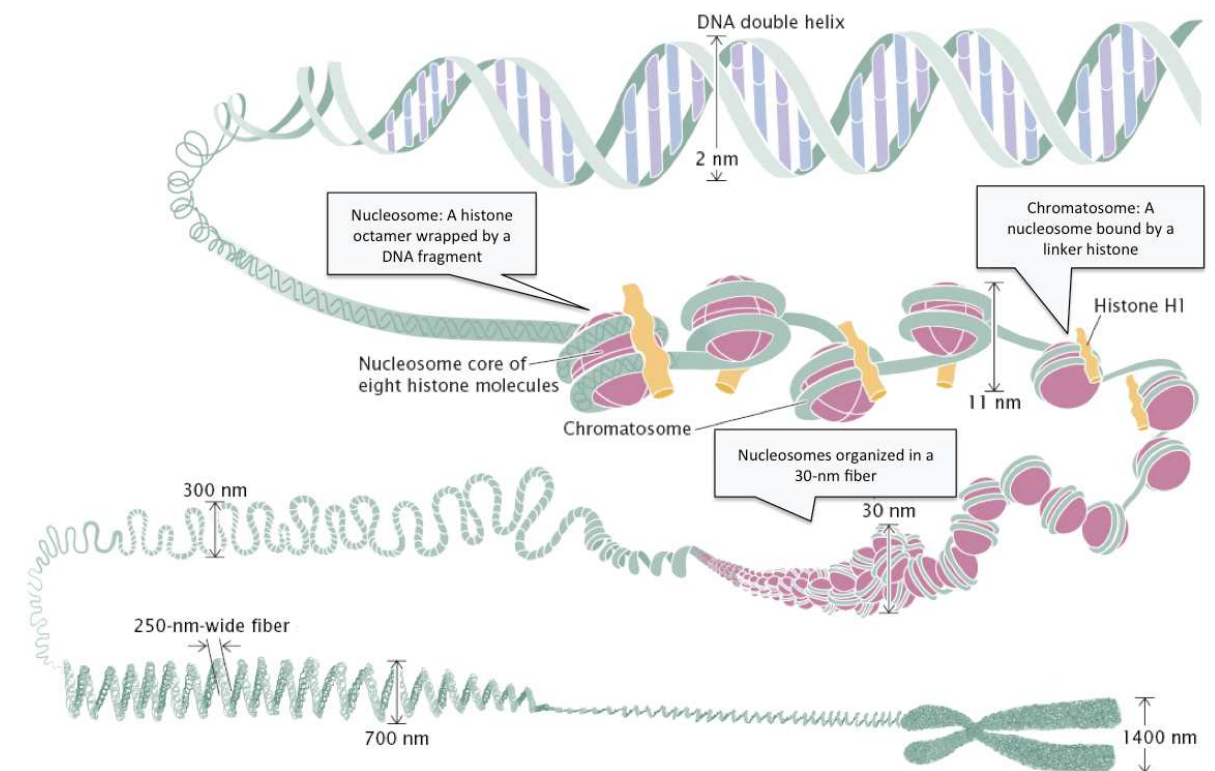


Figure 5. The different levels of chromatin condensation.

The nucleosome core particle consists of a histone octamer to which DNA wraps around ~ 1.7 times. Nucleosomes are further compacted and stabilized by linker histone (H1), which contributes to the formation of higher-order complexes. Chromatosomes, a nucleosome plus a linker histone, then organize in a 30-nm fiber. This fiber further condenses up to the formation of the chromosome.

Electron microscopy allowed visualization of the nucleosome, as well as other different structural organizations (Figure 6). DNA segments called “linker DNA” are present between nucleosomes. When observed at a low ionic strength, nucleosomes often arrange like “beads on a string”, which then organize to compose the 30-nm chromatin fiber. The precise structure of the levels of chromatin compaction higher than the 30-nm fiber is not yet well understood.

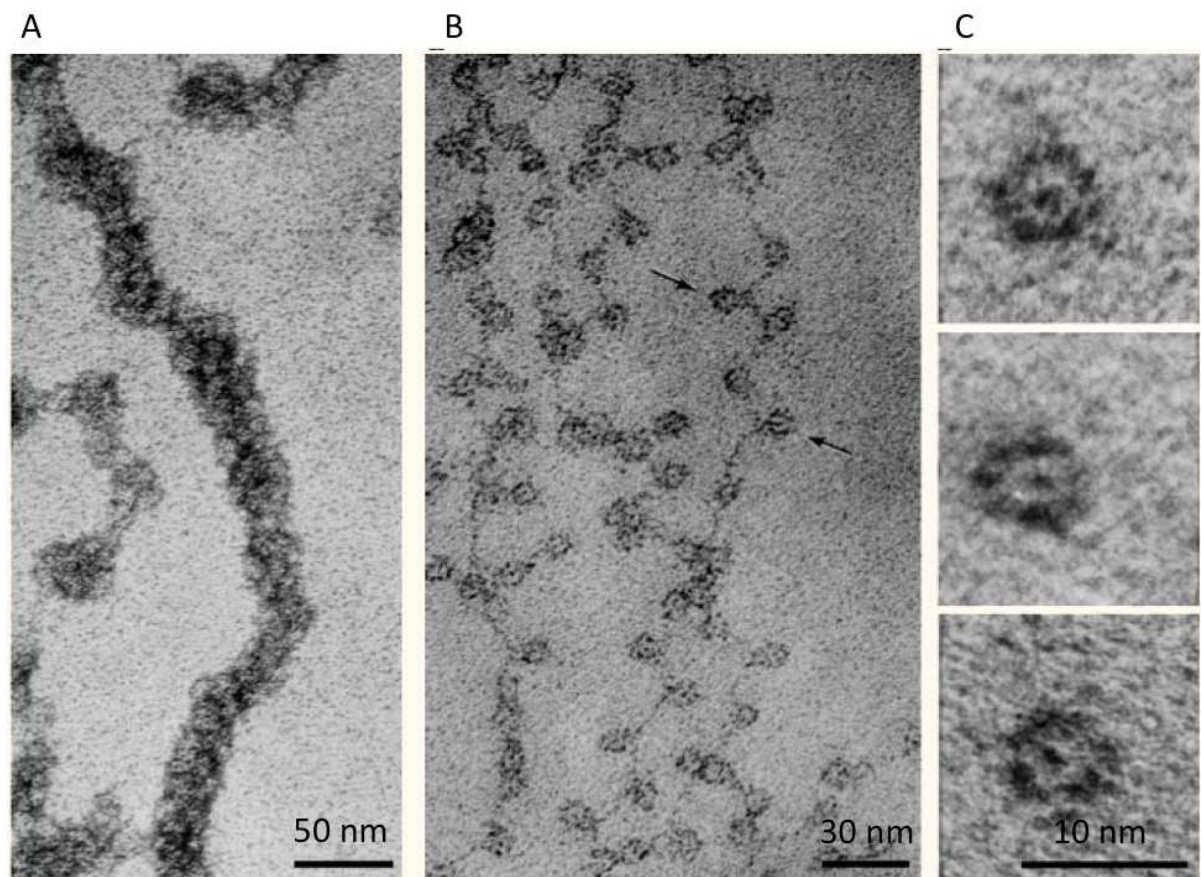


Figure 6. A gallery of electron micrographs of chromatin.

A | Chromatin spread at moderate ionic strength to maintain the 30-nm higher-order fiber. B | Chromatin spread at low ionic-strength, showing the 'beads on a string' structure. C | Isolated nucleosomes derived from nuclease-digested chromatin. (Olins & Olins, 2003).

- The nucleosome

This fundamental repeating unit of chromatin was extensively studied in relation to packaging. This basic unit is not only important for compaction, but it is equally essential for DNA-related processes, such as DNA transcription or DNA repair. The nucleosome is not a simple static unit; it has been shown to adopt a variety of structural rearrangements. Nucleosomes allow DNA compaction, which is itself a way of regulation, together with a variety of complementary systems. The main actors these regulatory systems are: modifying enzymes, histone chaperones, ATP-dependent chromatin remodelers, small RNAs, among others.

As mentioned before, it is a complex composed of a DNA segment of about 146 bp, wrapped around a histone octamer, approximately 1.7 times (Figure 7). The histone octamer is composed of two copies of each histones H2A, H2B, H3 and H4 (Becker et al., 1997). Histones H3 and H4 are arranged into a tetramer, and histones H2A and H2B forms two dimers. Then, the tetramer and dimers come into proximity to form a ramp around which DNA can wrap. The interacting surface between histones and DNA is made of several interactions such as electrostatic and hydrophobic contacts.

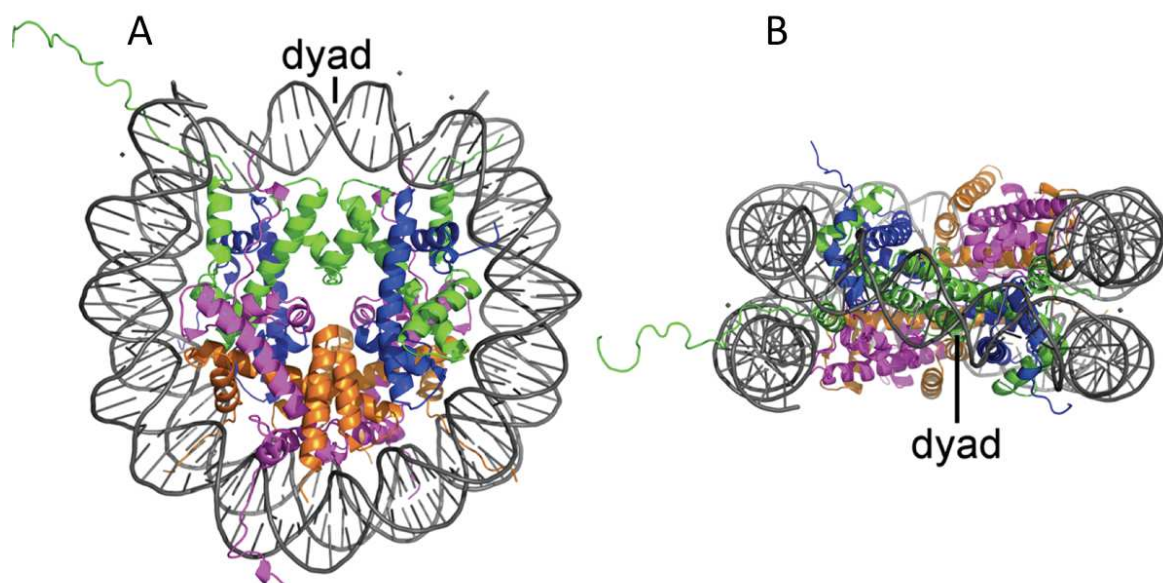


Figure 7. Structure of the nucleosome.

In grey, DNA wrapped around a histone octamer. The eight core histones composed of: H2A (magenta), H2B (orange), H3 (green) and H4 (blue) in a left-handed superhelix. Histones are proteins with a globular center and a disordered N-terminal that protrudes from the nucleosome core. In A, a side view, in B, a top view along the dyad axis. The images were created from the structural data in the RCSB Protein Data Bank (PDB) with the identification code 1AOI (Becker et al., 1997) using the PyMOL Molecular Graphics System (Ordu, Lusser, & Dekker, 2016).

- Core histones

This group of small and basic proteins is encoded by large families of genes and displays a strong degree of conservation throughout the plant and animal kingdom along evolution (Ingouff & Berger, 2010). Histones can be divided into two categories. Canonical histones is

the name for each of the four core histones, when they are expressed during the synthesis phase of the cell cycle and contribute to the packaging of the newly synthesized DNA (Henikoff & Ahmad, 2005). By contrast, the H2A, H2B, and H3 families contain variants, which are expressed during all phases of the cell cycle and can be assembled into nucleosomes in differentiated cells (Bernstein et al., 2006; Malik & Henikoff, 2003). These variants often receive epigenetic modifications mainly by chromatin modifiers and are often essential for development and differentiation processes. For instance, in plants, H3 variants sustain active transcriptional states and their diversification repertoire during the evolution of plants is striking (Ingouff & Berger, 2010).

It is important to stress out that while the term “structure” tends to connote a degree of permanence and stability, it has become abundantly clear that chromatin, at all hierarchical levels (as illustrated in Figure 5), is highly dynamic. In this way, chromatin can be seen as a platform for interactions between genetic and epigenetic factors (Woodcock, 2006).

“A complete understanding of the structure–function relationships of chromatin requires extending primarily one dimensional information, obtained from molecular genetic techniques and based on the underlying linear DNA sequence, to the three dimensional conformation” (Woodcock, 2006). Certainly, our inability to study events in real time inside living organisms has veiled a level of complexity in transcriptional mechanisms that is now emerging. (Voss & Hager, 2014). But there is an increasing awareness that complex dynamics on multiple timescales is central to the activation of appropriate transcriptional programs. In particular, studies that reveal many rapid and cycling molecular processes emphasize the central role of time-dependent events, which can be referred to as the “fourth dimension of gene regulation” (Kærn et al., 2005; Schubert et al., 2013).

2. CHROMATIN MODIFICATIONS

As an active and mobile system, chromatin presents a fascinating flexibility when responding to both internal and external signals in order to regulate DNA-related processes. This flexibility takes the form of changes in DNA accessibility. Proteins that make these changes can be classified in two groups. The first one refers to histone-modifying enzymes or

chromatin modifiers that alter the histone-DNA interactions upon chemical covalent modifications, which I will discuss first in the next section. The second one refers to chromatin remodelers, enzymes that alter the histone-DNA interactions non-covalently, (Hargreaves & Crabtree, 2011) which I will discuss later in the next section

A. Histone-modifying enzymes

These are enzymes that remove or add chemical groups to histones (Li, Carey, & Workman, 2007). The enzymes that perform these modifications chemically alter amino acids at the N-terminal tails of histones, exposed outside the nucleosome core particle. The N-terminal tails are able to easily interact with adjacent nucleosomes as well as with linker DNA. After the modification, protein-binding sites can be created or blocked (Pfluger & Wagner, 2007). In Table 1, the main histone modifications are presented.

Table 1. Overview of the six types of histone modification and their related functions.

In histones, only some amino acids are modified. Generally, acetylation of histone H3 occurs at lysines 9, 14, 18, 23, and 56. Methylation, at arginine 2 and lysines 4, 9, 27, 36, and 79. Phosphorylation is generally made at serine 10 and 28, as well as at threonine 3 and 11. Histone H4 is mainly acetylated at lysines 5, 8, 12 and 16, methylated at arginine 3 and lysine 20, and phosphorylated at serine 1 (Pfluger & Wagner, 2007).

Histone Modification	Examples of Histone-modifying Enzymes	Targer amino acids	Outcome for transcription
Ubiquitination	Ring2 (mammals), Rad6 (<i>S. cerevisiae</i>)	Lysine (K)	Activation (H2B) or repression (H2A)
Methylation	CLF, SWN, MEA (<i>A. thaliana</i>)	Lysine, Arginine (R)	Activation (H3K4, H3K36) or repression (H3K9, H3K27, H4K20, H4R3)
Acetylation	Histone Acetyltransferases (HATs) GNAT Family TAFII Family	Lysine	Activation
Phosphorylation	Ste20 (<i>S. cerevisiae</i>) WSTF (mammals)	Serine (S), Threonine (T)	Activation
Sumoylation	Ubc9	Lysine	Repression
Biotinylation	Biotinilase (<i>S. cerevisiae</i>)	Lysine	Activation

How exactly a histone modification can bring DNA decompaction? Acetylation, for example, which is the addition of acetyl groups to histones, can be explained by electrostatic interactions. When an acetyl group is added to a lysine of the histone tail, the positive charge of the histone is neutralized. The histone is therefore not as attracted to the negatively charged DNA, resulting in a more loose nucleosome structure. Another mechanism by which acetylation can be associated with transcriptional activity, that is true for other modifications as well, is that the acetyl group can be bound by proteins containing specific domains, which will be discussed in the next section. Decompaction by this mechanism can occur in several ways. One possibility is that a protein containing a specific recognition domain for the histone modification can bind the nucleosome and trigger a change in nucleosome stability, such as in the case for ATPases belonging to ATP-dependent Chromatin-remodeling complexes. It is also possible that this protein recruits other proteins, which will exert conformational changes.

Histone methylation, is can be more complex because it does not simply change the charge of the histone. This histone modification, which can exist in the form of mono-, di- or tri- methylation, is associated with either gene expression or gene repression depending on the target amino acid (Tianyi et al., 2015). Three of the most characterized histone modifications are for example, H3K4me (active mark), H3K9me (repressive mark) and H3K27me (repressive mark). These modifications specifically are done by lysine methyltransferases. The polycomb repression complex 2 (PRC2) contains enzymes that deposit the H3K27me mark. After deposition of these marks, other proteins are in charge to read these modifications, acting as docking sites for epigenetic factors.

Most of the roles of the individual histone modifications are conserved between plant and animal kingdoms. In plants, environmental signals such as temperature or light provoke the deposition or removal of modifications. In this way, gene expression is regulated in order to respond to these cues. For instance, prolonged exposure to low temperatures silences the flowering time repressor FLOWERING LOCUS C (FLC) in Arabidopsis by histone deacetylation and repressive histone methylation (Schmitz & Amasino, 2007; Sung & Amasino, 2006). Moreover, the histone deacetylase 19 (HDA19) and two histone acetyltransferases (HATs), HAF2 and GCN5, have negative and positive roles, respectively, in light-responsive gene

expression and in photomorphogenesis (Benhamed et al., 2006). Table 2, presents different types of histone modifications in Arabidopsis.

Table 2. Histone modifications in Arabidopsis and their effects on transcription.

Also listed are the enzymes known to add and remove each modification. In asterisk, (*) all modifications of a certain type, for example, H3K9me* denotes mono, di, and trimethylation of lysine nine of histone H3, while H3*ac denotes general acetylation of lysines in histone H3. (Adapted from Pfluger & Wagner, 2007).

Histone Modification	Established by	Removed by	Outcome for transcription
Ubiquitination			
H2BK143ub1	Ring-Type E3 ligases HUB1, HUB2	Deubiquitinases SUP32/UBP26	activation
Methylation			
H3K4me*	trxG class of histone methyltransferases (HMTs)	LSD1-type of histone demethylases (HDMs) FLD	activation
H3K9me*	Su(var) class of HMTs	JmjC-domain and LSD1-type HDMs	activation
H3K27me3	CLF, SWN, MEA		repression
H4R3sme2	Arginine methyltransferases SKB1/AtPRMT5	Deimination	repression
Acetylation			
H3K*ac/H4K* ac	Histone acetyltransferases (HATs) GNAT family: GCN5/HAG1 CBP/p300 family: HAC1, HAC5, HAC12 TAFII family: HAF2/TAF1	Histone Deacetylases (HDACs) RDP3 family: HDA19 HDA6 HDA1 family: HDA18 HD2 family: HD2A, HD2B, HD2C	activation
Phosphorylation			
H3S10ph H3S28ph H2T11ph	Kinases	Phosphatases	activation

Research on histone modifications is in constant evolution. Genome-wide studies are especially useful and have shown, for example, that looking at specific chromatin modifications can be a good indicator of the chromatin state. Recent work from (Engelhorn et

al., 2017) on *A. thaliana* revealed tri-methylation of lysine 4 on histone 3 (H3K4me3) as a greater predictor over H3K27me3 for transcription dynamics, which reveals chromatin mechanisms for gene activation and underlines the relevance of tissue- and time-specific epigenomics.

B. Chromatin remodeling complexes (CRCs)

Chromatin remodeling in eukaryotes is a fundamental mechanism that allows the transition from repressive to active states of the chromatin structure during the life of a cell. Chromatin-remodeling complexes (CRCs) can change the positioning, occupancy and composition of nucleosomes.

Here I will focus on ATP-dependent chromatin remodeling complexes. They were first discovered in yeast and homologs have been also found and studied in other organisms, like *Drosophila melanogaster*, *Xenopus laevis*, *Arabidopsis thaliana*, as well as humans (Tsukiyama & Wu, 1995). These findings illustrate the importance of the function of these proteins throughout evolution. ATP-dependent CRCs, use energy from ATP hydrolysis to destabilize and weaken interactions between histones and DNA (Han et al., 2015). They are large (>1 MDa) multi-component complexes, consisting of between 4 and 17 subunits (Tang, Nogales, & Ciferri, 2010).

Every ATP-dependent CRC contains an ATPase subunit and so these complexes have been classified, according to the type of ATPase subunit, in two super families: ISWI and SWI/SNF2. The latter super family can be divided in five sub families: INO80, SWR1, CHD, and SNF2, according to their conserved ATPase domains. A unique property of the SNF2 subfamily is its ability to disassemble nucleosomes and to induce stable alterations in the nucleosomal DNA (Phelan, Schnitzler, & Kingston, 2000). These domains as well as other conserved domains are indicated in Figure 8. The relation between the function of ATP-dependent complexes and chromatin was reinforced when the purified complex was shown to remodel nucleosomes in the presence of ATP *in vitro* (Peterson & Herskowitz, 1992).

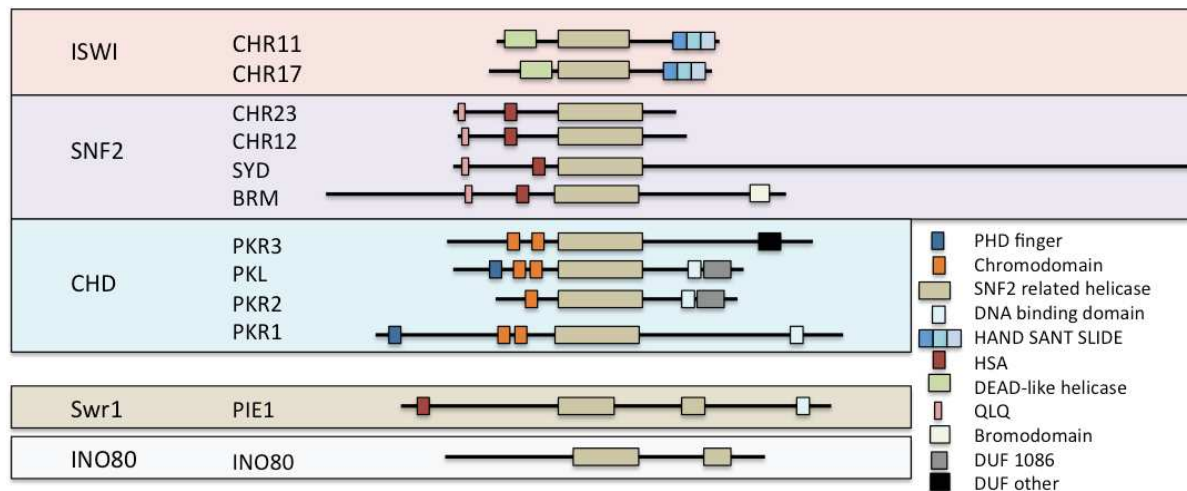


Figure 8. ATPase domains characterizing the major SWI/SNF2 family members in Arabidopsis.

Several common domains have been identified. The PHD finger domain has been found to be involved in protein-protein interactions. The Chromo domain is important for chromatin targeting (Shaked, Avivi-Ragolsky, & Levy, 2006). The SNF2 related helicase domain, binds DNA and hydrolyzes ATP. The HAND SANT SLIDE domains are involved in nucleosome and DNA binding. The HSA domain, for helicase/SANT-associated, is involved in actin binding (Tang et al., 2010). The DEAD-like helicase domain, participates in the unwinding of nucleosomal DNA. The QLQ domain is important for protein-protein interaction (Shaked et al., 2006). The bromodomain, interacts specifically with acetylated lysines in histone tail peptides (Dhalluin et al., 1999). Finally, the DUF's domains are of unknown function (Adapted from Gentry & Hennig, 2014).

- Mode of action

Mechanically, the remodeling process begins with the binding between the nucleosome and the whole remodeling complex, composed of several subunits, apart from the ATPase subunit. Based on single molecule experiments, a denominated translocase subunit, comprises a torsion subdomain and a tracking subdomain. This domain is able to bind nucleosomal DNA situated at approximately two turns from the dyad. (The dyad is the axis of symmetry close to which the entering and exiting DNA duplexes lie within a nucleosome). After ATP hydrolysis, the torsion subdomain initiates nucleosome unwrapping, repositioning or ejection of the histone octamer (Havas et al., 2000; Zhang et al., 2006), (Figure 9).

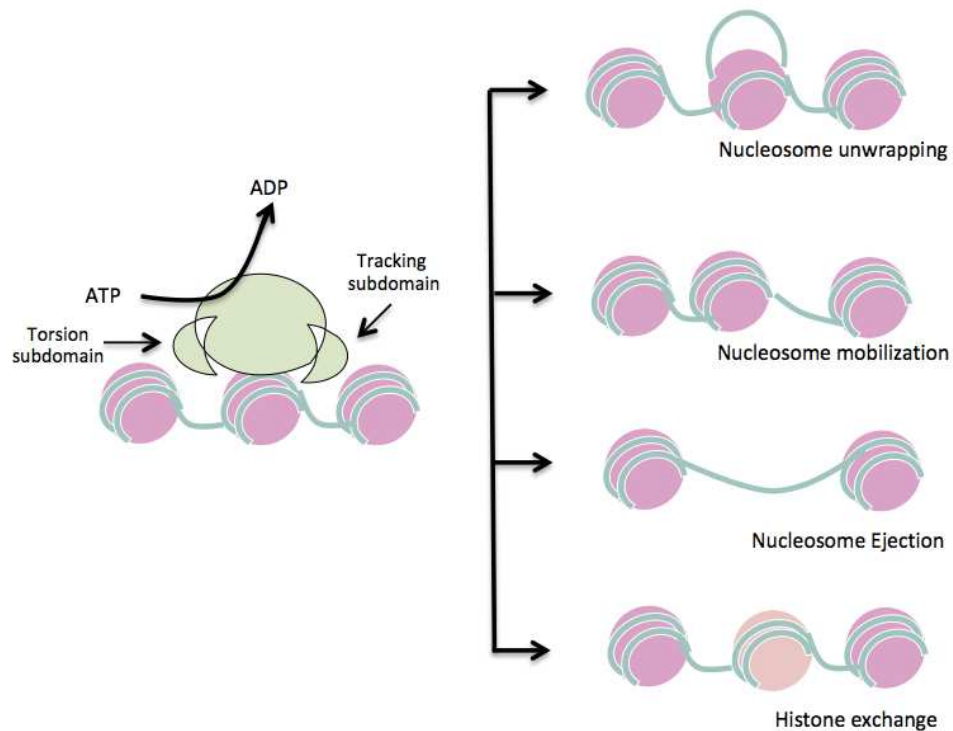


Figure 9. Model of the nucleosome remodeling process.

Different effects of an ATP-dependent CRC (in green) on nucleosomal DNA. The torsion and tracking subdomain are the main actors for the disruption of histone-DNA interactions. Nucleosomes can be unwrapped, mobilized or ejected. In some cases ATP-dependent remodeling complexes can use ATP to introduce histone variants within the nucleosome by a process called dimer exchange (Adapted from Tang et al., 2010).

Concerning the specificity of these CRCs toward genomic regions or specific loci, several targeting mechanisms have been observed. CRCs can be recruited at specific loci by transcription factors or Long non-coding RNAs (LncRNAs) (Wu et al., 2012). In Figure 10, a model for specific recruitment is illustrated.

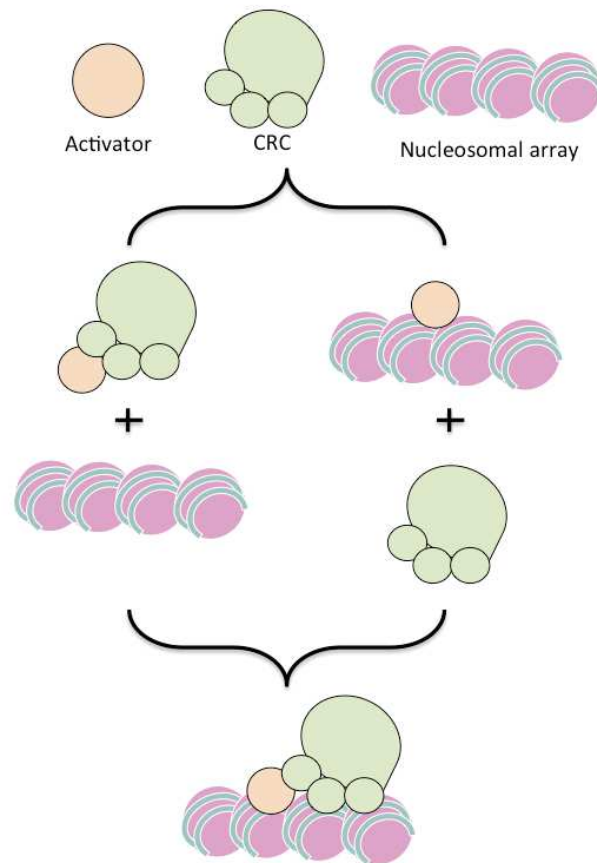


Figure 10. Activator-mediated targeting model.

Transcriptional regulatory factors (yellow circle) may bind first either to the target loci or to CRCs (green circles representing sub-units of the CRCs). (Adapted from Vignali et al., 2000).

- SPLAYED and BRAHMA

During my thesis project I was particularly interested in two plant CRCs: SPLAYED (SYD) and BRAHMA (BRM). They have been proposed to act as Trithorax Group (TrxG) proteins because their function seems to counteract Polycomb Group (PcG) proteins' function. Particularly, BRM from Arabidopsis (AtBRM) is an activator of floral organ identity genes such as *AP2*, *AP3* and *PI*, which are controlled by PcG proteins. (Li et al., 2015). SYD is required for stem cell maintenance in the shoot apical meristem and directly promotes expression of *WUS* (Kwon, 2006). Both CRCs play a variety of functions and are of great importance not only for flower development but also throughout the plant's life. *brm-3 syd-5* double mutants show a striking pin inflorescence phenotype (Wu et al., 2015) and *brm-1 syd-5* double mutants are embryonic lethal (Bezhani et al., 2007), (Figure 11).

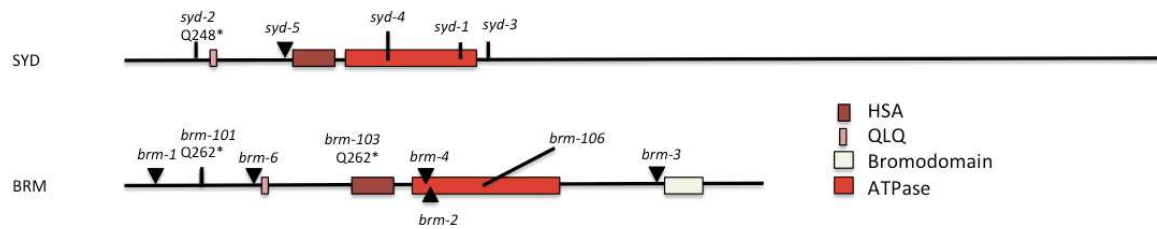


Figure 11. Domain architecture of Arabidopsis SYD and BRM.

Domain prediction is based on InterPro (<http://www.ebi.ac.uk/interpro/>). Mutations and T-DNA insertion sites are indicated. The asterisk indicates a stop codon. The black triangle indicates a T-DNA insertion. (Han et al., 2015).

Even though the genetic analysis of SYD and BRM can give precious information about their function and besides the fact that yeast homologs of these subunits alone have been shown to remodel nucleosomes *in vitro*, several other subunits are necessary to form the whole complex (Figure 12). The other subunits play a role in enhancing the ATPase activity of SYD and BRM, also, they mediate interactions with other protein complexes.

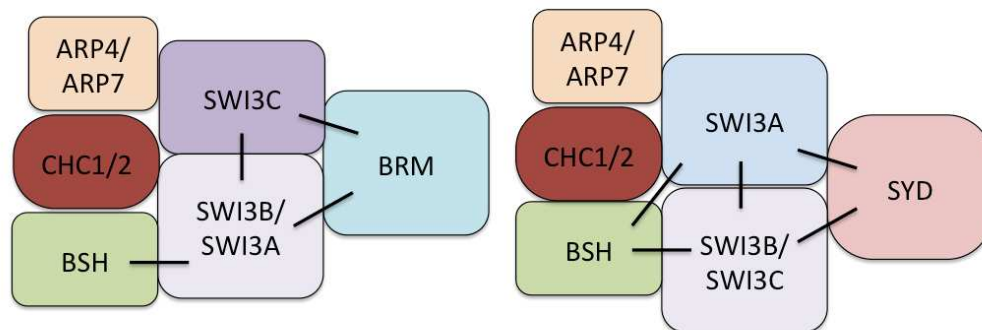


Figure 12. Hypothetical composition of BRAHMA (BRM) and SPLAYED (SYD) CRC complexes.

Association of known subunits with BRM (left) and SYD (right) chromatin remodeling complexes based on phenotypic analyses, binary interaction tests and proteomic data in yeast (Bezhani et al., 2007; Hurtado, Farrona, & Reyes, 2006; Sarnowski et al., 2002; Smaczniak et al., 2011; Vercruyssen et al., 2014). Interacting proteins are connected by dark lines. The actin-related protein (ARP) subunits are involved in interactions with actin proteins. The clathrin heavy chain (CHC) subunit is involved in DNA-binding. SWI3A/B/C subunits participate in nucleosome positioning. Bushy (BSH) is involved in the recruitment of the CRCs to target genes.

In conclusion, chromatin remodelers play pivotal roles in developmental switches. As mentioned in (Gentry & Hennig, 2014) “The overlapping roles and diversification of these

different subfamilies are likely a reflection of the increased complexity of regulatory mechanisms affecting chromatin in plants ". Some important questions that remain open concerning these protein complexes is their role in relation to euchromatin, and what is their implication in non-regulatory regions of chromatin.

3. TRANSCRIPTION FACTORS (TFs)

TFs play important roles in the regulation of gene transcription. Even if they are not the only proteins required, TFs influence positively or negatively the transcription of a particular gene at a specific time, mainly by binding to specific sequences on DNA. Some transcription factors bind to a DNA promoter sequence near the transcription start site, others bind to regulatory sequences, such as enhancer sequences (Phillips & Hoopes, 2008).

TFs often have certain specific DNA-binding domain. Some of them such as the basic helix-loop-helix (bHLH) structure, the helix-turn-helix, the zinc finger, and the leucine zipper structures, are represented in (Figure 13) (Phillips & Hoopes, 2008). However many other domains exist and nowadays, databases such as the Database for Arabidopsis Transcription Factors (DATF) (<http://atrm.cbi.pku.edu.cn/with>) which contains information known and predicted TFs in Arabidopsis.

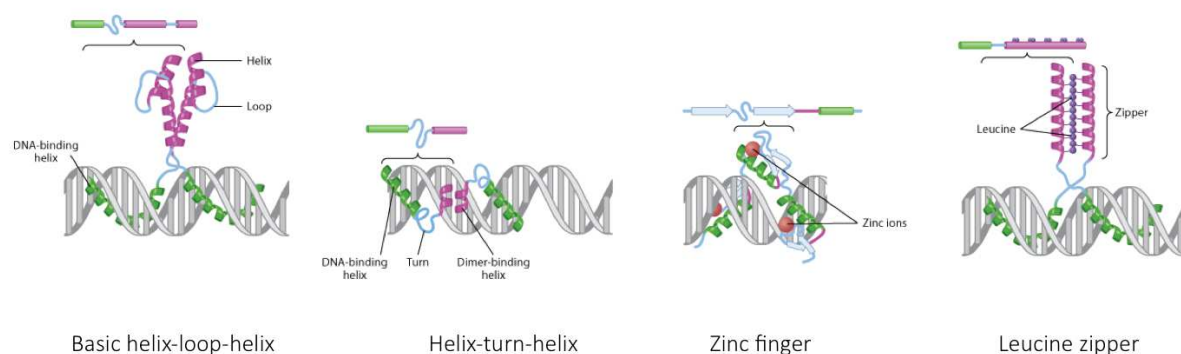


Figure 13. Common DNA-binding motifs.

The primary structure of each motif is shown above a representation of the structure interacting with DNA.

A. Pioneer transcription factors (PTFs)

As explained before, DNA is carefully packed into nucleosomes and further condensed by linker histone to stabilize this compaction. This state is refractory for transcription machinery. Nonetheless, chromatin has to open in regulatory regions, so that RNA polymerase can perform its function. How TFs access to these compacted regions? In the last two decades, a sub set of transcription factors, called Pioneer Transcription Factors (PTFs) that can bind to compacted regions of chromatin have been, and continue to be studied.

Genome-wide assays such as ChIP-seq, *in vitro* chromatin reconstitution as well as genetic studies have generated data that revealed the distinct ability of these PTFs to target their binding sites on nucleosomal DNA. Meaning, they are able to overcome the steric constraints that nucleosomes provide. They are in consequence able to establish a “competent state” in a particular region, so it can be further regulated by other partners, as illustrated in Figure 14. (Iwafuchi-Doi & Zaret, 2014), such as additional TFs, CRCs, among others. It is however important to mention that CRCs do not possess DNA-binding specificity like a TF does.

ATP-dependent CRCs behave in a similar way, unwinding nucleosomal regions so they can be accessible to general transcription factors. But remodelers’ activity on the genome seems to be concentrated at regulatory sequences, where they can be recruited by transcription factors. Plus, remodeling complexes are large complexes compared to nucleosomes and use one ATP per base pair of “remodeled” DNA, which represents a high energy cost (Clapier & Cairns, 2009). Therefore, PTFs act in a complementary mode of action with CRCs to open closed regions of chromatin.

In contrast to active chromatin, heterochromatin has revealed to be refractory to pioneer transcription factors. Such closed domains have repressive covalent modifications of the core histones including H3K9 or H3K27 methylation. (Becker, Nicetto, & Zaret, 2016; Ho et al., 2014). The molecular mechanism behind the blocking for PTFs’ binding heterochromatin remains an open question.

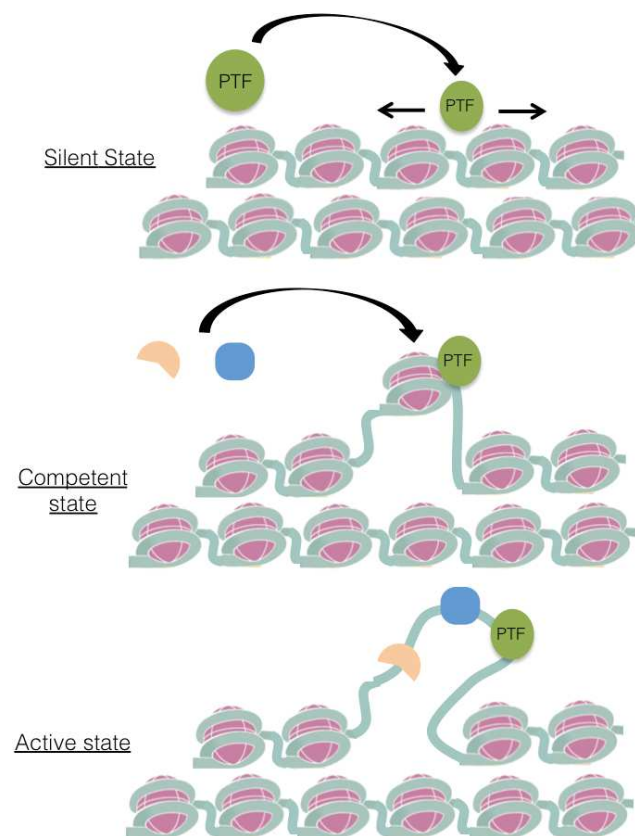


Figure 14. Model of Pioneer Transcription Factors mechanism.

(Top) Pioneer factors can scan nucleosomal DNA for their target sites in silent chromatin. (Middle) This allows for cooperative interactions with other factors (indicated in blue and beige) and more stable binding. (Bottom) Binding of PTFs can lead to activated sequences with open chromatin features (Zaret & Mango, 2016).

- Examples of PTFs and their molecular mechanism

The first described pioneer factors were FoxA (Forkhead box) proteins, steroid receptors observed to occupy liver-specific enhancers prior to hepatic induction (Zook et al., 2016). Other main PTFs are briefly described in Table 3 adapted from (Iwafuchi-Doi & Zaret, 2014).

Table 3. Predicted/Validated Pioneer Transcription Factors (From Iwafuchi-Doi & Zaret, 2014).

Pioneer Factors	Cellular/Biochemical context (model organism)	Predicted/Validated pioneer activity	References
FoxA	Trans differentiation from fibroblast to iHep (mouse) Liver/foregut development (mouse/ <i>Caenorhabditis elegans</i>) <i>In vitro</i> reconstituted nucleosome or nucleosome array	Inducing trans differentiation Establishment of the competence for liver development Binding and increasing DNA accessibility in compacted nucleosome array independent from other factors <i>in vitro</i>	Huang et al. 2011; Sekiya and Suzuki 2011 Lee et al. 2005
Class V Pou	Zygotic genome activation (zebrafish/mouse)	Inducing zygotic genome activation	Foygel et al. 2008; Lee et al. 2013
Klf4	Reprogramming from fibroblast to iPSCs (mouse/human)	Inducing reprogramming Binding to chromatin that is non-hypersensitive and not pre marked by histone modifications <i>in vivo</i>	Takahashi and Yamanaka 2006 Soufi et al. 2012
Ascl1	Trans differentiation from fibroblast/hepatocyte to iN cells (mouse/human)	Inducing trans differentiation Binding to chromatin <i>in vivo</i>	Vierbuchen et al. 2010
Pax7	Pituitary melanotrope development	Establishment of the competence for melanotrope development	Budry et al. 2012; Drouin 2014
GATA1	Mitotic bookmarking (mouse)	Binding chromatin mitotic	Kadauke et al. 2012

As the paradigm pioneer transcription factor, FoxA has been studied extensively. *In vivo*, genetic, and *in vitro* biochemical studies indicate that FoxA can bind on one side of the DNA helix, leaving the other side available to bind core histones. Also, its binding displaces linker histones from the local chromatin and keeps nucleosomes accessible (Cirillo et al., 1998; Iwafuchi-Doi et al., 2016). This is because FoxA's DBD (DNA-binding domain) is structurally very similar to the winged-helix DBD of linker histone. This domain is surrounded by two basic disordered N- and C- terminal domains, which acquire secondary structures when bound to DNA.

Furthermore, the C-terminal domain of FoxA can bind directly to core histone proteins without ATP-dependent chromatin remodelers and is required for loosening nucleosomal arrays. *In vitro* and footprinting assays showed that HNF3 (Hepatocyte Nuclear Factor 3), a homologue of FoxA, can directly bind to H3 and H4 via its C-terminal domain and this core-histone binding is essential for nucleosome loosening (Cirillo et al., 2002).

Another example of PTF is the early B-cell factor 1 (EBF1), a key transcriptional determinant of B-lymphocyte differentiation. Its unique DBD adopts a pseudoimmunoglobulin-like fold when bound as a dimer to its palindromic recognition site (Treiber et al., 2010). Boller et al., (2016), recently showed that EBF1 transcription factor, has the ability to create DNase I hypersensitive sites. Also, it was demonstrated that its C-terminal domain is required to access chromatin. This illustrates, how the EBF1 C-terminal domain, which lacks a predicted secondary structure (Figure 15), can have an effect on the TF's binding to chromatin even if it does not directly interact with DNA.

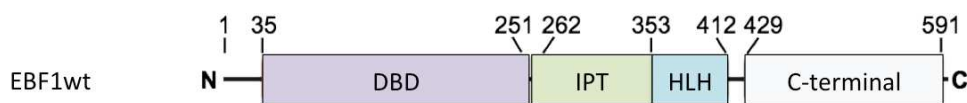


Figure 15. Schematic domain representation of EBF1.

Four main domains of EBF1, the DNA-binding domain (DBD), the Immunoglobulin-like fold, Plexins, Transcription factors (IPT) domain, the helix-loop-helix (HLH) domain, and the C-terminal domain with amino acids positions indicated.

Interestingly, Soufi et al., (2015), studied a group of potential TFs that together reprogram somatic cells to pluripotency. These are: octamer-binding transcription factor 4 (Oct4), (sex determining region Y)-box 2 (Sox2), krueppel-like factor 4 (Klf4) and myc proto-oncogene (c-Myc). This group was analyzed in an integrative way by three approaches to assess their interactions with nucleosomes: Biochemical assays, genomic analysis (in the context of ectopic expression of PTFs) and structural analysis. The study provided interesting insights of PTFs activity. Oct4, Sox2 and Klf4 but not c-Myc, which contain the structural DBDs: Pit-Oct-Unc (POU), High Mobility Group (HMG), Zinc fingers (ZF) and basic-Helix-loop-Helix (bHLH), respectively, were shown to have a pioneer function in cell reprogramming. Their results also show that among a group of transcription factors important for cell-fate changes,

there can be a synergistic behavior of pioneer (Oct4, Sox2 and Klf4) with non-pioneer (c-Myc) factors. A common feature of these PTFs is the ability of their respective DBDs to adapt and bind to DNA on a nucleosome surface (Figure 16).

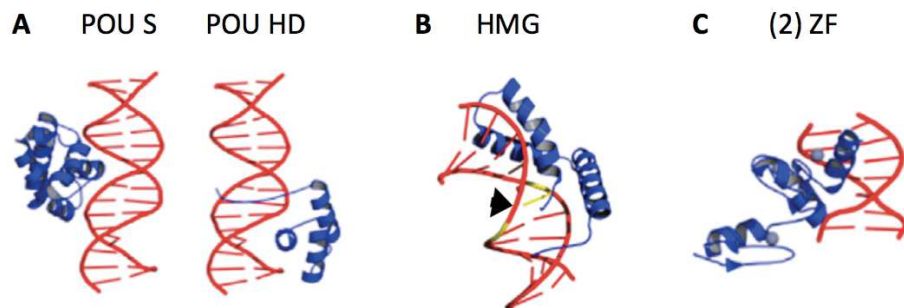


Figure 16. Schematic representations of the 3D DBDs of Oct4, Sox2 and Klf4 interacting with DNA.

A | Oct4 contains a POU domain composed of an N-terminal POU-specific (POUS) and a C-terminal POU-homeodomain (POUHD), separated by a linker region. The X-ray structure of Oct4-POU-DNA complex confirms that the POUS and POUHD each bind one-half of the octameric motif on DNA (Esch et al., 2013). The two POU domains do not target directly adjacent half sites on nucleosomes, as seen in free DNA, but the exposure of the separate half sites on nucleosomes is enough for Oct4 initial targeting. B | Sox2 binds DNA through its HMG box, inducing a sharp bend and widening of the minor groove. C | Klf4 uses its two most C-terminal ZFs, out of three, to recognize its hexameric DND binding motif, occupying one side of the DNA double helix and leaving more than half of the opposite surface potentially free to interact with histones in a nucleosome (Soufi et al., 2015).

In this section, I explained how chromatin is organized and compacted, the nucleosome, as the basic unit of compaction, as well as the mechanisms and proteins involved in the regulation of this dynamic structure. Now, I will describe the hypothesis about LFY's ability to trigger changes in chromatin organization that are required for floral initiation and floral organ development.

III. LEAFY AS A POTENTIAL PIONEER TRANSCRIPTION FACTOR

This section will focus on LFY as a potential plant-specific pioneer transcription factor (PTF). To date, there has not been any PTF confirmed in the plant kingdom. As introduced in section I, LFY is a master regulator of floral development. It integrates external as well as internal cues and is also able to coordinate the floral network (Posé, Yant, & Schmid, 2012).

LFY is expressed at low levels in vegetative tissue but it is upregulated in response to floral inductive signals. Afterwards, LFY triggers the emergence of a group of stem cells, the floral meristem (FM), on the flanks of the shoot apex. Later, it specifies their floral identity by inducing floral homeotic genes of the MADS box family, in charge of specifying flower organ identities. Moreover, one of the main roles of LFY is to bring the important and precise balance between the proliferative stem cell activity of the floral meristem and its eventual termination to form the determinate flower (Irish, 2010). Overall, because it responds to converging upstream signals, and is fundamental to bring a flowering response it has been classified as a floral pathway integrator (Simpson & Dean, 2002).

1. A MASTER FUNCTION IN A GENETIC CONTEXT

As mentioned before, PTFs are transcriptional regulators that have the ability to engage with silent chromatin, increase target site accessibility and cooperative association to factors to bind (Guo & Morris, 2017). These DNA-binding characteristics go in hand with their ability to initiate developmental programs. Several, have been shown to be fundamental in distinct developmental contexts. For example, the paired box protein 7 (Pax7) transcription factor, in pituitary melanotrope development and the hematopoietic TF, PU.1, for myeloid and lymphoid development (Iwafuchi-Doi & Zaret, 2014). Remarkably, the most studied pioneer transcription factor FoxA, involved in liver induction, has been used together with other transcription factors, in the most impressive example of cellular reprogramming: the conversion of differentiated cells into induced pluripotent stem cells (Takahashi & Yamanaka, 2006).

- The particular case of LFY

For LFY, several experiments are particularly illustrative of its crucial function in flower development. As reported previously, *lfy* mutant flowers display floral homeotic defects. The *lfy* phenotype is flower to shoot conversion (Weigel et al., 1992). On the contrary, when LFY is expressed from a constitutive promoter, plants flower earlier than the wild type and ectopic flowers also appear in the axils of rosette leaves where axillary shoots normally arise in the wild type (Weigel, Detlef & Nilsson, 1995). Besides, when a modified form of the F-Box

protein UFO is co-expressed, it can trigger in a LFY-dependent manner, the formation of flowers from leaf tissue (Risseuw et al., 2013). Also, it is able to induce flowers from roots, together with co-regulators such as the homeodomain-TF, WUS (Gallois et al., 2004). Finally, LFY activates floral MADS genes expression such as *AG* and *AP3*, which are normally repressed by the PRC2 complex in Arabidopsis. The PRC2 complex contributes to chromatin compaction by tri-methylation of histone H3 at lysine 27 (Engelhorn et al., 2014).

As mentioned above, it is clear nowadays that cell type determination is established by a collection of regulatory proteins rather a single one. In relation to this, it has been suggested that SEP3 and AP1 TFs play a pioneer role during flower development (Pajoro et al., 2014; Smaczniak et al., 2011).

2. A UNIQUE STRUCTURE

The LFY protein displays unique features. It is present in all land plants but exerts a unique function in angiosperms, even if a gene network involving LFY and central to the control of floral development may have already existed before the appearance of flowers (Moyroud et al., 2017). In other species, its role is much less understood. LFY is composed of approximately of 400 amino acids, depending on the species and it contains two conserved domains connected by a highly variable and less structured linker domain (Figure 17).

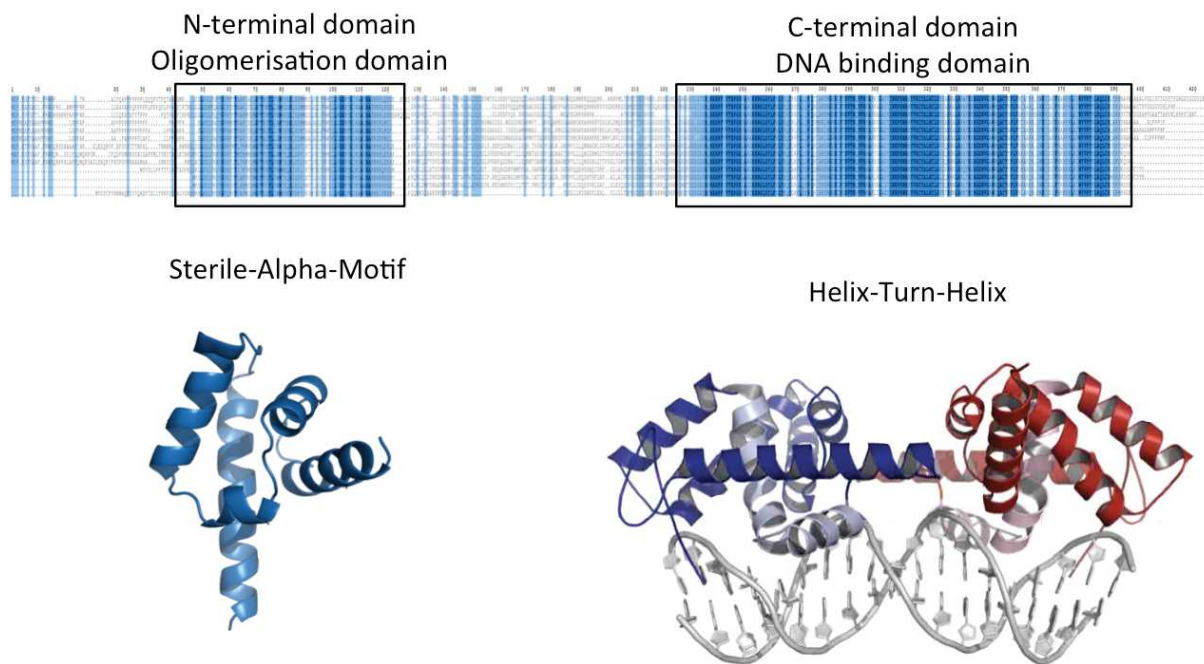


Figure 17. LFY contains two conserved domains.

On top, sequence alignment providing from four angiosperm species (*A. thaliana*, *P. hybrida*, *A. trichopoda* and *O. sativa*), three gymnosperm species (*G. biloba*, *P. radiata* and *W. mirabilis*), two ferns (*C. richardii* and *A. lygodiifolia*) and the moss *P. patens*. The different blue colors indicate the degree of amino acid conservation. The darkest blue represent identical amino acids among all species. At the bottom, the crystal structures of the two conserved domains are observed. First, a sterile alpha motif (SAM) domain in the N-terminal domain, which is essential for LFY floral function and allows oligomerization on DNA then, a helix-turn-helix (HTH) DNA-binding domain (DBD) that binds DNA as a dimer.

- The DBD domain of LFY

The LFY C-terminal DNA-binding domain was crystallized with a DNA sequence taken from the regulatory region of the *AP1* or *AG* genes. It was demonstrated that LFY binds as a dimer, and recognizes semi-palindromic DNA sites of 19 bp, formed by two inverted “half sites” (Figure 18).

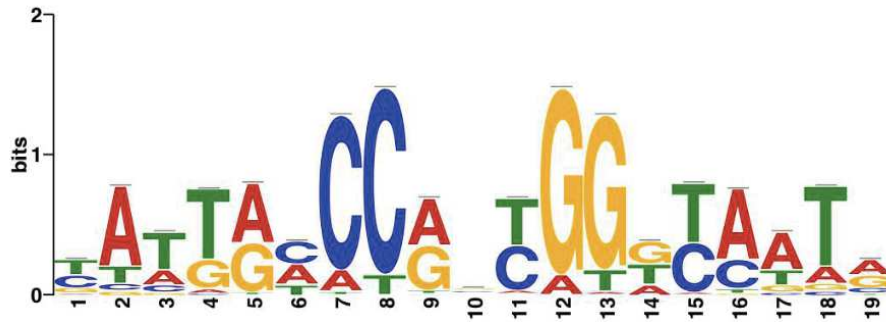


Figure 18. Logo representing the 19-bp motif targeted by LFY's DBD.

The number of bits of each base shows the total information at a given position and within this, the height is the base frequency (Moyroud et al., 2011).

In solution, LFY's C-terminal domain is a monomer and dimerization takes place when DNA is present. Once a monomer binds DNA, the binding of the second one is facilitated due to an interaction between the two monomers through a histidine and an arginine residues (Hamès et al., 2008). This dimerization interface between two DBDs is different from the oligomerization mediated by the N-terminal domain, which is described in the next section.

- The SAM domain, an oligomerization domain

This particular domain is important for LFY's function since it is essential for a correct flower development. Crystallization of the N-terminal domain of LFY indicates that it forms a head-to-tail helicoidal structure as shown in Figure 19 (Sayou et al., 2016). The residues that enable this interaction are well conserved through diverse species, from algae to angiosperms, indicating that oligomerization must have exerted an important function throughout evolution.

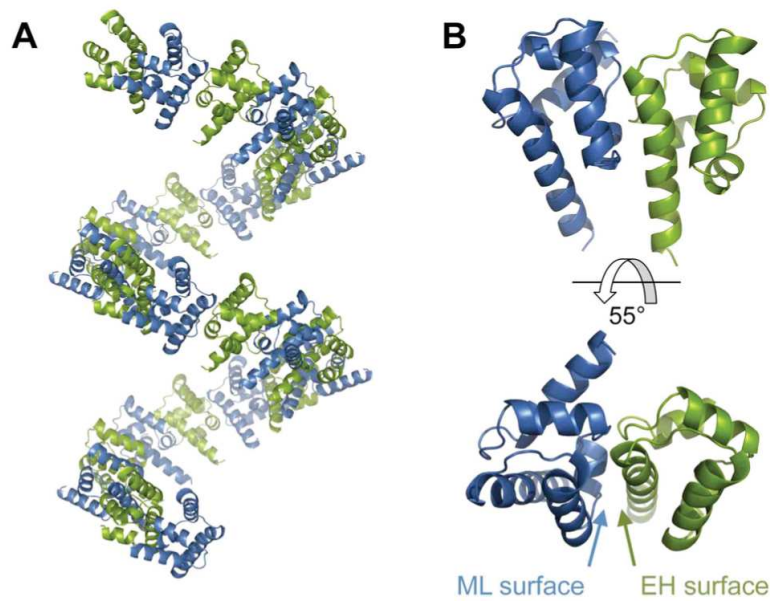


Figure 19. Oligomerization of the N-terminal LFY crystal structure from the gymnosperm *Ginkgo biloba*.

A | SAM monomers make contact via two polar surfaces making a helical polymer through a mid-loop (ML) and an end-helix (EH) surfaces. B | On the top, a side view and on the bottom, a top view of a dimer (Sayou et al., 2016).

The resolution of the crystallographic structure allowed the identification of the residues involved in the interaction between each monomer. From this, a mutated version of LFY, LFY_{TERE}, unable to oligomerize, was designed. The LFY_{TERE} protein consists in a substitution of amino acids at both, the ML and the EH surface, with the objective of disrupting the interaction. Two genetic experiments prove the relevance of this domain in flower development. First, complementation of the weak allele *lfy-12*, which presents flower replacement by shoots or sterile shoot/flower intermediate structures, is much less efficient with LFY_{TERE} compared to wild type LFY. Second, in p35S:LFY_{TERE} plants, where LFY_{TERE} is expressed under the control of the 35S constitutive promoter, no early flowering and no ectopic flowers were observed compared to p35S:LFY plants (Figure 20), (Sayou et al., 2016).

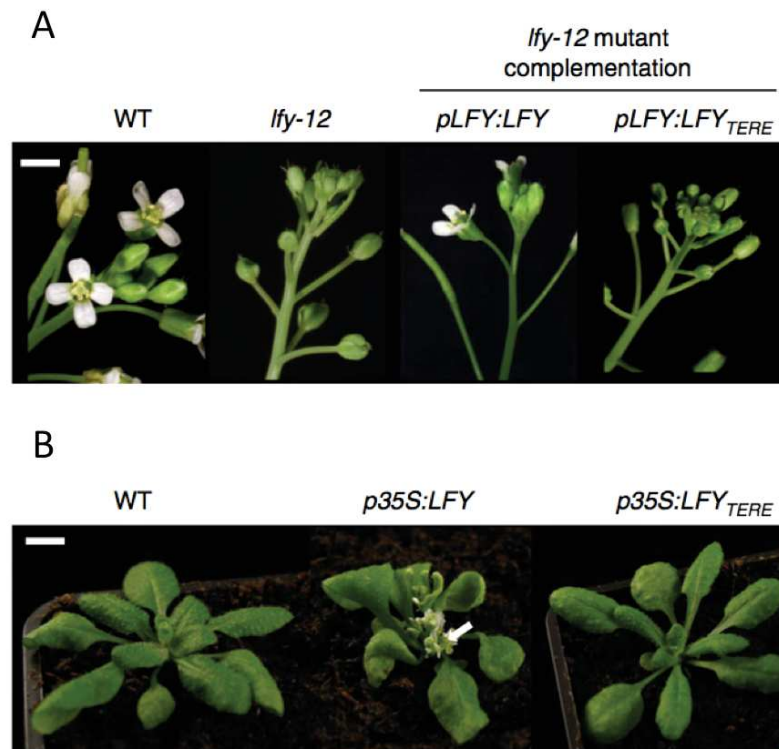


Figure 20. The LFY-SAM domain is essential for proper flower development in Arabidopsis.

A | Wild type (WT), *lfy-12* mutant and *lfy-12* expressing pLFY:LFY or pLFY:LFY_{TERE} inflorescences from 40-day-old plants grown under long day conditions (scale bar, 2 mm). B | WT and transgenic plants expressing p35S:LFY or p35S:LFY_{TERE} grown for 30 days under long-day conditions. Scale bar, 1 cm (Sayou et al., 2016).

3. GENOME-WIDE INSIGHTS OF LFY'S PIONEER ROLE

Recent work from our group showed interesting results obtained by a combination of structural, biochemical and genomic approaches that shed light on the function of LFY SAM oligomerization domain's function. Here, I pay special attention to the ability to access closed chromatin regions that this domain confers to LFY. Interestingly, a transcription factor's domain that does not make contact to DNA directly has an important effect on its DNA binding at a genome-wide scale (Sayou et al., 2016).

Chromatin immunoprecipitation coupled to high-throughput sequencing (ChIP-Seq) assays, allows to determine binding sites of a transcription factor *in vivo*. ChIP-Seq data were obtained on 2-week-old Arabidopsis seedlings from a) p35S:LFY transgenic plants expressing LFY (AtLFY) from the p35S constitutive promoter, b) p35S:LFY_{TERE} plants overexpressing the mutant version of LFY protein unable to oligomerize (AtLFY_{TERE}), and c) non-transgenic

Arabidopsis (Col-0) which served as a negative control, since at this stage LFY endogenous expression is negligible. To identify LFY-binding sites (LFYbs), a previously validated position weight matrix that computes a score between 0 (highest affinity site) and -56 (lowest affinity site) was used (Moyroud et al., 2011).

When comparing the intensity of LFY binding in p35S:LFY and p35S:LFY_{TERE} bound regions, it was found that the TERE mutation drastically reduced the LFY binding *in vivo* especially in regions displaying multiple low-affinity LFY binding sites. The score to determine a lost in LFY binding was defined as -25. However, the lost of LFY_{TERE} binding was at a score of -20 (Sayou et al., 2016). As a reference, the score for the *AP1* target is -7. Therefore, the difference in these two values suggests that oligomerization promotes binding to sites with low-affinity. In Figure 21, three relevant genomic regions of LFY targets are shown for illustration purposes.

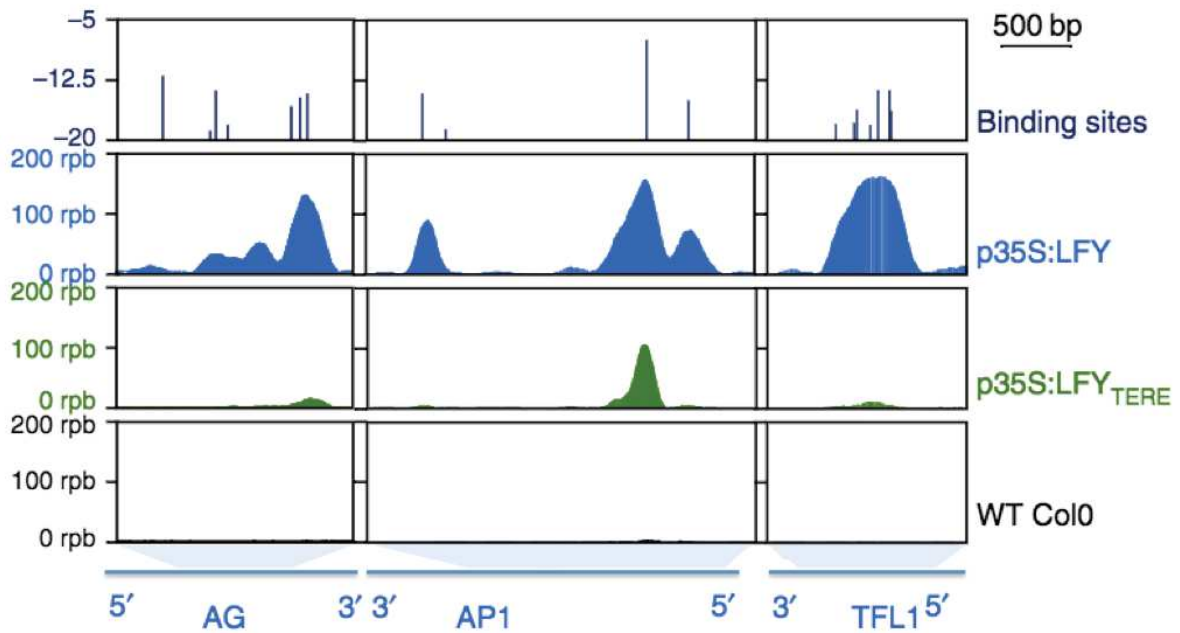


Figure 21. Impaired LFY oligomerization affects genome-wide LFY DNA-binding *in vivo*.

AtLFY and AtLFY_{TERE} binding profiles on two floral homeotic genes induced by LFY, *AG* and *AP1* and one meristem identity gene repressed by LFY, *TERMINAL FLOWER 1 (TFL1)*. On the top, position in the genome of theoretical LFY binding sites and their respective score on the y-axis. Below, ChIP-Seq read coverage, defined as the area above the read number curve, for p35S:LFY (blue), p35S:LFY_{TERE} (green) and Col-0 (black, background signal).

On the y-axis, reads per base (rpb). LFY_{TERE} displays an overall diminished DNA binding, nevertheless the extent of this reduction in binding can largely fluctuate in different genomic regions (Sayou et al., 2016).

A parameter used in all further analyses is the coverage fold reduction (CFR), defined as the ratio between the LFY and the LFY_{TERE} Chip-Seq coverage.

$$\text{CFR} = \text{coverage}(\text{p35S:LFY}) / \text{coverage}(\text{p35S:LFY}_{\text{TERE}})$$

It was shown that the bound regions with low-affinity LFYbs have in general a higher CFR than regions with high-affinity LFYbs. Also, it was found that regions with high LFYbs density have higher CFR than regions with less LFYbs. Therefore, these results suggest that a functional oligomerization domain helps LFY bind to DNA regions with either sub-optimal LFYbs or clusters of LFYbs. Furthermore, binding sites spaced only by 1 bp were overrepresented in regions with high CFR. Accordingly, regions with one or more sites spaced by 1 bp significantly increased the CFR value. Collectively, these evidences indicate that LFY's oligomerization domain helps binding suboptimal-binding affinity sites or regions that have multiple and adjacent sites (Sayou et al., 2016).

- LFY's oligomerization domain allows access to closed chromatin regions.

The accessibility of a given DNA region can be assessed by DNaseI hypersensitivity (DHS) assays coupled to high-throughput sequencing (DNaseI-Seq) assays. These data were used to test whether there was a link between the CFR and the accessibility of DNA regions. In the low CFR regions, only 10–15% of the LFY-bound regions were closed. Remarkably, this number reached 70% in regions with high CFR (Figure 22), this suggests that genomic regions where LFY DNA-binding is strongly affected if LFY cannot oligomerize, are largely closed genomic regions (Sayou et al., 2016).

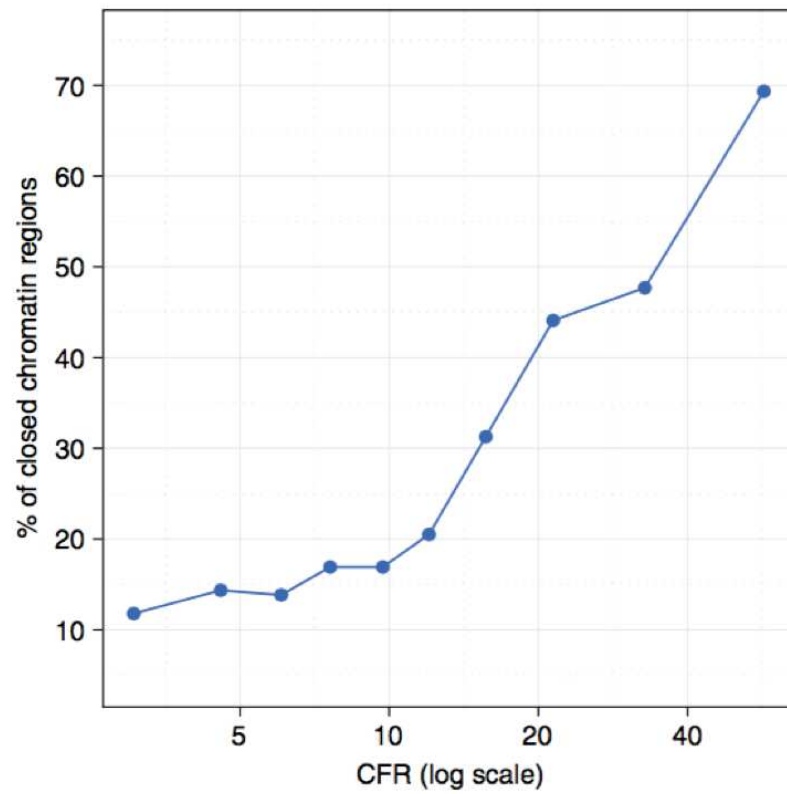


Figure 22. Relation of chromatin accessibility to genome-wide LFY DNA-binding.

Percentage of closed regions in deciles of p35S:LFY-bound regions sorted according to their CFR (Sayou et al., 2016).

In addition, as shown in Figure 23A, the CFR values are remarkably lower when the accessibility is higher. Similarly, regions with hampered accessibility have higher CFR values than accessible ones (Figure 23).

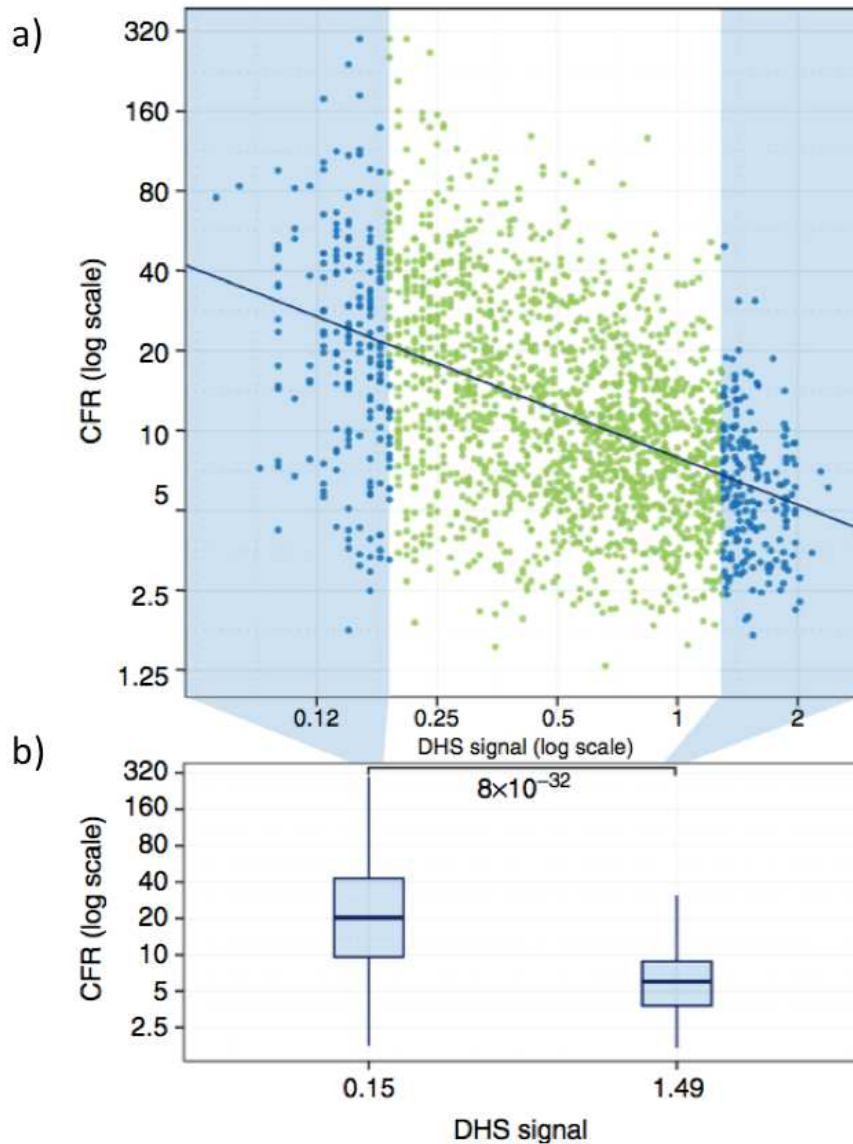


Figure 23. Genome-wide LFY DNA-binding in relation to chromatin state.

A | Comparison of the CFR of the LFY-bound regions in p35S :LFY and the chromatin accessibility in 15-day-old seedlings (DHS signal). Data are shown in light blue and green, and fitted values are shown in dark blue. B | Boxplot representing the CFR signal in the regions of highest and lowest DHS (first and last deciles). P-value for Mann–Whitney rank test on the two extreme groups is indicated on the graph (Sayou et al., 2016).

As mentioned in (Sayou et al., 2016), “the chromatin state of the bound regions has a strong impact on the need for LFY’s oligomerization”. These results, coupled with another DNase-Seq data set from 7-day-old seedlings that yielded similar results (Sullivan et al., 2014), provide strong evidence that this oligomerization domain is essential for genome-wide DNA binding *in vivo* of this transcription factor.

In brief, the previous results suggest that oligomerization can confer to LFY the ability to bind closed chromatin regions, which it could not bind without its SAM domain. Since one main characteristic of these particular regions is to be low-affinity and high density in binding sites, we could hypothesize that LFY might exert a mechanical effect on nucleosomes, via oligomerization, favoring binding to nucleosome-enriched regions. Consistent with this, the crystal structure of the complex LFY C-terminal and DNA shows how the binding of several dimers can only take place on one side of the double helix (Figure 24).

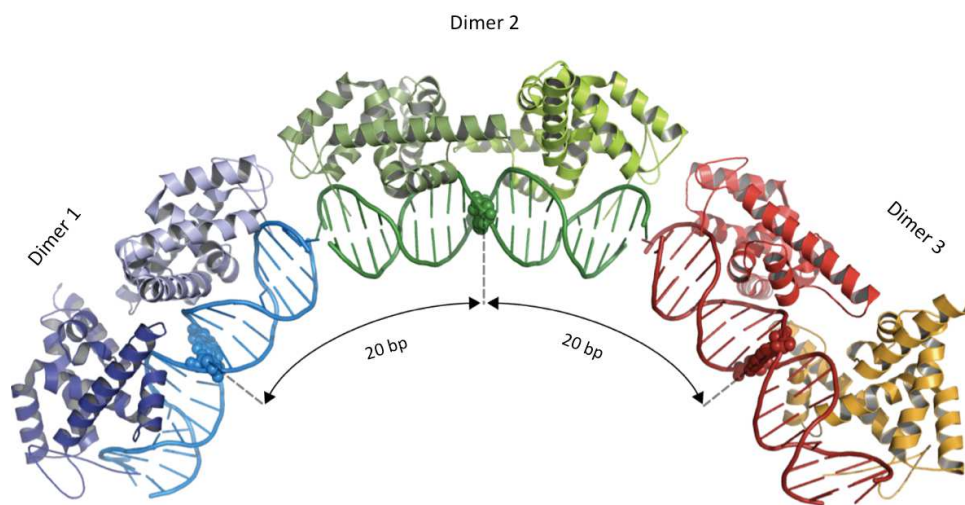


Figure 24. Crystal arrangement of LFY DBD on DNA.

The central base pair of each binding site is indicated by spheres (Hamès et al., 2008).

This kind of binding could represent a “pioneer feature” of LFY. This is because the binding takes place on only one side of the DNA double strand, and therefore does not hinder the other side to interact with core histones (See Figure 16). Other TFs bind that bind DNA in a tweezer-like manner are not good candidates to be PTFs. For instance, this LFY’s feature has been observed in the recently identified GAGA-associated factor (GAF) pioneer transcription factor, which plays a key role in early *Drosophila* development. It contains a Pox virus and Zinc finger (POZ) domain that allows oligomerization and which binding leads to DNA bending observed by electron microscopy (Moshe & Kaplan, 2017).

It is moreover compelling that LFY has been shown to interact with two chromatin-remodeling complexes (Wu et al., 2012) that, after being recruited to specific closed loci, could engage their decompaction.

4. LFY AND CHROMATIN REMODELING COMPLEXES

The interaction between LFY and the two catalytic ATPase-subunits of ATP-dependent CRCs, SYD and BRM has been mostly studied in relation to the regulation of homeotic genes *AP3* and *AG*. These two genes are LFY targets that specify the development of floral organs in whorls 2, 3 for *AP3* and whorl 4 for *AG*. It was observed that SYD and BRM have partially overlapping roles in flower patterning. The cross of the two null mutants *syd-2* (Wagner & Meyerowitz, 2002) or *brm-101* (Kwon, 2006) to the weak *lfy-5* allele displayed, both, strongly enhanced floral homeotic defects compared to *lfy-5* (Figure 25A).

It was further observed that these homeotic defects were caused by a failure of *AP3* and *AG* induction, since *in situ* hybridization experiments showed a drastic reduction in *AP3* and *AG* expression compared to the *lfy-5* mutant (Figure 25B) (Wu et al., 2012). A similar outcome was observed in a conditional *syd brm* double mutant in flower primordial. This was obtained by designing a microRNA (aMIR) directed against BRM which phenocopied the *brm* null mutant. When aMIR is expressed under the *LFY* promoter in the *syd-2* background (*syd-2 pLFY:aMIRBRM*), flower primordial greatly reduced *AG* and *AP3* expression similar to *syd-2 lfy-5* and *brm-101 lfy-5* plants, (Figure 25B). Impressively, when *AG* was constitutively expressed in these mutants, stamen and carpel identity defects were rescued (Wu et al., 2012).

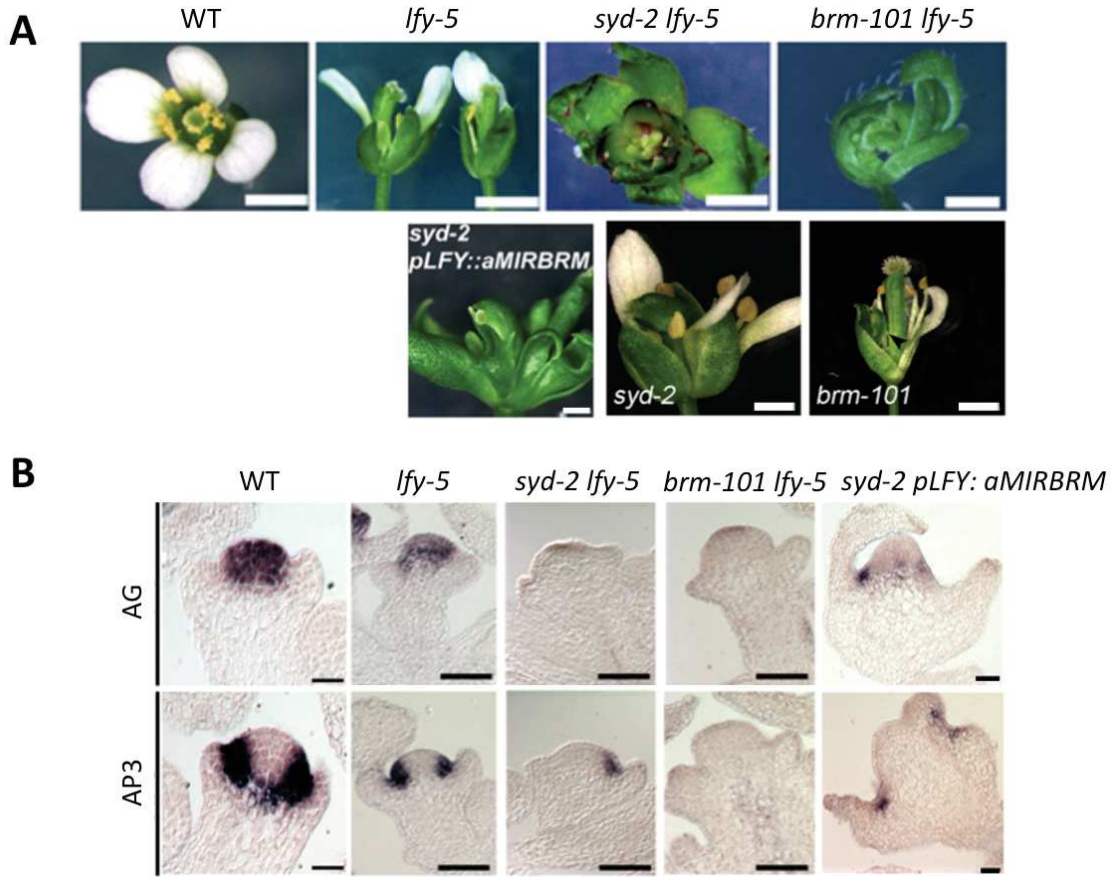


Figure 25. *brm* and *syd* mutants strongly enhance the floral homeotic defects observed in *lfy-5* mutant.

A | Floral homeotic defects in *syd-2 lfy-5* and *brm-101 lfy-5* are enhanced compared to *lfy-5*. In a similar way those defects are enhanced in *syd-2 pLFY::aMIRBRM* conditional double mutant. (Scale bars: 1mm) B | *In situ* hybridization of *AP3* and *AG* expression in stage 3 flowers of *lfy-5*, *syd-2 lfy-5*, *brm-101 lfy-5* and *syd-2 pLFY::aMIRBRM* mutants. Expression is strongly reduced in all mutants. (Scale bars: 30 μm) (Wu et al., 2012).

Interestingly, LFY, SYD and BRM strongly and specifically associate with the *AP3* and *AG* regulatory regions during flower patterning. Occupancy by LFY, SYD and BRM was assessed by ChIP assays before and during flower patterning using a semi synchronized flower induction system (*ap1 cal p35S:AP1-GR*). LFY, SYD and BRM, specifically associated to the *AP3* and *AG* loci when flower patterning was initiated (Wu et al., 2012).

Further ChIP experiments showed that, relative to the wild type, SYD association to *AP3* and *AG* loci was strongly reduced in *lfy-5* mutants. On the contrary, LFY association to the two loci was not dependent on SYD or BRM (Figure 26A), suggesting a recruitment mechanism of SYD and BRM, by LFY. Finally, physical interactions were assessed by pull-down assays with SYD and BRM N-terminal domains (SYDN and BRMN) and the recombinant LFY

protein fused by the C-terminus to the glutathione-S-transferase (GST) protein, (GST-LFY). Experiments revealed that SYDN and BRMN interacted with LFY (Figure 26B). Also, in yeast-two hybrid (Y2H) assays, the N-terminal regions of SYD or BRM weakly interact with LFY (Figure 26C). Bimolecular Fluorescence Complementation (BiFC), also confirmed the mentioned interactions (Figure 26D).

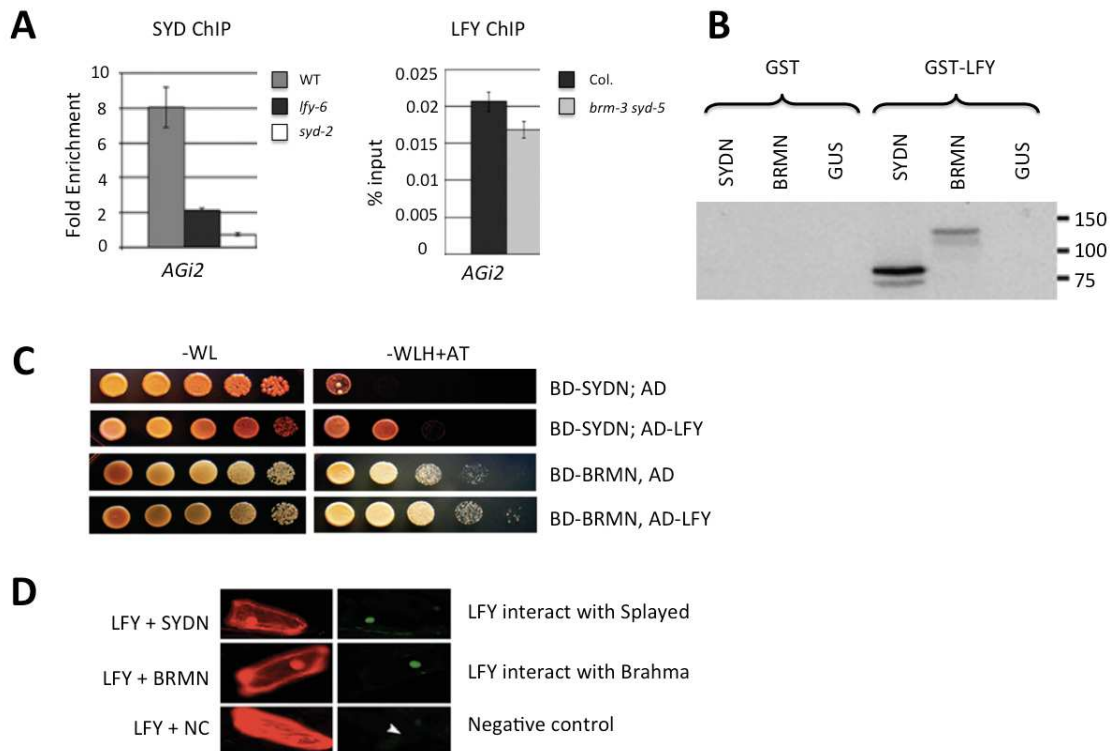


Figure 26. Interactions between LFY, SYD and BRM.

A | LFY binding to *AG* regulatory region. Anti-LFY ChIP-PCR in wild type and *brm-3 syd-5* inflorescences normalized over LFY levels in the two genotypes. B | GST-LFY pull-down assay showing physical interaction between LFY and the [35S]methionine-labeled SYDN and BRMN. On the right, approximate molecular mass in kilodaltons. C | The N-terminal regions of SYD and BRM (SYDN or BRMN) interact with LFY. On the left, growth on media lacking tryptophan and leucine (-WL, not selecting for interaction). On the right, growth on media that selects for interaction by further omitting histidine and adding 0.1 mM 3-AT (a competitive inhibitor of the *HIS3* gene product used as a reporter gene) for SYD or 0.5 mM 3-AT for BRM. BD, binding domain. AD, activation domain. Interaction of LFY with SYD and BRM is weak since it is close to the background signal. D | *In planta* interactions between LFY and N-terminal domains of SYD or BRM (SYDN or BRMN) by bimolecular fluorescence complementation (BiFC). On the left, transformed cells expressing p35S:2xmCherry. On the right, yellow fluorescent protein (YFP) fluorescence in the same cells (Wu et al., 2012).

Additionally, the role of PcG proteins in the regulation of *AP3* and *AG* was studied, and evidences suggest that SYD and BRM act antagonistically to PcG proteins. *AP3* and *AG* regulatory regions are known to be repressed via PRC2/PRC1 complexes (Zheng & Chen, 2011). PRC2 causes the trimethylation of lysine 27 of histone 3 (H3K37me3) of these loci. Then, PRC1 is recruited to these regions to further stabilize this repression.

To assess their antagonistic role in respect to SYD and BRM, a mutant of a methyltransferase subunit of the PRC2 complex, CURLY LEAF (CLF) was used. The *clf-2* null mutant displays curled leaves due to ectopic expression of *AG* and *AP3* (Chanvivattana, 2004; Goodrich et al., 1997). While *syd-2 pLFY:aMIRBRM* plants display floral homeotic defects as described above, *clf-2 syd-2 pLFY:aMIRBRM* plants showed a rescue phenotype. Also, *clf-2* caused a very strong increase of *AG* and *AP3* expression compared to the double mutant *syd-2 pLFY:aMIRBRM*, observed through qRT-PCR (Figure 27A). In addition, *clf-2 syd-2* double mutants, showed a reduction a 70% of *AP3* and *AG* expression compared to *clf-2* (Figure 27B). Similar results were obtained with BRM. Thus, these results suggest that CLF and SYD/BRM have antagonistic effects at the regulatory regions of the common target genes *AP3* and *AG* (Wu et al., 2012).

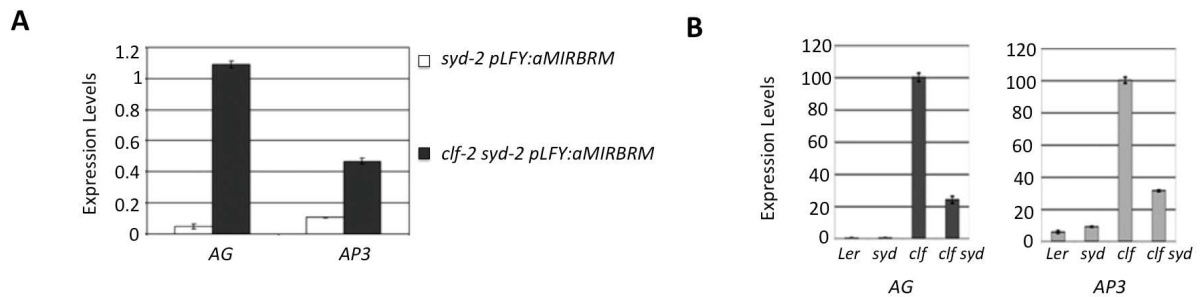


Figure 27. Antagonistic effects of CLF and BRM/SYD in *AP3* and *AG* regulatory regions.

A | qRT-PCR of *AG* and *AP3* expression in 28-days-old *syd-2 pLFY:aMIRBRM* and *clf-2 syd-2 pLFY:aMIRBRM* inflorescences. In the triple mutant, where CLF is not present to repress both loci, *AG* and *AP3* expression is greatly recovered. B | qRT-PCR of *AG* and *AP3* expression in 13-day-old wild-type (Ler), *clf-2*, *syd-2*, and *clf-2 syd-2* seedlings. The antagonistic effect of mutations, on one hand, of the PcG protein CLF and on the other hand, the TrxG protein, SYD are observed based on expression levels of *AG* and *AP3* (Wu et al., 2012).

Even though the exact molecular mechanism by which polycomb repression is overcome at *AP3* and *AG* loci is still not fully understood, the combined results suggest a mechanism in which LFY might be able to activate its target genes *AP3* and *AG* by recruiting

SYD and BRM, and that these two last, act antagonistically to PcG proteins in order to overcome the repressive state of the loci. Certainly, other TrxG proteins participate at the same time, or later in this process such as the H3K4me3 methyltransferase ARABIDOPSIS HOMOLOG OF TRITHORAX1 (ATX1), the chromatin remodeler PICKLE (PKL) as well as the TrxG protein, ULTRAPETALA1 (ULT1) (Aichinger et al., 2009; Carles & Fletcher, 2009; Saleh et al., 2007).

OBJECTIVES

LEAFY is a master regulator for early stages of flower development. This transcription factor has two conserved domains across species. The first one is a Sterile Alpha Motif (SAM) domain in the N-terminus, which confers LFY the ability to oligomerize across the DNA chain and which has been recently proved to be essential for its floral function. Plus, combined *in vitro* assays and genome-wide analyses, have demonstrated that oligomerization is key for LFY's DNA-binding because it grants access to regions with multiple or low affinity binding sites and also to closed chromatin regions. The second conserved domain is a Helix-turn-helix (HTH) DNA-binding-domain (DBD), which allows LFY to bind as dimer to only one face of the DNA chain. This makes it suitable to bind on nucleosomes.

In order to perform its function, LFY interacts with tissue-specific co-activators but there are other regulation mechanisms where LFY is involved, that are not yet well understood. One of these mechanisms is related to two CRCs from the SWI/SNF (Switch/Sucrose Non-Fermentable) family, SPLAYED (SYD) and BRAHMA (BRM). They possess an ATPase activity that allows them to destabilize histone-DNA interactions. LFY was shown to recruit these two previously mentioned CRCs to some regulatory regions (Wu et al., 2012) and could thus indirectly drive sequence-specific chromatin remodeling.

All these different pieces of information strongly suggest that LFY could act as a pioneer transcription factor. Therefore, the goal of my project was to use and integrate these evidences to unravel LFY's mode of action at the molecular level.

Chapter I is focused on LFY's ability to interact with the basic unit of compaction of chromatin, nucleosomes. The main questions that I tried to answer are:

- Is LFY able to directly bind nucleosomes *in vitro*?
- Is LFY able to specifically bind nucleosomes containing LFY target sites?
- What domains of LFY are implicated in this interaction?

In Chapter II, I was interested in assessing by different biophysical techniques the interactions between LFY and the two CRCs, SYD and BRM. Furthermore, through an *in vivo* approach, I studied to what extent LFY's function requires SYD and BRM.

The main questions that I tried to answer are:

- Which are LFY's domains able to interact with the CRCs, SYD and BRM?
- Which is the minimal domain of BRM and SYD that interacts with LFY?
- When ectopically expressed from the p35S constitutive promoter, does LFY require both of these CRCs to exert its function in plants?

CHAPTER I. LFY INTERACTION WITH NUCLEOSOMES

I. THE STUDY OF PIONEER TRANSCRIPTION FACTORS THROUGH EMSAs

In this chapter, I will discuss about the axis of research that I pursued during my PhD in order to assess a possible interaction between LFY and nucleosomes. *In vitro* nucleosome reconstitution is a way to obtain nucleosome particles with a given DNA sequence of interest. A valuable aspect of this approach for my work is that we were able to compare the binding of different LFY proteins. This approach has already been used in order to assess « pioneer » characteristics of other TFs.

In vitro nucleosome reconstitution is often verified by electrophoretic mobility shift assays (EMSA). This assay allows the visualization of the formation of protein-DNA complexes. Proteins are incubated with DNA. Either one or both are tagged with a fluorophore. Then, the mix is loaded into a polyacrylamide gel. The bigger the complex is, the slower its migration on the gel (Figure 28).

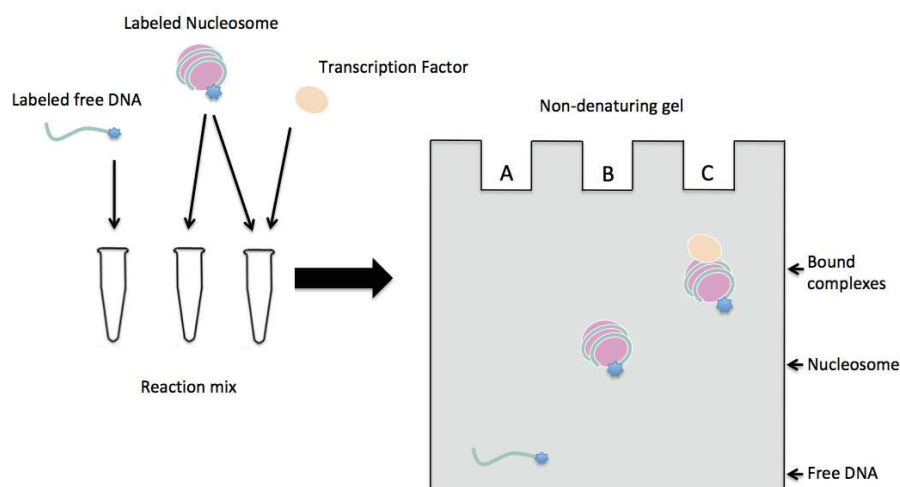


Figure 28. Schematic representation of a conventional EMSA.

The labeled DNA is represented as blue line. The star at one end of the DNA represents the label. Nucleosomes reconstituted with labeled-DNA can be mixed with protein, represented by the simple oval pink shape. After mixing and incubating the sampled, they are loaded into the gel (Adapted from (Alves & Cunha, 2012)). In an EMSA, rather than determining the exact size of a protein-DNA complex, we observe and compare shifting patterns to a known reference.

In the following sections, the work of different research groups concerning the study of pioneer transcription factors and their properties, mainly through EMSAs, will be presented. This will facilitate the understanding of the results obtained for the particular case of LFY.

1. SPECIFIC AND NON-SPECIFIC NUCLEOSOME BINDING

It has been demonstrated that PTFs are able to interact with nucleosomes through both specific and non-specific interactions. Two decades ago, (Li & Wrangé, 1997) showed through EMSAs that a glucocorticoid hormone-activated TF (GR), which binds to DNA as a homodimer, is able to target its binding site (a Glucocorticoid Response Element (GRE)) on nucleosomal DNA. Even though today it is still uncertain how this factor can serve as a PTF, genomic sites where its binding creates DNase I hypersensitivity and recruitment of other TFs have been identified (Voss & Hager, 2014).

Three different nucleosome particles were tested against the GR transcription factor. A nucleosome without any GRE (nucleosome called nTG), and two other nucleosomes containing a GRE, but flanked by different DNA sequences (nucleosome n3GoA1 with GRE flanked by an AT-rich sequences and nucleosome n3Go1 with GRE flanked by a GC triplet on each side) (Figure 29). The experiment revealed, that the relative affinity of GR to n3GoA1 is about threefold higher compared to n3Go1 and about tenfold higher compared to nTG. This illustrates that GR has a higher affinity for nucleosomes with GRE and that the flanking DNA sequence can have a strong impact on the TF's accessibility to a nucleosomal DNA site (Li & Wrangé, 1997).

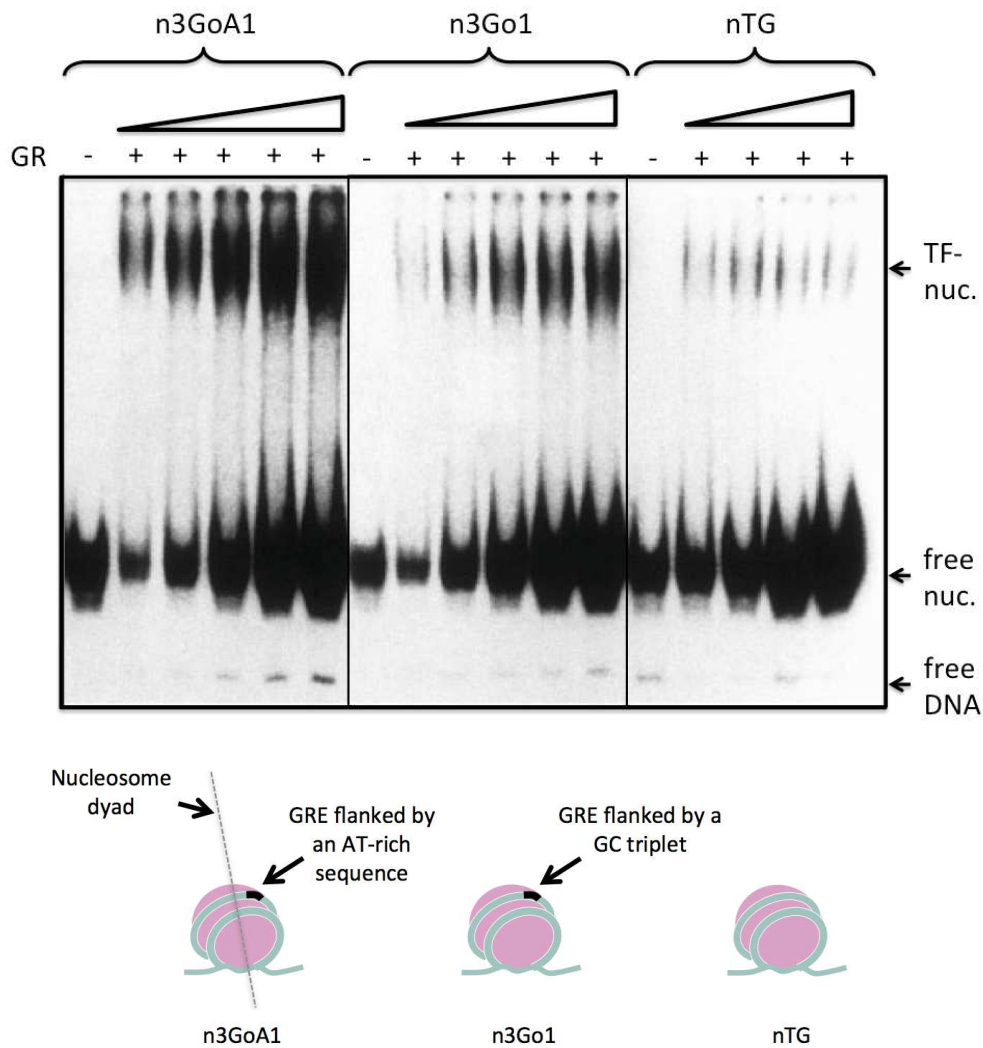


Figure 29. Relative GR affinity to 161-bp nucleosomal templates n3GoA1, n3Go1 and nTG.

On top, EMSA of the binding of GR to nucleosomes. Nucleosomes n3GoA1 and n3Go1 contain a binding site for GR but not nTG. Increasing amounts of the indicated nucleosome probe (9.4, 18.8, 37.5, 75.0, and 112.5 pM) were incubated with a constant concentration of GR (1.2 nM) (Li & Wrangé, 1997). Below, a diagram of the different nucleosome probes. In n3GoA1 the GRE, shadowed segment, is positioned 30 bp away from the nucleosome dyad and is flanked by an AT-rich dark segment. Next, in n3Go1 a GC triplet is present on each side flanks the GRE. Last, nTG does not contain any GRE. The flanking sequences of the GRE on nucleosomes, had an impact on GR's ability to bind them.

Besides the fact that the flanking sequence of a binding site positioned on a nucleosome can have an effect on TF's binding, Lone et al., (2013) showed that the position of the binding site within a nucleosome can also have an effect on TF's binding.

Another example of the use of *in vitro* reconstituted nucleosomes to assess pioneer TFs and in particular their specificity, is the work of Sekiya et al., (2009), in conjunction with *in*

in vivo approaches. In this study, several TFs involved in mouse liver cells' specification: c-Myc, GATA-4, FoxA1, high-mobility group protein B1 (HMGB1) and nuclear factor 1 (NF-1), were studied.

Because FoxA1 moves more slowly than other TFs in nuclei of cells, it was hypothesized that this low mobility could be related to its ability to bind nucleosomes non-specifically. EMSAs were performed in order to assess this. The TFs mentioned were tested against nucleosomes made with two so-called “positioning” sequences: NCP147 and 5S rDNA (Becker et al., 1997). Positioning is a measure of the extent to which a population of nucleosomes resists deviating from its consensus location along the DNA and can be thought of in terms of a single reference point on the nucleosome, like its dyad (Pugh, 2010). These sequences did not contain target sites for any of the TFs (Figure 30). The FoxA1 TF showed a higher affinity for nucleosomes compared to the other TFs tested. This inversely correlates with its mobility in nuclei.

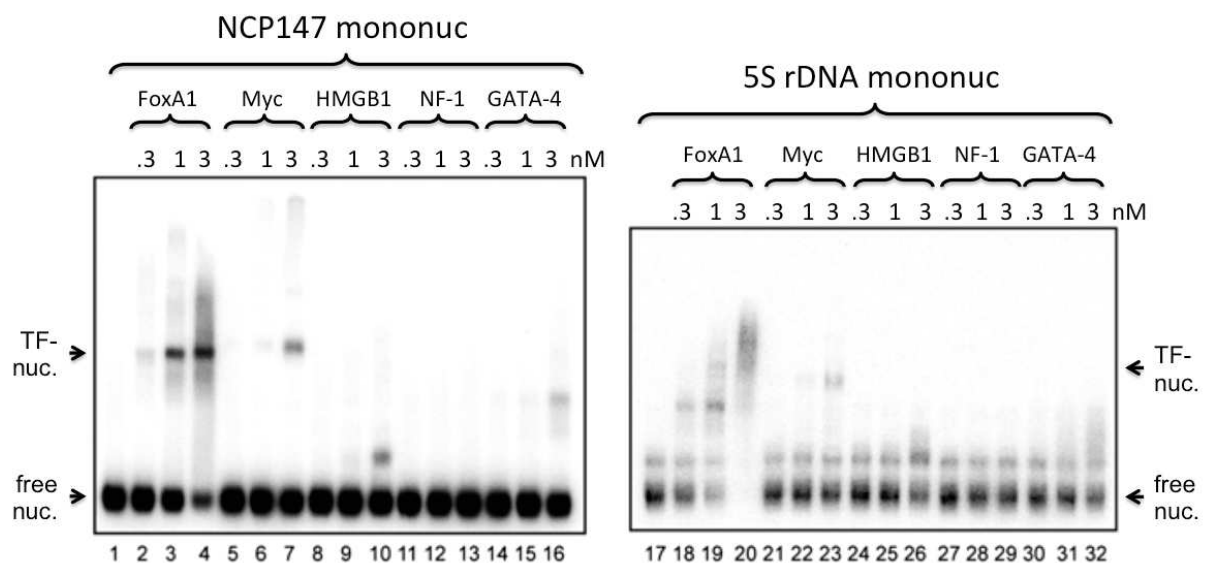


Figure 30. EMSA assessing the interaction of c-Myc, GATA-4, FoxA1, HMGB1 and NF-1, against two nucleosomes. In the left, NCP147 nucleosome, in the right, 5S rDNA nucleosome, both without any binding site for any of the tested TFs. Increasing amounts of TFs were used, in the presence of a fixed amount of nucleosomes, at 2 nM. In these conditions, TFs interact to a different extent, in a non-specific way with nucleosomes. FoxA showed the strongest binding ability to both NCP147 and 5S rDNA nucleosomes. The non-specific binding component to nucleosome binding is a characteristic that some PTFs display (Sekiya et al., 2009).

Complementary experiments were done to look in detail at the specific binding of FoxA1 to nucleosomes. For this, FoxA1 binding sites were inserted in nucleosome NCP147, at different positions in relation to the dyad axis. The binding of FoxA1 to NCP147 and five other nucleosomes containing target sites was observed in EMSA. FoxA1 presented a slight stronger binding to a nucleosome containing a specific binding site, close to the dyad axis (D+2) than to the control (NCP147) (Figure 31).

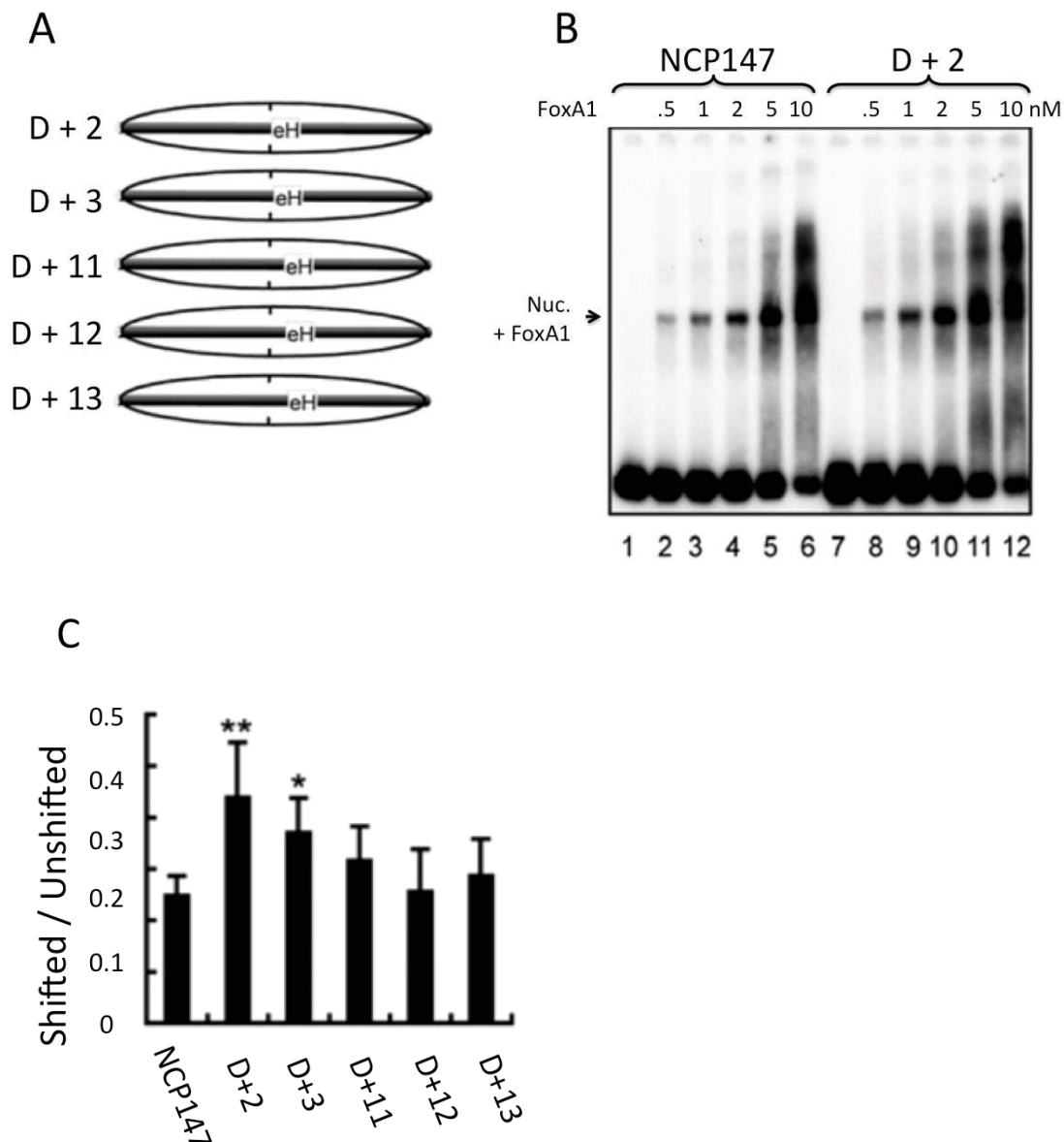


Figure 31. EMSA assessing the specific and non-specific binding of FoxA1 to nucleosome NCP147.

A | Diagram of different versions of NCP147 with inserted FoxA binding sites at indicated base pair positions with respect to the dyad axis (indicated in the center on the nucleosome). B | EMSA with increasing amounts of FoxA1 with respect to a fixed amount of nucleosome (2 nM). C | Quantification of the EMSA in B plus EMSAs performed with the other NCP147 versions shown in A. The ratio of shifted/unshifted nucleosomes at 2 nM

FoxA1 is shown. Standard deviation (SD); n = 3; (*) P < 0.005; (**) P < 0.01, student's t-test (Sekiya et al., 2009). FoxA1 displays a slight stronger binding to a nucleosome containing a specific binding site (D+2) than to the control (NCP147). Interestingly, the strength of this specific binding is decreased in the four other versions (D+3, D+11, D+12, D+13) where the binding site is displaced away from the nucleosome dyad axis. FoxA1 is a PTF that clearly presents a component of specific and non-specific binding to nucleosomes.

In conclusion, FoxA1 binding to nucleosomes is composed of a specific as well as a non-specific binding. Even though FoxA1 is indeed able to specifically bind to nucleosomes, its non-specific binding to nucleosomes is a major determinant of FoxA1 movement in the nucleus.

2. PIONEER AND NON-PIONEER TRANSCRIPTION FACTORS

As explained before, one characteristic that differentiates PTFs from other TFs, is the ability to bind nucleosomes. A recent study assessed by EMSAs the nucleosome-binding ability of a group of TFs with the combined skill to bind closed chromatin. Often not only one, but a collection of transcription factors is required to engage a developmental program. Do all reprogramming TFs are pioneer? Do they all share the same molecular mechanism for nucleosome binding?

Soufi et al., (2015) assessed this question concerning a group of four TFs necessary for liver induction by using EMSAs, among other techniques. A 162-bp DNA sequence containing target sites for each of the TFs, was specifically selected from genome-wide data analysis for nucleosome reconstitution. Interestingly, the three TFs: Oct4, Sox2, and in a lesser extent Klf4, but not c-Myc, showed the capacity to target sites located on nucleosomes (Figure 32).

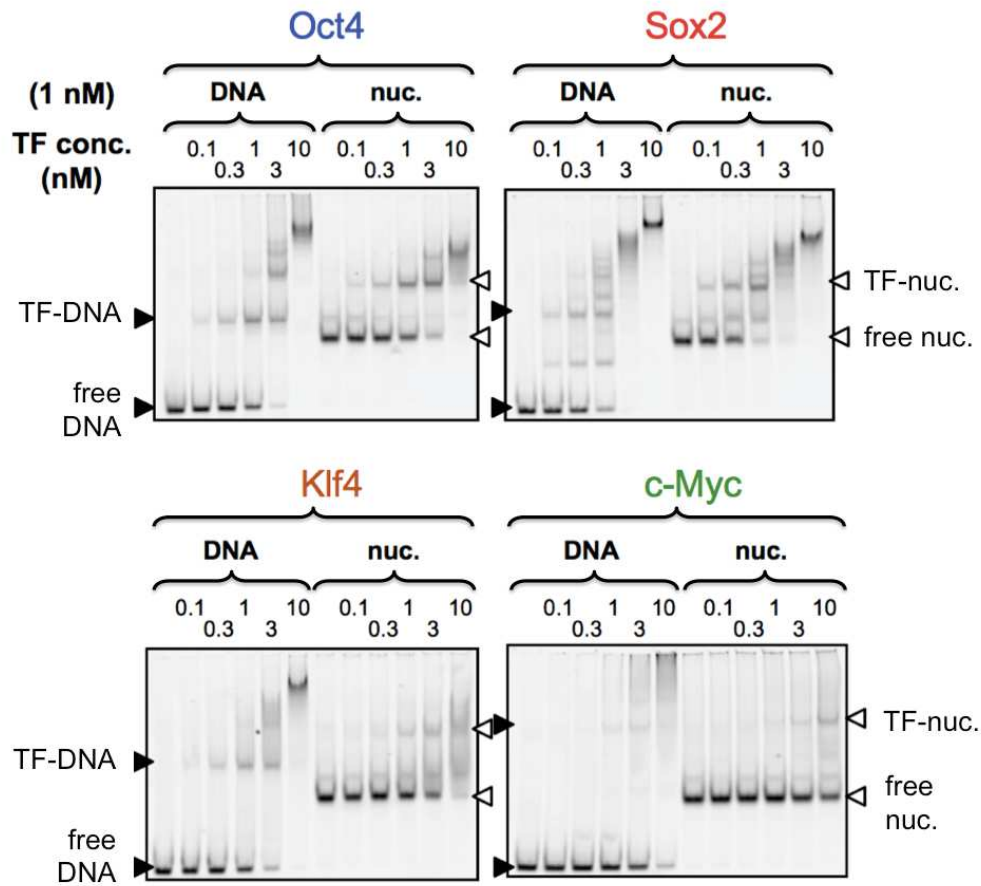


Figure 32. EMSA showing the shift patterns of increasing amounts of recombinant TFs to free DNA and nucleosomes.

Increasing amounts of recombinant (from left to right and top to bottom) Oct4, Sox2, Klf4, and c-Myc proteins to Cy5- labeled free DNA as well as, to *in vitro* reconstituted nucleosomes. Free DNA migrates the fastest in the gel and is indicated by black arrows at the bottom of each gel. Bands appearing above the free-DNA bands represent TF-DNA complexes, of larger size and thus migrating slower than free DNA alone. Nucleosomes, on the right side of each gel, migrate slower than free DNA due to their size. Bands appearing above the free-nucleosome bands represent TF-nucleosome complexes, of large size and thus migrating slower than nucleosomes alone. All TFs are able to completely shift the band corresponding to nucleosomes at the highest concentration of TF (10 nM), except for c-Myc, where only a faint band corresponding to a c-Myc-nucleosome complex is observed. In complementary experiments, Oct4, Sox2, Klf4, and c-Myc proteins showed less affinity to DNA probes containing non-specific sites (Soufi et al., 2015). This EMSA allows to assess one of the main pioneer characteristics of a TF, its ability to bind nucleosomal DNA, and therefore separate pioneer from non-pioneer TFs.

In order to verify that the shifted patterns observed in the figure above represented a TF-nucleosome complex, Soufi et al., (2015) performed a western blot after an EMSA. This allows tracking not only the migration of DNA but also of histones and the TF (Figure 33). Histone H3 and H2B detected in the western blot assay and their migration correspond to the nucleosome-TF complex. Besides they were detected in an equimolar amount, meaning the nucleosomes conserved their histone octamer and were not destabilized by Oct4 binding.

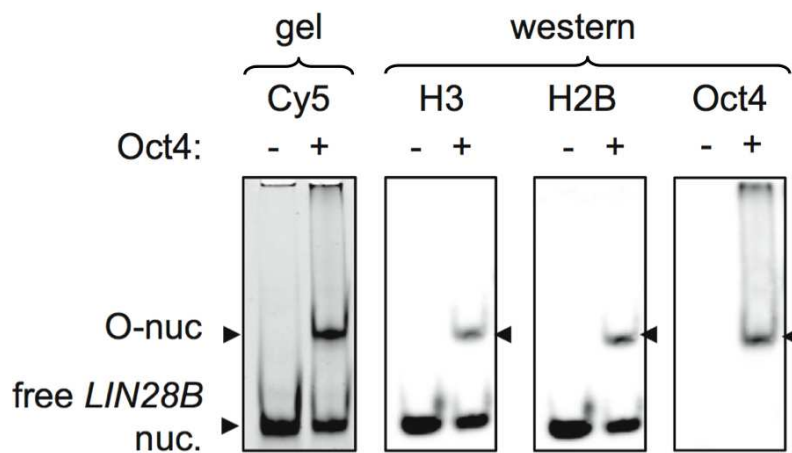


Figure 33. WEMSA of Oct4 binding to a nucleosome containing H2A and H3 histones in equimolar amounts. The EMSAs shows the binding of Oct4 at 1 nM (left panel). After the EMSA, the proteins were transferred onto a PVDF membrane (WEMSA) and blotted for H3, H2B, Oct4, as indicated in the three panels on the right. The upper black arrowheads indicate the observed TF-nucleosome complexes. The lower black arrowhead indicates only nucleosome LIN28B (Soufi et al., 2015). The western blots after the EMSA allowed to verify that the shifted band in the presence of Oct4 represents a TF-nucleosome complex, since two of the four core histones have the same migration that the Cy5-labeled nucleosomal DNA and were detected in equal amounts, as expected in nucleosomes.

Overall, the mentioned examples of EMSAs to study PTFs were useful to illustrate some of their properties. A pioneer transcription factor binding to nucleosomes is composed of both specific and non-specific interactions. The position of a binding site within a nucleosome core particle can affect the affinity of the PTF's binding. Furthermore, the flanking sequences of a binding site were shown to also have an impact on the PTF's affinity. In the following section, I will present our methods and results concerning the LFY TF.

II. NUCLEOSOME RECONSTITUTION, OUR BIOPHYSICAL APPROACH

1. THE NUCLEOSOME AS A FAMILY OF STRUCTURES

In the previous section I explained the reason for choosing the *in vitro* approach of nucleosome reconstitution for assessing LFY's nucleosome binding. In this section, I would nonetheless like to point out that nucleosomes, as mentioned in the introduction, are not static simple units. On the contrary, these units are always subjected to changes in their structure by several mechanisms.

One of them involves spontaneous conformational changes that have been revealed by high-speed atomic force microscopy (HS-AFM) studies (Ordu et al., 2016). Due to conformational changes after torsional stress, nucleosomes can change their "chirality" from a left-handed to a right-handed DNA-wrapping to give a reversed nucleosomal structure called reversome (Bancaud et al., 2007).

Also, the interactions that nucleosomes can have with different kind of proteins (histone-modifying enzymes, ATP- dependent CRCs, incorporation of histone variants, interactions with chaperones, among others) can alter their structure. For instance, the transcription enzyme RNA polymerase II (Pol II) can produce a hexameric sub-complex missing one H2A/H2B dimer, termed a hexasome (Bintu et al., 2012). The subnucleosomal structure that only contains the (H3-H4)₂ tetramer is called a tetrasome. Interestingly, Human SWI/SNF CRCs acting on mono-nucleosomes are able to generate altered di-nucleosomes. In these structures, DNA is partially released from one nucleosome and is thought to bind to the exposed histone surface of another nucleosome (Lavelle & Prunell, 2017).

Finally, the nature of a DNA sequence influences the strength of DNA-histone interactions. DNA sequence can affect the static curvature, flexibility and stability (Ngo et al., 2015). In an effort to take this dynamics into account, we identified nucleosome-enriched regions in Arabidopsis' genome that are efficiently targeted by LFY and in a closed state,

according to MNase data. Mnase-seq datasets are the most widely used data to map nucleosome positions throughout the genomes. DNase-seq data sets can also contribute to the identification of nucleosomes since they give information about the accessibility of a given genomic region. Briefly, DNase-seq data provides inferred information on chromatin-binding proteins that lie between hypersensitive sites, whereas Mnase-seq data maps protected chromatin regions.

2. SELECTION OF DNA SEQUENCES

The 601 DNA sequence was identified by Lowary & Widom, (1998) using the Systematic Evolution of Ligands by Exponential enrichment (SELEX) method. It is one of the strongest known positioning sequence to constitute a nucleosome. Because of its high affinity for histones it has been useful for *in vitro* nucleosome reconstitution. My initial work of nucleosome reconstitution was made using this sequence, as well as a modified 601 sequence, with an inserted LFY's target site at the nucleosome dyad, as shown in Figure 34.

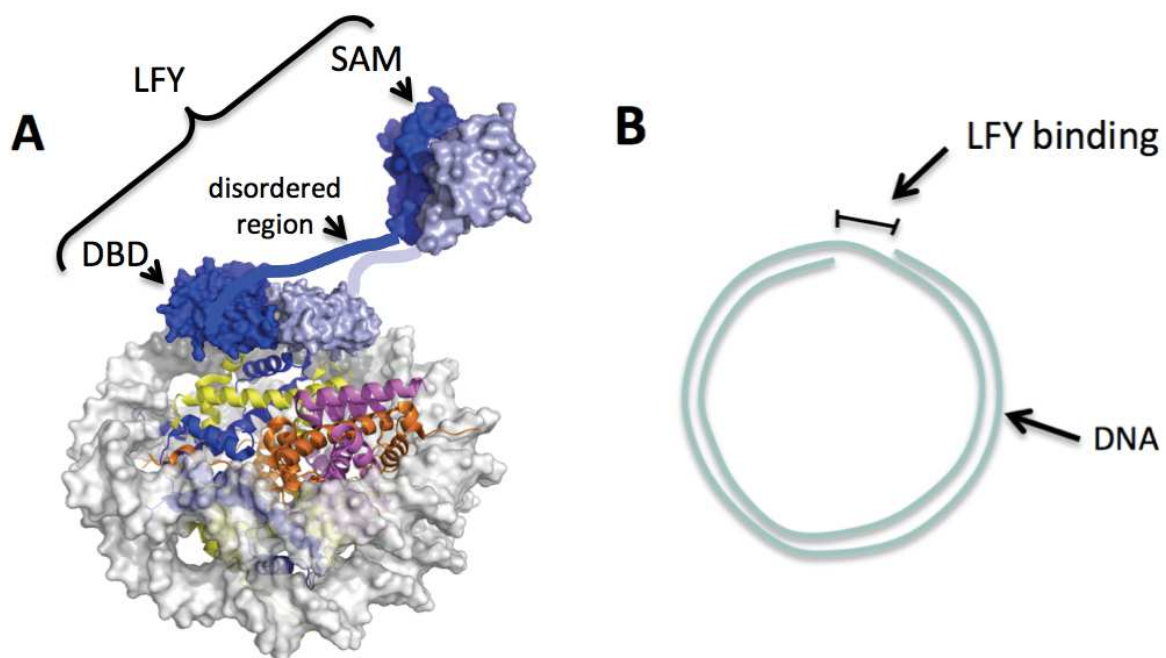


Figure 34. Model for LFY binding in a nucleosome core particle (NCP).

A | On top, LFY binding as a dimer (one monomer in blue, the other in grey) on a region in the nucleosome where there is only one strand of DNA. On the bottom, the crystal structure of a nucleosome core particle with DNA 601 is shown. In light grey, the DNA double strand, in blue, yellow, orange and purple, two copies of each of

the core histones. B | Schematic representation of the DNA wrapped around the histone octamer (not shown) to show the location of the LFY-binding site in the nucleosome. In this nucleosome 601, the LFY binding site is positioned at its central position, which is in this case, a region in the nucleosome with only one DNA strand, instead of two, facilitating the access for a TF.

Apart from these sequences, with the help of bioinformatic analysis done by Laura Grégoire, we identified nucleosome-enriched regions in the Arabidopsis genome, efficiently targeted by LFY. To identify these regions, we used genome-wide data from Chromatin immunoprecipitation and sequencing (ChIP-seq) (submitted to GEO with accession number GSE64245) made in our group, using plants p35S:LFY, p35S:LFY_{TERE} and wild type Arabidopsis.

Genomic regions for nucleosome reconstitution were selected to have both a strong ChIP-seq signal in p35S:LFY plants and a substantial loss of this binding in p35S:LFY_{TERE} plants. This approach thus selected genomic regions in which LFY binding is facilitated by oligomerization and potentially more prone to favor a PTF activity of LFY.

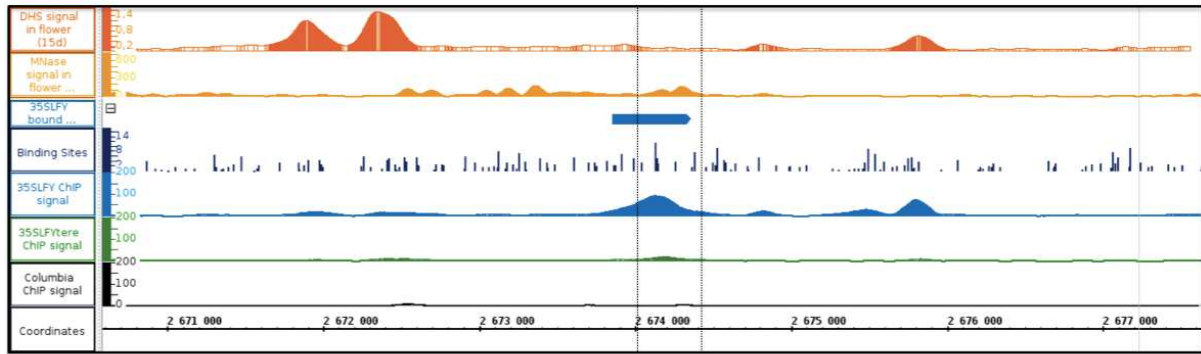
We also used, DNase-seq data from Zhang et al., (2012) as well as MNase-seq from Zhang et al., (2015). MNase-seq technology consists of a digestion made by a micrococcal nuclease (MNase) combined with high-throughput sequencing (MNase-seq), a powerful technique useful to map the genome-wide distribution of nucleosome occupancy. The thresholds used for DNase and MNase data were used according to the previous work established by Zhang et al., (2012) and Zhang et al., (2015).

Altogether, the criteria used to select genomic regions where LFY could act as a PTF were the following:

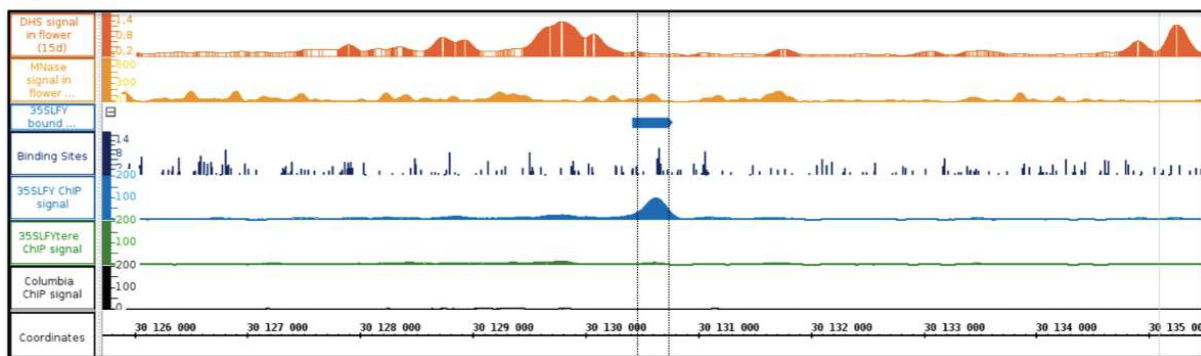
- +/- 75 bp around an MNase peak for each nucleosome core particle
- A value > 200 reads per base (rpb) for each MNase peak
- At least one LFY target site with an affinity score > -15
- Low DNase-seq signal indicating a closed region
- A loss of LFY binding comparing ChIP-seq data (p35S:LFY to p35S:LFY_{TERE})

The profiles of the selected regions are shown in the next Figure 35.

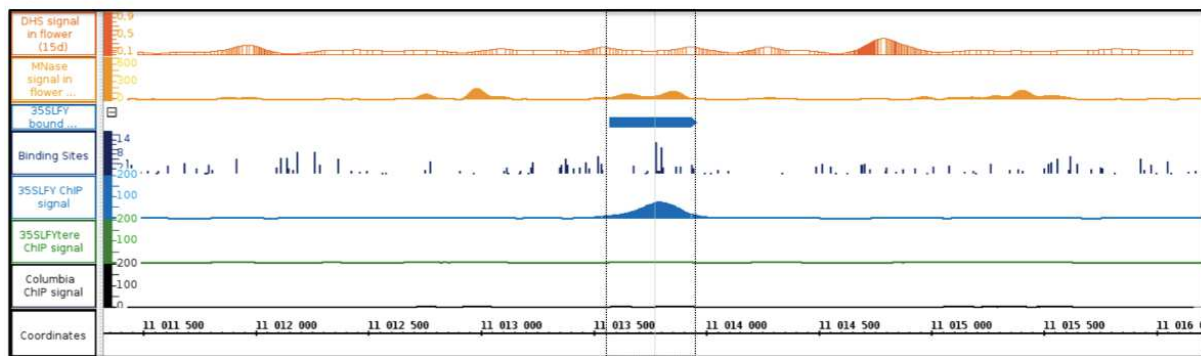
1



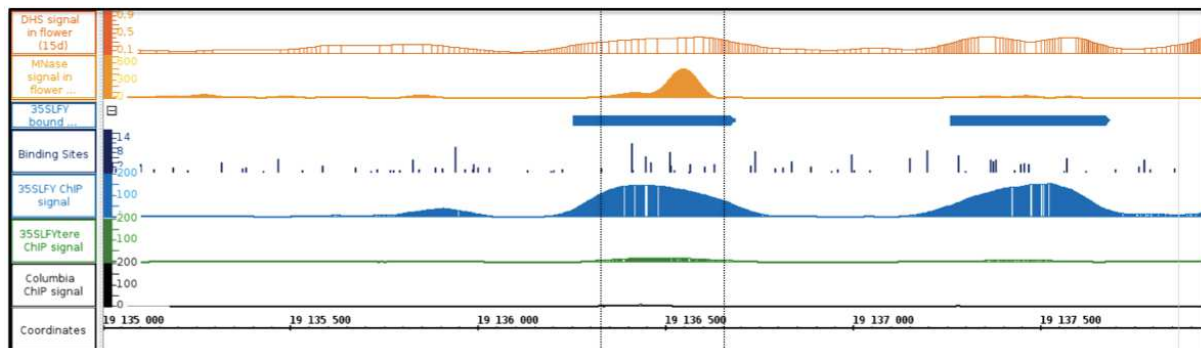
2



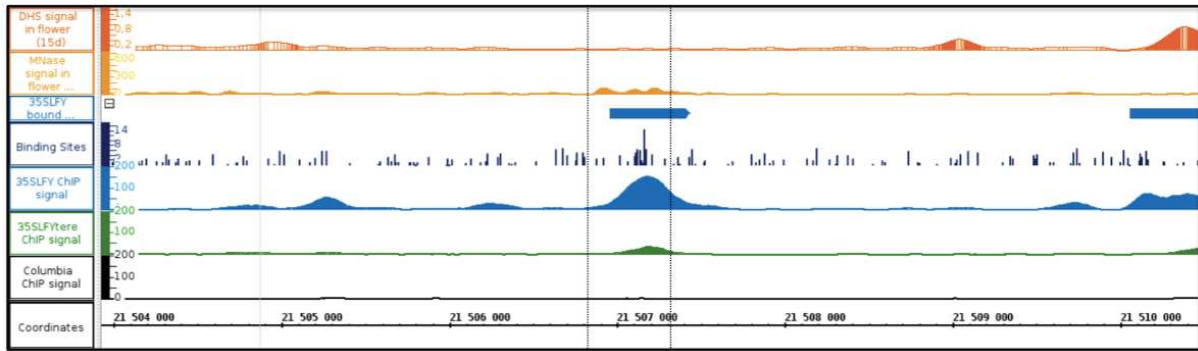
3



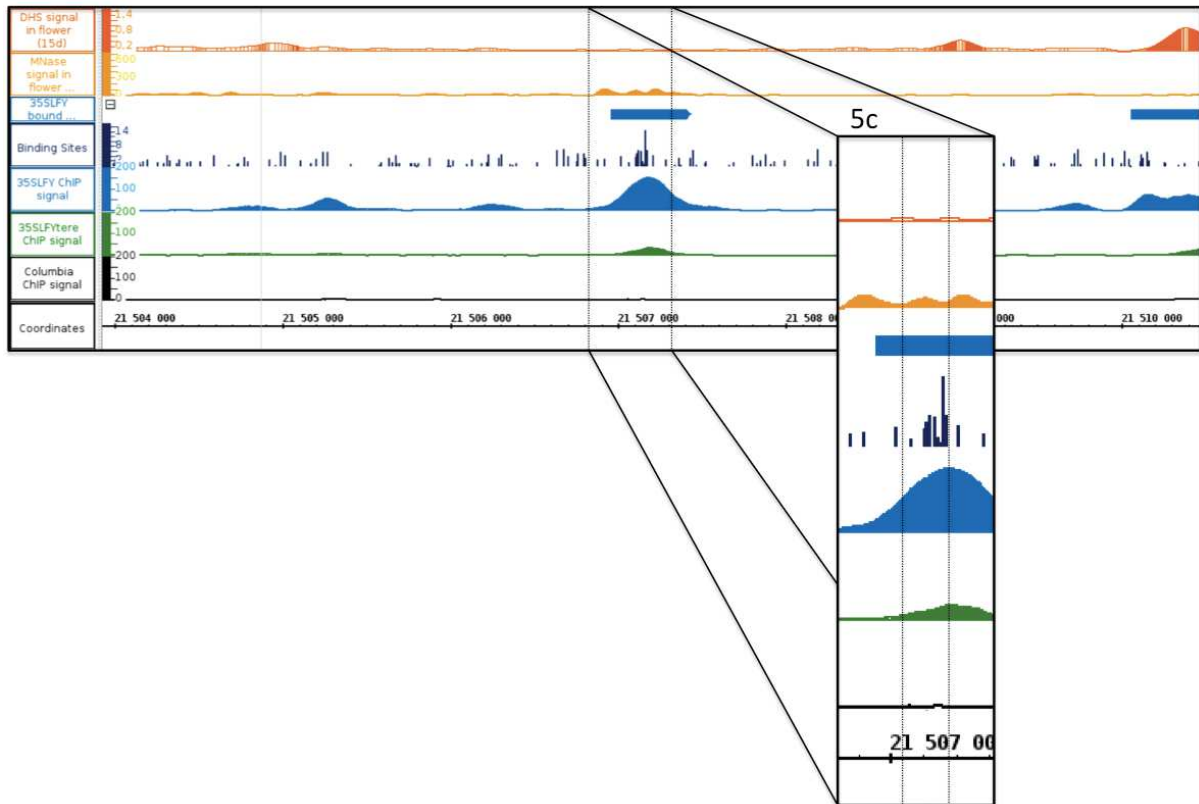
4



5



5



6

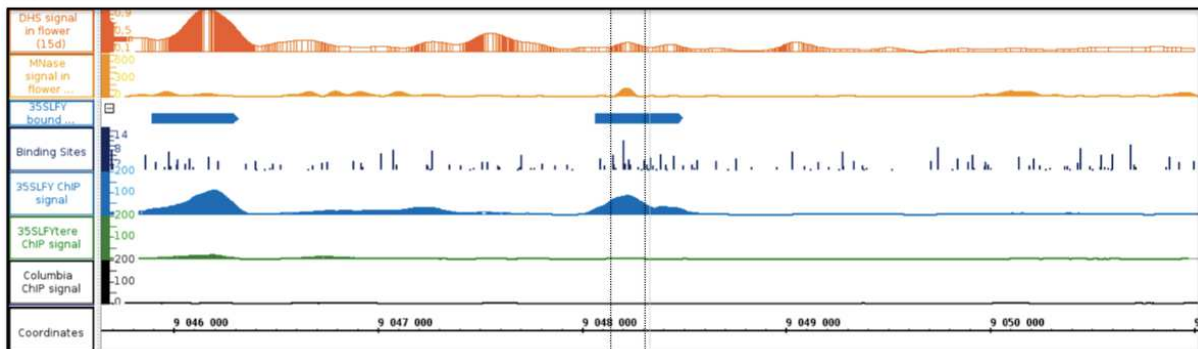


Figure 35. Identification of six DNA regions suitable for nucleosome formation.

Inside the fine black boxes are the DNA regions selected for later nucleosome reconstitution. From top to bottom, the DNase signal (in orange) allows to map open genomic regions. The MNase signal (in yellow) allows mapping occluded regions where DNA is packed into nucleosomes. In blue, the bound regions for the p35S:LFY transgenic line. The row showing LFY binding sites (in dark blue), shows the exact starting position where the protein binds, according to the previously validated LFY's position weight matrix (PWM). The ChIP signal for p35S:LFY, p35S:LFYTERE as well as WT (in blue, green and black respectively). Sequence 5 genomic location is chr3-21506957-21507437 bp. Sequence 2 genomic location is chr1-30130416-30130778 bp. Sequence 1 genomic location is chr1-2673853-2674360 bp. Sequence 3 genomic location is chr2-11013571-11013958 bp. Sequence 6 genomic location is chr4-9048065-9048502 bp. Sequence 4 genomic location is chr3-19136254-19136688 bp. Neither of these locations are on or nearby genes. Fragments number 2, 5c and 6 contain one nucleosome-positioning sequence. Fragments 1, 3 and 4 contain two nucleosome-positioning sequences and fragment number 5 contains three.

Thus, I worked with DNA sequences of different nature. Initially, the 601 sequence allowed the assembly of a well-positioned nucleosome already used and studied in the past, containing a LFY target site. However, the other DNA sequences represent also good candidates for *in vitro* reconstitution with the advantage of being real sequences selected from the Arabidopsis genome. Sequences 2 and 5c were chosen for reconstitution and further EMSAs because the reconstitution process for a single nucleosome is more straightforward than the reconstitution of nucleosome arrays (more than one nucleosome) and because these fragments contained the strongest LFY binding sites of all the six selected DNA fragments. However, since these regions are taken from Arabidopsis genome, they have often low-affinity sites within them. Therefore, mutations were made in order to eliminate all strong LFY binding sites and be then able to measure non-specific binding of LFY.

3. NUCLEOSOME RECONSTITUTION

A. Production and purification of histone octamers

Once having selected the DNA sequences for nucleosome reconstitution, how can we reconstitute nucleosomes? Nucleosomes are composed of DNA and histones. To produce recombinant histones, we initially tried to use a polycistronic vector containing Arabidopsis core histones but the plasmid synthesis was unsuccessful. However, Yoonjung Shim, kindly

provided a polycistronic vector containing *Xenopus Laevis* core histones (Shim et al., 2012), which are highly similar to *Arabidopsis* histones, as shown in Figure 36. Therefore, the core histones for nucleosome reconstitution, were recombinant *Xenopus Laevis*: H2A, H2B, H3, and H4.

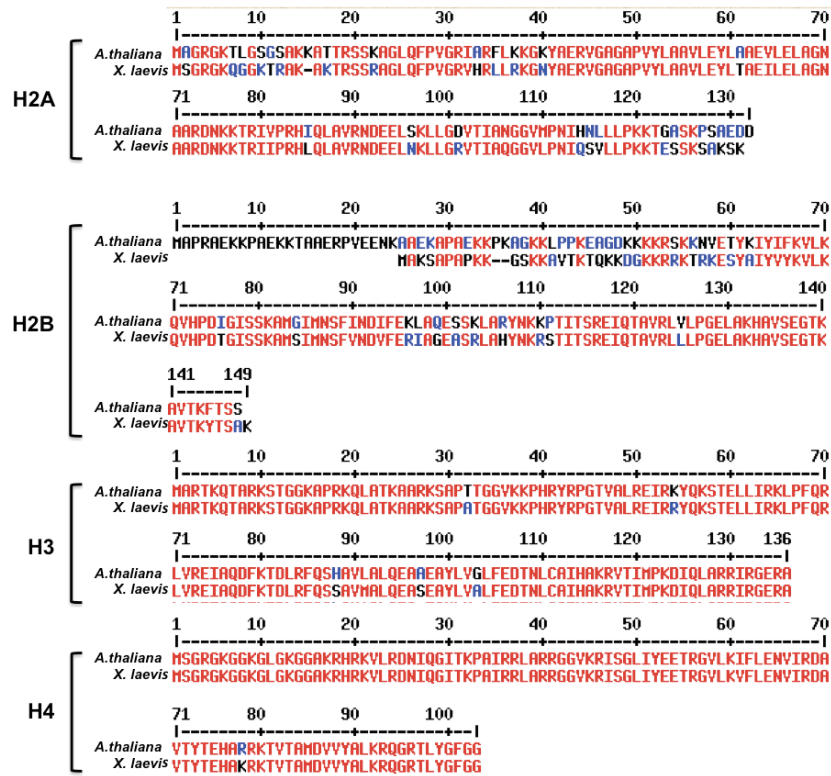


Figure 36. Protein sequence alignment of the four core histones using Multialin.

In red, conserved amino acids in both proteins. In blue and black different amino acids having the same position. Core histones are highly conserved across eukaryotes in terms of sequence and structure (Mariño-Ramírez et al., 2006). Here, it can be observed in detail that core histones from *A. thaliana* and *X. laevis* are highly conserved with the exception of the N-terminal of H2B.

To obtain purified histones octamers, the methods of Shim et al., (2012) were followed. It is the first published protocol of purification in non-denaturing conditions. (The overall steps are shown in Figure 37. This protocol, allows the co-expression and co-purification of the four core histones from a single polycistronic vector. This vector was constructed by incorporating the four histone genes into the multiple cloning site of a pET vector under the control of a single T7 promoter. To facilitate the purification, an hexahistidine (His₆) was added to histones H2A and H4. As H2A forms a dimer with H2B and

H4 forms a tetramer with H3, H2B and H3 are co-purified. A thrombin site in the N and C termini, was also added to histones H2A and H4, respectively, to remove the His₆ tag after purification. Since all histones are co- expressed, they can immediately form stable histone octamers. Purification is performed in conditions of 2 M NaCl. Under these conditions, histones can dissociate from bulk bacterial DNA and remain mostly in solution instead of inclusion bodies. This allows bypassing denaturation and refolding procedures. Indeed, in most protocols, the four subunits (H2A, H2B, H3, and H4) are overexpressed individually and purified under denaturing conditions from insoluble inclusion bodies (Lee et al., 2015). Octamers are then formed by refolding the denatured histones together in equimolar ratios.

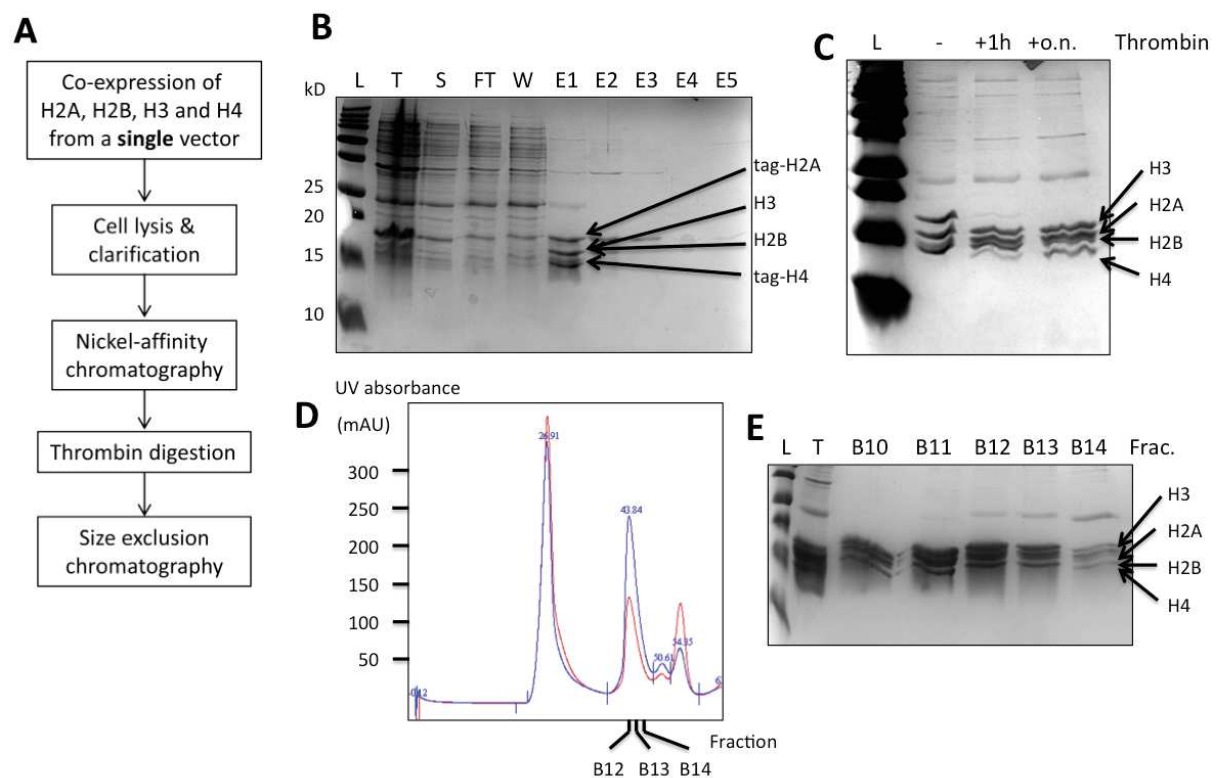


Figure 37. Results obtained following the global procedure for the co-expression and co-purification of histone octamers of Shim et al., (2012).

A | General workflow of histone octamer preparation. B | SDS-PAGE gel of the nickel affinity column purification, starting from cell supernatant. The purification is possible thanks to two 6-histidine tags in H2A and H4. Histones are eluted at 30 and 50 mM of imidazole. L, protein ladder. T, total fraction. S, soluble fraction. FT, flow-through, W, wash. E1-E5, elution at 30, 50, 100, 150, 200 mM imidazole. C | The affinity tags on H2A and H4 were removed by thrombin digestion. (-) and (+1h, + o.n.) thrombin indicate histone samples before and after 1h and overnight thrombin digestion, respectively. D | Thrombin-digested histones were concentrated to approximately 3 mg/mL, therefore around 3 mg were purified over a Superdex 200 10/300 (GE Healthcare)

column. The chromatogram (D) and SDS–PAGE gel (E) show that stoichiometric histone complexes elute at around 12.8 ml, in the B12-B14 fractions. UV, ultraviolet. E | SDS–PAGE gel of the Superdex 200 column nickel purification. L, molecular weight marker. T, sample before purification. B10 to B14 indicate fraction numbers. In this method, all four core histones (H2A, H2B, H3, and H4) are expressed together from a single polycistronic vector and purified with two chromatography steps under non-denaturing conditions.

B. Nucleosome assembly *in vitro*

In vivo, nucleosome assembly happens thanks to the help of specialized enzymes called chaperones such as assembly protein 1 (Nap1) and anti-silencing factor 1(Asf1), called chaperones. They facilitate and regulate this process. Even though how exactly NCPs are assembled is still an open question, recent studies from Fan et al., (2015), propose a model where histone chaperones mediate nucleosome assembly along a path that leads to thermodynamically-favored products.

In vitro nucleosome reconstitution can be made without histone chaperones. Instead, salt concentration can be used to play the role of mediator in this assembly. This happens in a stepwise manner. Initially, DNA and core histones are mixed in high salt concentration solutions. At 2 M salt, interaction between histones and DNA does not take place. Salts associate with DNA strands preventing binding to histones. Afterwards, with a progressive decrease in salt concentration the initial binding of the (H3–H4)₂ tetramer to the DNA, followed by the incorporation of the two H2A-H2B dimers can take place (Polo & Almouzni, 2006). In agreement with this, nucleosomes have been observed to lose their H2A-H2B dimers under an increase in salt concentration, while (H3–H4)₂ tetramers remain bound to the DNA (Ordu et al., 2016). Two different protocols were tested in this work to reduce salt concentration. The first one, and most commonly used, consist of a salt dialysis (Becker et al., 1997). The second one, of a salt dilution (Okuwaki et al., 2005) (Detailed in Materials and Methods).

In order to label DNA, PCR was used to generate Cy5-labeled DNA fragments. Reconstitution of nucleosomes with each DNA fragment had to be optimized. Normally, the nucleosomes are obtained with a molar ratio of histone octamer to DNA of 1 to 1. Nonetheless, in practice, each DNA fragment needed a different ratio. In Figure 38, an EMSA

of several preparations for nucleosomes made with sequence 601 as well as 601_{LFY} are shown, with an increasing amount of histone octamer compared to DNA. When the histone octamer/DNA ratio is too high, aggregates can form that do not enter the gel. In Figure 39, the reconstitution of a nucleosome with sequence 2 is presented as an example.

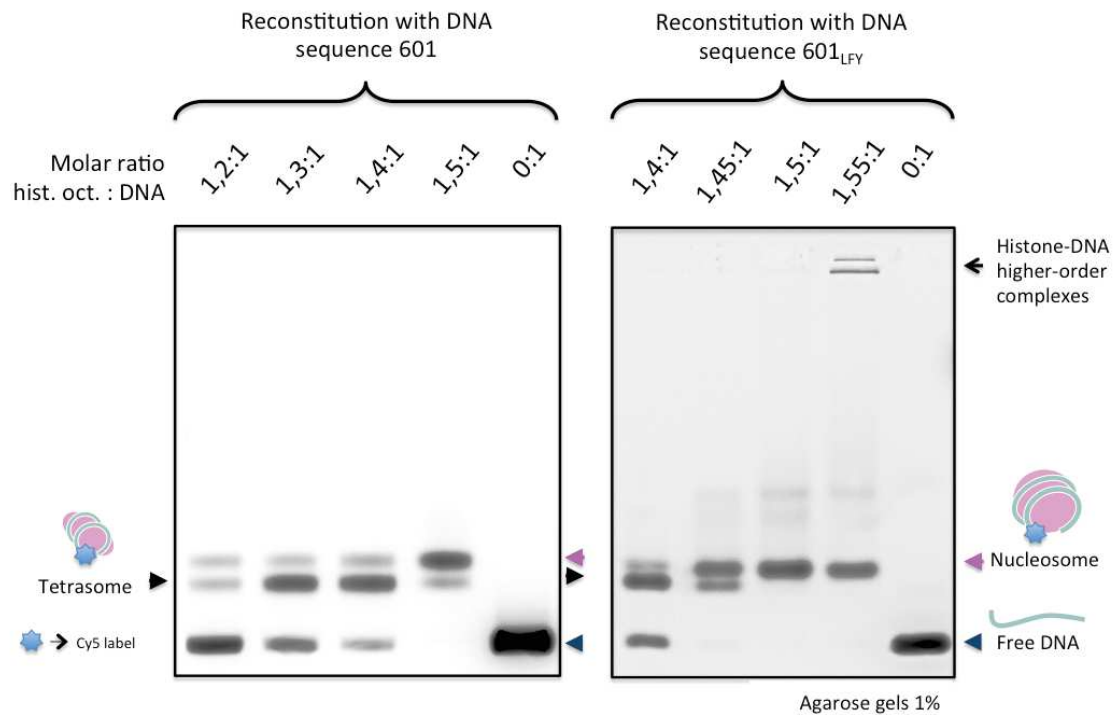


Figure 38. Nucleosome reconstitution by salt dialysis.

The formation of two species is observed, a tetrasome (indicated with a black arrow), is formed at the lower histone octamer : DNA ratios. The formation of this species naturally occurs during the process of nucleosome assembly, it corresponds to a H3-H4 tetramer and it can be the majority species at a certain amount of histone octamer. Nucleosomes, favored at higher histone octamer molar ratios, are indicated with a magenta arrow.

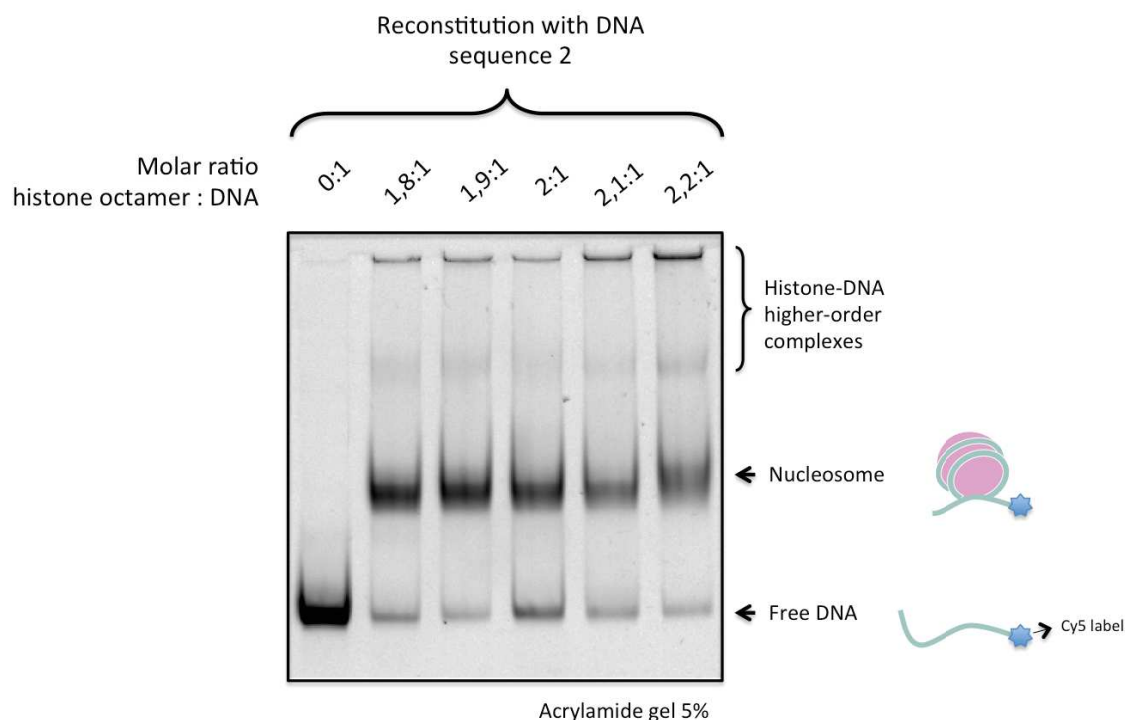


Figure 39. Nucleosome reconstitution by salt dilution.

Increasing histone octamer : DNA molar ratios favor the formation of higher-order aggregates.

C. AFM as a tool to verify nucleosome reconstitution

The reconstitution process was confirmed by Atomic Force Microscopy (AFM), a high-resolution type of scanning probe microscopy, that allows an atomic or molecular-level resolution, in the order of fractions of a nanometer. This was possible thanks to collaboration with Jean-Luc Pellequer and Jean-Marie Teulon in the context of a GRAL project. Samples were imaged with the ScanAsyst software, on a dry mode with a tip of Silicone nitride (ScanAsyt-AIR-HR). Figure 40, shows reconstituted nucleosomes with sequence 4 (482 bp). Finally, Figure 41 shows a schematic representation of nucleosome core particles used for interaction assays.

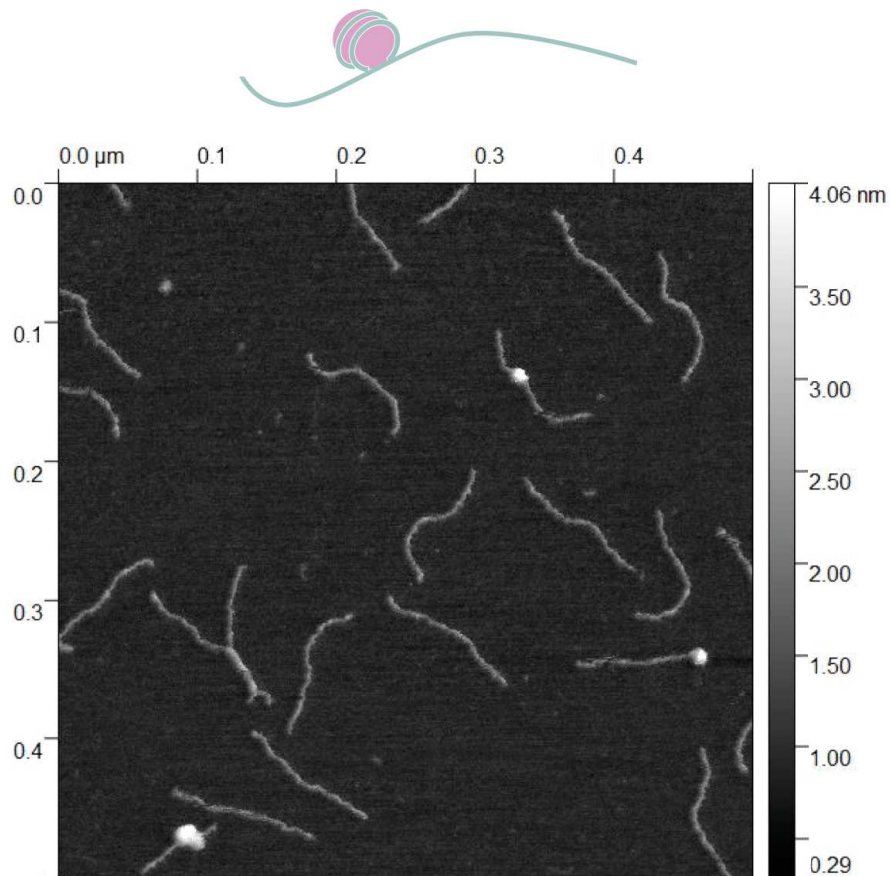


Figure 40. Nucleosome reconstituted with sequence 4 (482 bp).

A 50 μL drop of nucleosomes in an initial buffer of 10 mM Tris-EDTA (TE) drop diluted to a final concentration of 1 nM with distilled water, was deposited on a mica layer, and then dried in a desiccating cabinet. This visualization allows to distinguish molecular objects based on their height. The ladder on the right of the image shows the color given to molecules according to their height on the mica. Nucleosomes appear brighter than free DNA strands because they are bigger complexes containing histones. In some cases, the nucleosome is observed around the center of a DNA molecule, but it can also be at one side of the fragment. During the drying process, molecules on the mica can undergo position and minor conformational changes.








DNA fragment size	Nucleosomes with at least one LFY binding site	Nucleosomes without any LFY binding site
147 bp	601 _{LFY} 	601 
182 bp	2 	2m3 
157 bp	5c 	5cm3 
482 bp	4 	

Figure 41. Schematic representation of nucleosomes made with DNA sequences of different lengths for EMSAs. For nucleosome 601_{LFY}, a strong-affinity LFY binding site (score -7) was inserted. For nucleosomes 2m3 and 5cm3, the strongest-affinity LFY binding sites were mutated. In nucleosome 2m3, the strongest-affinity site had a score of -14.5. In nucleosome 5cm3, the strongest-affinity site had a score of -10. However, all nucleosomes bear some low-affinity binding sites with scores varying of -22 to -27. In particular, nucleosome 5cm3 contains a zone of a series of low-affinity binding sites that could facilitate LFY's cooperative binding (DNA sequences in Materials and Methods). Nucleosome 5c and 5cm3 have 10 extra bps that can protrude from the nucleosome core particle. In a similar way, nucleosome 2 and 2m3 have 35 extra bps, and nucleosome 4 has 335.

To summarize this section, the materials and protocols for nucleosome reconstitution were presented. We reconstituted nucleosomes with DNA sequences 601, 601_{LFY}, 2 and 5c as well as the mutated versions 2c and 5cm3. In the next section I will present the interaction tests that we performed between LFY and these different reconstituted nucleosomes.

III. INTERACTIONS BETWEEN LFY AND NUCLEOSOME CORE PARTICLES

The hypothesis that LFY could directly interact with chromatin is a fascinating open question. This hypothesis raised several questions about whether LFY binds the DNA in the nucleosome, whether it binds specifically compared to other TFs and what is the consequence on nucleosome stability.

1. LFY'S BINDING AND SPECIFICITY FOR NUCLEOSOMES

In order to test the hypothesis that LFY could directly interact with nucleosomes, we used two distinct LFY proteins. A purified version of LFY without the first 40 amino acids (aa) from *A. thaliana* as well as from *G. biloba* were used (AtLFY-40 and GbLFY-40, respectively). These truncated versions where a disordered region is missing, are more stable than the full-length proteins. In this section, I will present and discuss the results obtained from EMSAs assessing LFY's interactions with free DNA sequences and nucleosomes reconstituted with these DNA sequences.

A. Comparative binding of AtLFY and GbLFY to free DNA

These two proteins used have been described previously to bind highly specifically in EMSAs to oligonucleotides carrying one or two LFY-binding sites (Sayou et al., 2016). To determine how LFY binds to DNA in experimental conditions where nucleosomes are formed and stable, we looked at both proteins AtLFY-40 and GbLFY-40 (Figures 42 and 43 respectively) bindings and specificities for sequences 601, 2 and 5c. The assays were carried out in the presence of genomic fish DNA competitor, a non-labeled DNA, used in order to eliminate LFY non-specific binding.

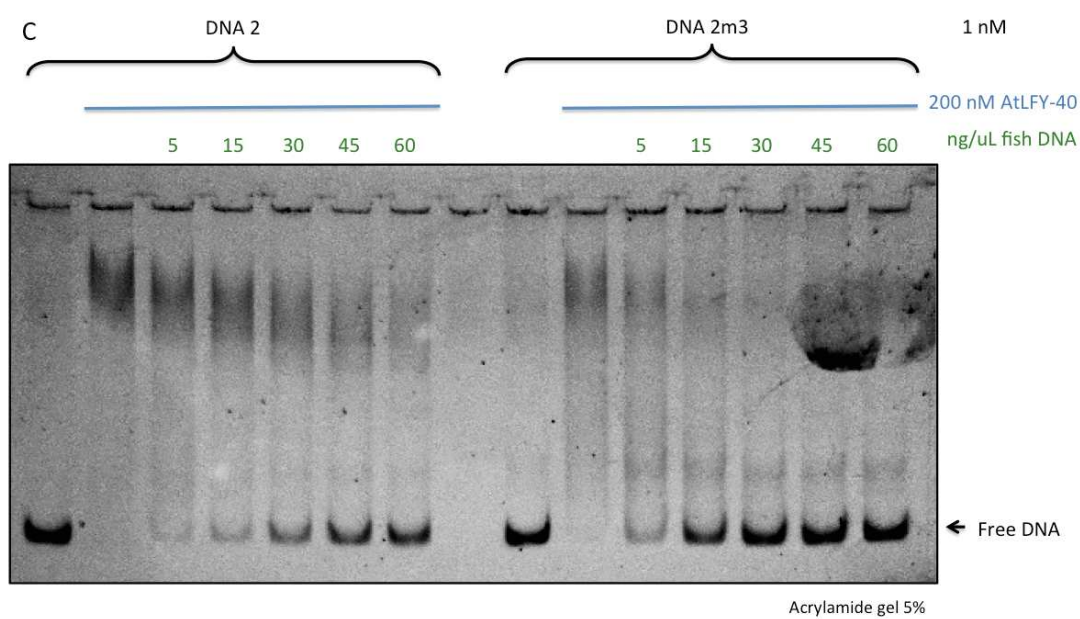
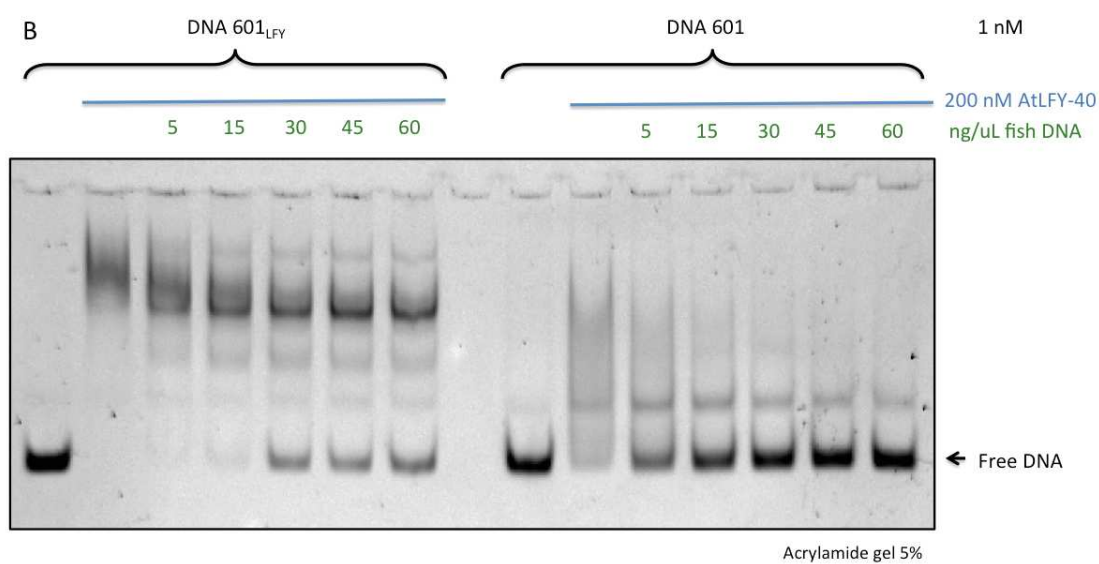
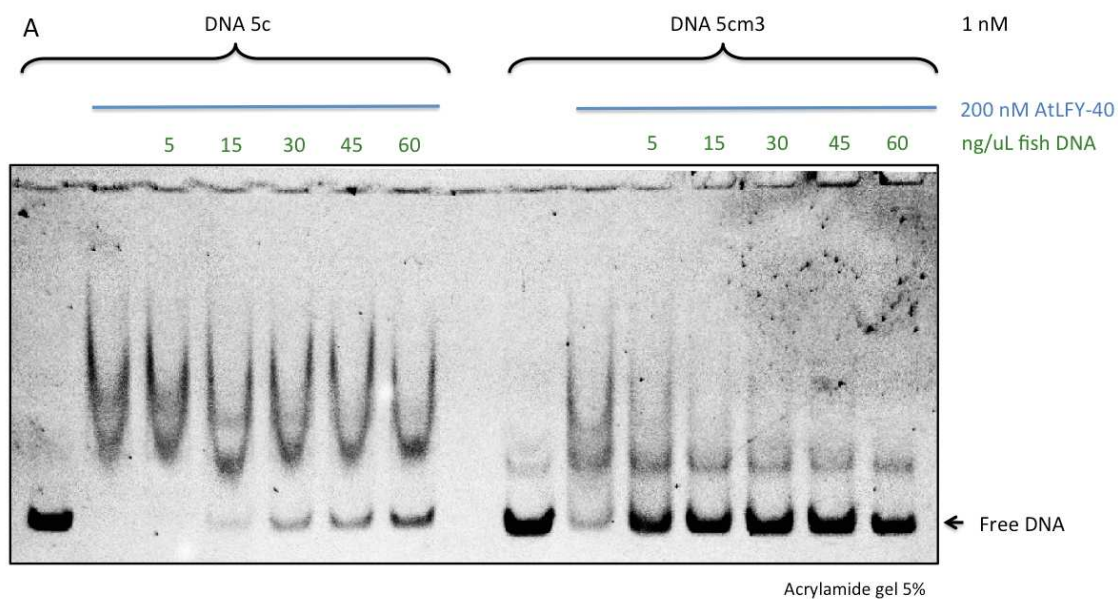


Figure 42. EMSAs with AtLFY-40 and different DNA probes, in the presence of increasing concentrations of fish DNA competitor.

A | Free DNA 5c and its mutated control 5cm3 where LFY binding sites are mutated. B | Free DNA 601_{LFY} and its mutated control with no LFY binding site. C | Free DNA 2 and its mutated 2m3 control. In all three gels, AtLFY-40 concentration is maintained constant and fish DNA competitor concentration is increased in order to test the binding specificity. We can observe that DNA 5c, 601_{LFY} and 2, are shifted at a greater extent by the protein than the respective control DNAs with no LFY binding sites, when the same conditions of protein and competitor DNA are used. Therefore AtLFY-40 can bind specifically to DNA fragments..

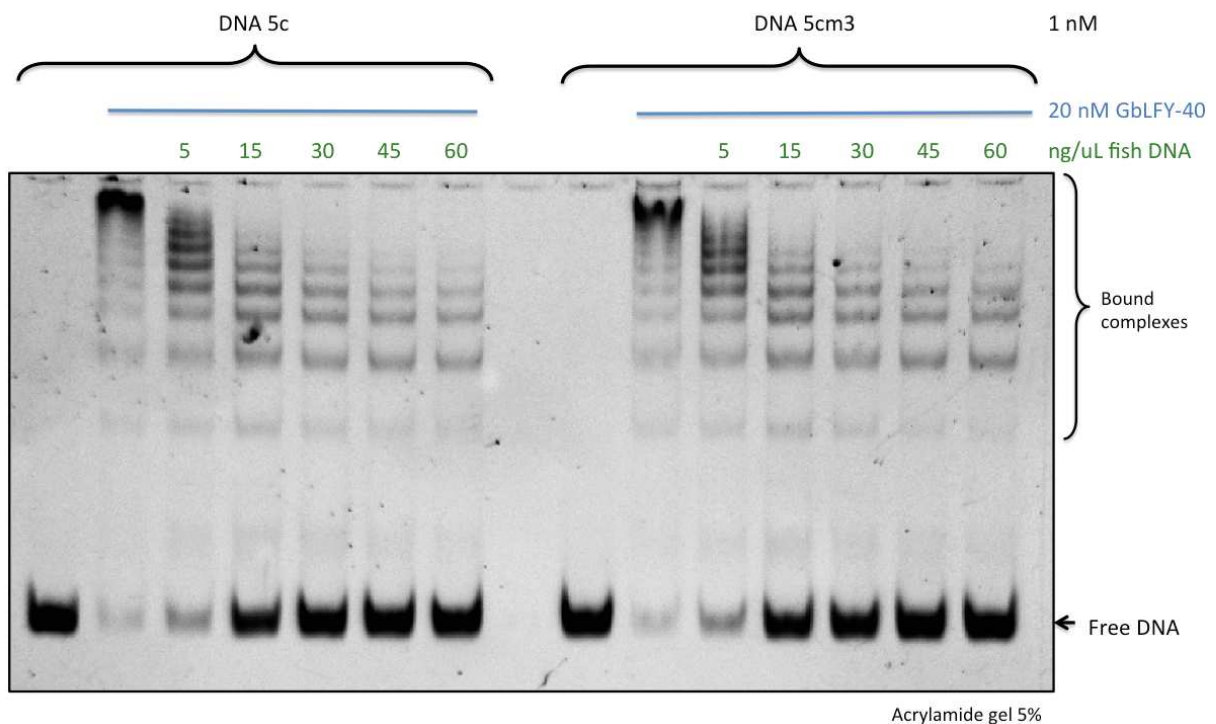


Figure 43. EMSA of free DNA 5c and 5cm3 (as a negative control) with increasing amounts of GbLFY-40.

At all concentrations of competitor fish DNA, no significant differences were observed in the band-shifted pattern, indicating a non-specific binding. The same behavior was observed when tested with the other two sequences and their controls (601_{LFY}, 601, 2 and 2m3). GbLFY-40 binds DNA in our experimental conditions but does not seem to show a preference for sequences containing strong target sites in comparison to the same sequences where the strong binding sites were mutated and only remain low-affinity binding sites. Therefore, GbLFY-40 binds in a non-specific way. EMAs were performed in the same conditions with free DNA probes 2, 2m3 (negative control), 601_{LFY} and 601 (negative control) and similar results were obtained, indicating the GbLFY-40 binds to these probes also in a non-specific way.

The assays were carried out in the presence of genomic fish DNA competitor, a non-labeled DNA, used in order to eliminate part of LFY's non-specific binding. We could observe that DNAs 5c, 601_{LFY} and 2, are shifted at a greater extent by AtLFY-40 than their respective controls DNA 5cm3, 601 and 2cm3, for a given concentration of protein and DNA competitor. AtLFY-40 is able to bind DNA fragments and shows a higher affinity for sequences containing its target sites; therefore AtLFY-40 binds in a specific way. In conclusion, AtLFY-40 does bind specifically to the mentioned DNA sequences but not GbLFY-40.

B. LFY binding to nucleosomes

We next studied the interactions between LFY and nucleosomes. The first question we asked was: is LFY able to directly bind to nucleosomes? In the case of AtLFY-40, we observed practically no binding to nucleosome 601_{LFY}, which is composed of a histone octamer and DNA 601_{LFY} with a strong LFY binding site (Figure 44A). Only free DNA in the nucleosome preparation was completely shifted.

Interestingly, AtLFY-40 is however able to bind nucleosome 2 (Figure 44B). To determine whether the N-terminal region of LFY, which is involved in LFY oligomerization, is required for the binding to nucleosome 2 we performed EMSAs with AtLFY's DBD (Figure 44C). It was interesting to observe that this domain alone is able to interact with nucleosomes. However, it is difficult to compare to the interaction of the complete with nucleosome 2. In any case, this suggests that LFY DBD alone can still bind, to a certain extent, nucleosome 2.

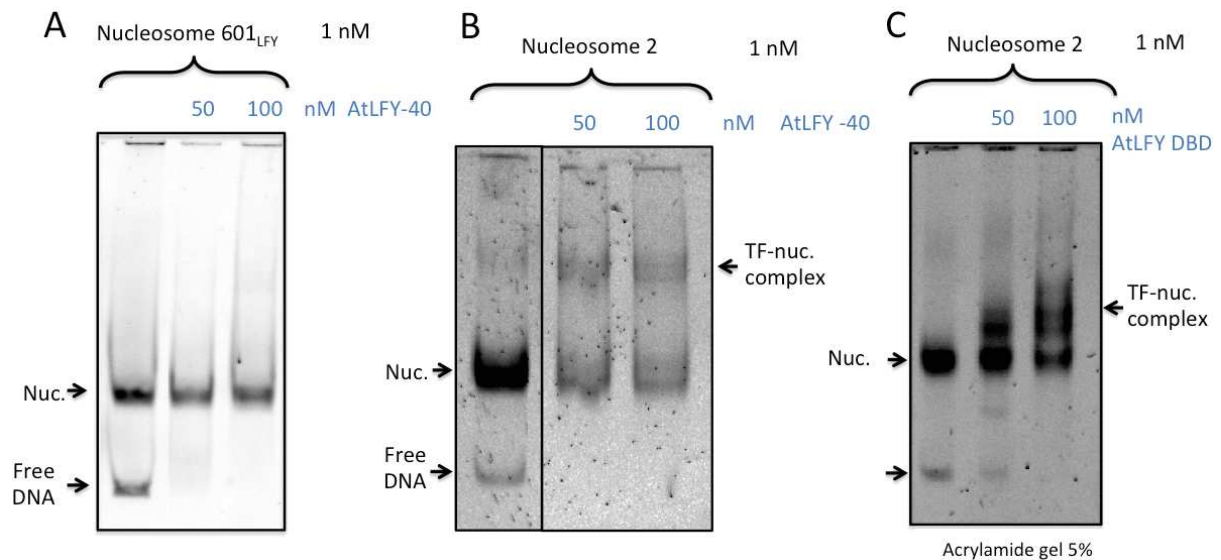


Figure 44. EMSA of nucleosome 601_{LFY} and 2 in the presence of increasing amounts of AtLFY-40 protein.

A | Nucleosome 601_{LFY}, is not bound by AtLFY-40 since the intensity of the band corresponding to the nucleosome does not decrease in the presence of LFY and no new bands are observed that could indicate the presence of a nucleosome 601_{LFY}-LFY complex. However, most free DNA in the nucleosome preparation does disappear in the presence of LFY, indicating that LFY binds DNA. B | AtLFY-40 is able to interact with nucleosome 2. Bands corresponding to higher-order complexes are observed in lanes where LFY was added. C | AtLFY-40's DBD is able to bind nucleosomes.

EMSAs performed in order to assess AtLFY-40's ability to bind nucleosome 5c showed practically no interaction (Figure 45). In this EMSA, different conditions were used to test the binding. Protein was set to a higher concentration and a range of five increasing concentrations of fish DNA competitor was added.

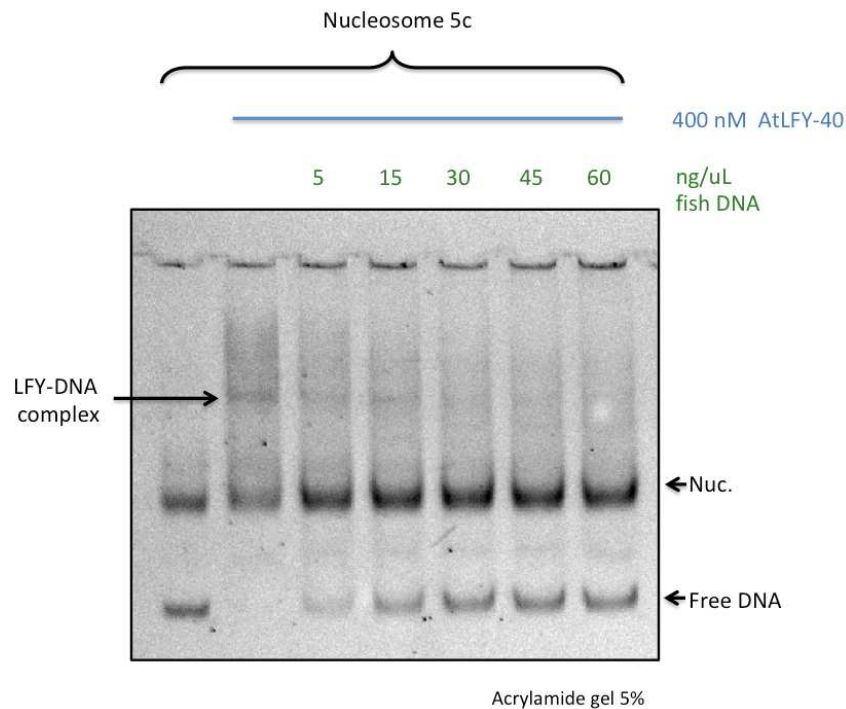


Figure 45. EMSA of nucleosome 5c with AtLFY-40, in the presence of increasing concentrations of fish DNA.

Fish DNA competitor, a non-labeled DNA, was added to the reaction mix in order to eliminate part of LFY's non-specific binding. There is a poor binding of AtLFY-40 to nucleosome 5c. The band corresponding to nucleosome 5c remains practically the same in concentrations of DNA competitor between 15 and 60 ng/μL, indicating almost no interaction with the nucleosome. In the lane with no DNA competitor, as well as in the lanes with 5 ng/μL a weak band can be observed that very likely represents free DNA that has been shifted.

Overall, EMSAs from Figures 44 and 45, show that AtLFY-40 presents almost no interaction with nucleosomes 601_{LFY} and 5c but it does interact clearly with nucleosome 2 and an interaction can still be observed when testing only the protein's DBD. Nucleosome 601_{LFY} is a nucleosome core particle that does not have any free base pairs at the entrance of the nucleosome. Nucleosome 5c has 10 extra bps that protrude from the nucleosome core particle as illustrated before in Figure 41 and finally, nucleosome 2, has 35 base pairs that overhang. This means that AtLFY-40 could be actually binding to this free DNA, since it can bind either as a monomer on a single site (9 bp) and as a dimer, on a semi-palindromic site (19 bp) (Moyroud et al., 2011). This could explain why this protein interacts only with nucleosome 2 but not with nucleosome 601_{LFY} or nucleosome 5c. Further experiments with a shorter version of sequence number 2 (without DNA overhangs) would allow to confirm this interpretation and determine if these extra base pairs play a role in this binding.

After carrying out the previous test for AtLFY-40, we wondered about the other LFY protein's behavior. Is GbLFY-40 able to bind nucleosomes? For this, we tested the binding to nucleosomes 2 and 5c compared to nucleosome 601 without any LFY binding site, which was used as a negative control for this experiment (Figure 46A). Using again a range of DNA competitor to test the specificity we observed more binding to nucleosome 2 and 5c than to nucleosome 601. Afterwards, we found conditions in which all binding to nucleosome 601 is eliminated while binding to the other two nucleosomes was kept (Figure 46B). Indeed, in our experimental conditions at the concentration of 20 nM competitor DNA and 50 nM GbLFY-40 and, this protein is able to bind specifically to nucleosomes 2 and 5c but not to 601.

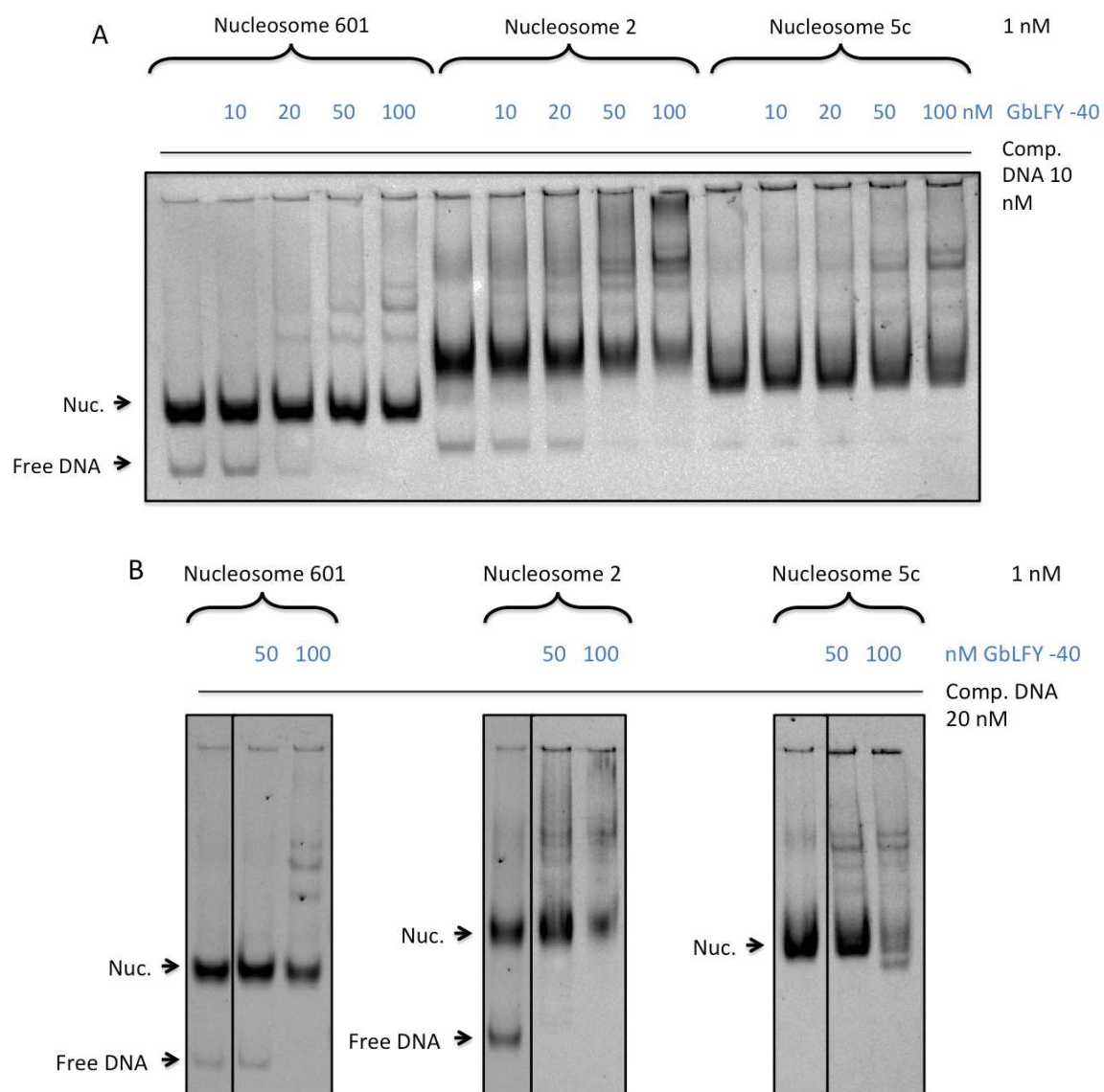


Figure 46. EMSA of the interaction between GbLFY-40 and nucleosome 2 and nucleosome 5c compared to nucleosome 601 (negative control).

A | EMSA showing the different shift patterns of LFY and nucleosomes, in the presence of 10nM of competitor for all the lanes. First, nucleosome 601, then, nucleosome 5c and finally nucleosome 2. The concentration of GbLFY-40 is 0, 10, 20, 50 and 100 nM. B | The same experiment but with 50 nM of GbLFY-40, all binding to nucleosome 601 (control) is eliminated by the competitor DNA. Some shifted patterns still appear in the same condition (20 nM competitor DNA) for the nucleosomes containing a LFY binding site (2 and 5c). This indicates a specific binding of GbLFY-40 to nucleosomes 2 and 5c.

In conclusion, as summarized in Figure 47, AtLFY-40 and GbLFY-40 display different behaviors concerning nucleosome binding. GbLFY-40 is able to bind both nucleosome 2 and 5c while AtLFY-40 is only able to bind nucleosome 2. In the case of nucleosome 2, we cannot rule out that both AtLFY-40 and GbLFY-40 actually bind the free DNA overhangs of this nucleosome particle.

Binding ability of LFY proteins to nucleosomes			
Nucleosome Protein	601 _{LFY}	2	5c
AtLFY-40	✗	✓	Very weak
GbLFY-40	Not tested	✓	✓

Figure 47. Summary of the binding ability of the two LFY proteins AtLFY-40 and GbLFY-40 to nucleosome 601_{LFY}, 2 and 5c.

C. LFY specificity for nucleosomes

Since some PTFs have specific and non-specific interactions with nucleosomes, we looked to answer the question, are the LFY proteins able to bind nucleosomes specifically? Given the results obtained with the set of EMSAs concerning nucleosomes, summarized in Figure 47, we continued working only with nucleosome 5c. This was in order to avoid the possibility of a binding to free overhanging DNA present in nucleosome 2 rather than a

binding to the DNA wrapped around the histone octamer. We looked at AtLFY-40's specificity for nucleosome 5c and its negative control 5cm3, where all strong LFY binding sites were mutated (Figure 48). No differences were observed between the binding to the two nucleosomes (5c and 5cm3) as expected from our previous results where LFY could not bind nucleosome 601, and very weakly nucleosome 5c (both of this nucleosomes, lacking free DNA overhangs). This experiment, confirms that the higher-order bands possibly indicating a weak interaction between AtLFY-40 and nucleosome 5c, actually corresponds to a non-specific binding to free DNA. We performed the same EMSA with GbLFY-40 and we did not observe any specificity for nucleosome 5c as the nucleosome 5cm3 showed a similar pattern in EMSAs (Figure 49).

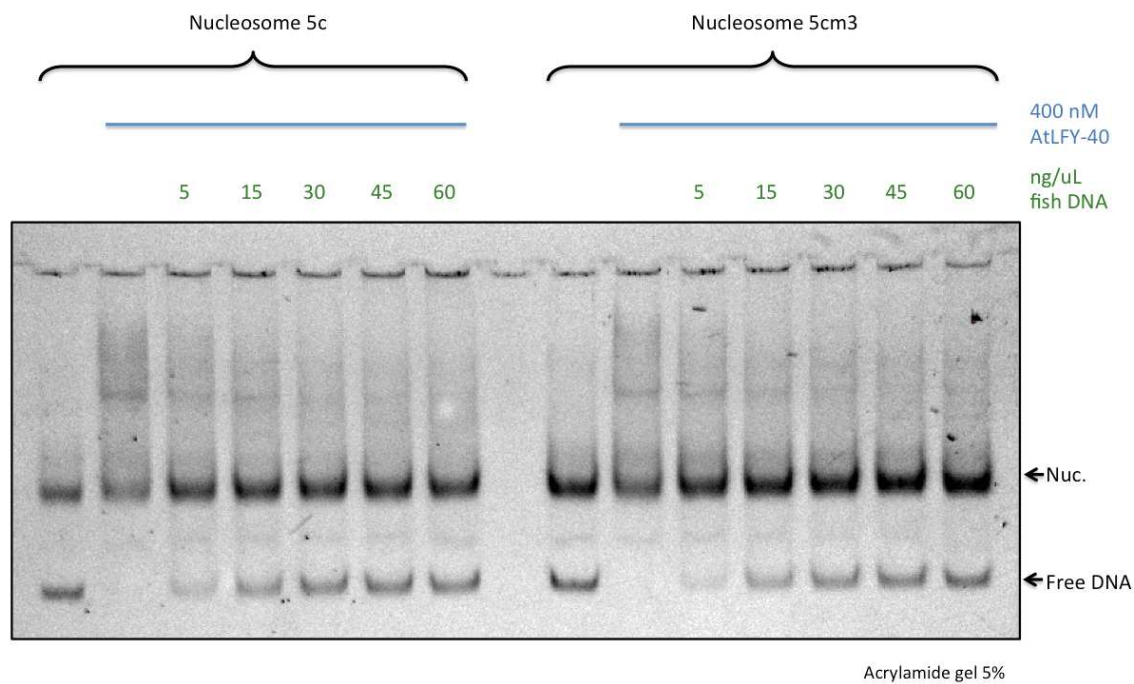


Figure 48. EMSA of nucleosome 5c and 5cm3 (negative control) with AtLFY-40, in the presence of increasing concentrations of fish DNA.

There is a poor binding of AtLFY-40 to nucleosomes when no DNA competitor is used (second lane from the left) and the same pattern is observed with 5cm3 (negative control). This indicates that AtLFY-40 besides binding poorly to nucleosome 5c does not present any specificity.

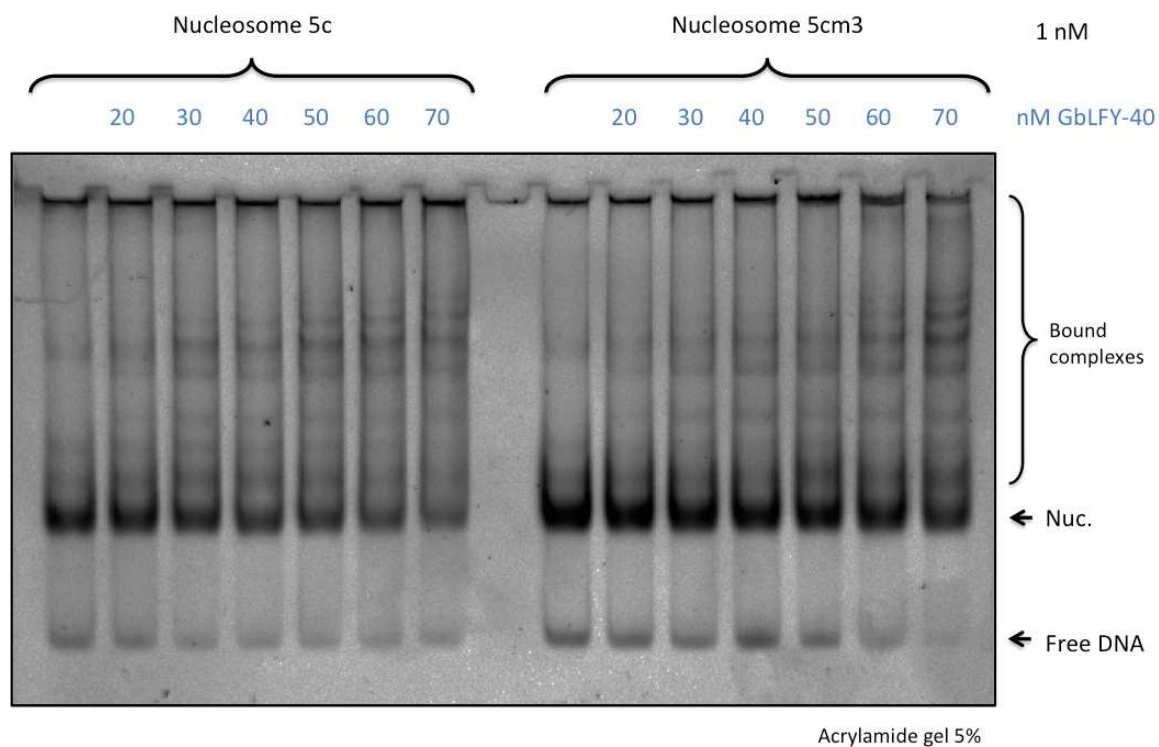


Figure 49. EMSA of nucleosome 5c and 5cm3 (negative control) with increasing amounts of GbLFY-40. No significant difference is observed indicating a non-specific binding. There is no different in the shifted pattern between the two nucleosomes 5c and 5cm3 even though nucleosome 5c contains a strong LFY-binding site. GbLFY-40 is able to bind nucleosome 5c but in a non-specific way.

In conclusion, none of the two LFY proteins were able to bind nucleosome 5c specifically. From all the experiments presented so far some main questions emerge. First, why is there a difference of specificity for DNA between GbLFY-40 and AtLFY-40? (Figures 42 and 43).

In the aim of answering this question, a supplementary EMSA with GbLFY-40 and 5c free DNA was carried out but in different experimental conditions in order to see if we could observe some degree of GbLFY-40 specificity. We chose conditions where DNA was more abundant (100 nM) than protein (up to 60 nM) and an interesting result was observed (Figure 50). Even if the amount of bound DNA by the protein does not change between 5c and 5cm3 we did observe a dominant band in the case of DNA 5c.

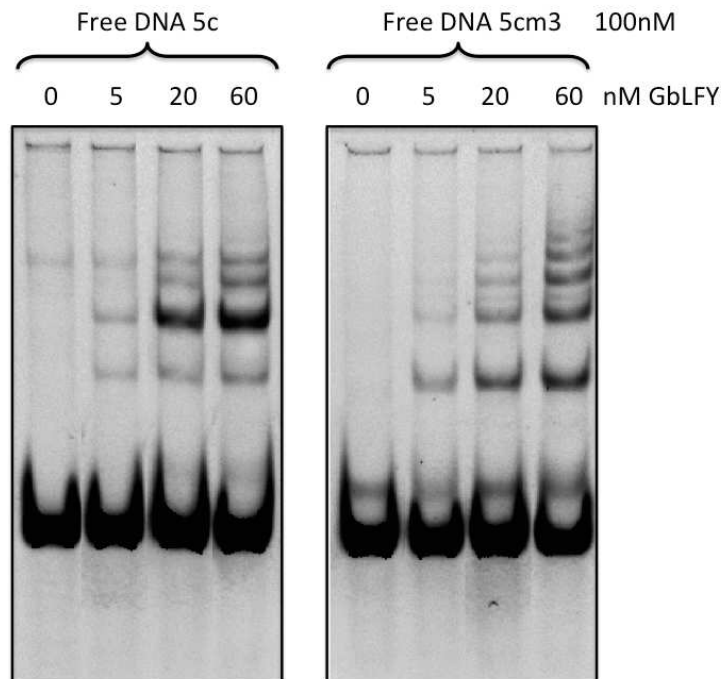


Figure 50. EMSA of GbLFY-40 binding on free DNA 5c and 5cm3 with a high DNA to protein ratio.

A different profile is observed between free DNA 5c and 5cm3. The major band in the condition of DNA 5c at 20 and 60 nM GbLFY-40 could indicate a complex corresponding to the protein bound specifically to the strongest LFY target site within the DNA sequence. On the other hand, in the case of 5cm3 DNA, the ladder-like pattern could represent the formation of various complexes on LFY binding sites with low affinity. Thus, GbLFY-40 does seem to display specificity in these conditions. This experiment reveals the effect of the mutation of the strongest LFY binding site. The fact that we can still see binding of GbLFY-40 to low-affinity binding sites on free DNA 5cm3 could explain why there is no difference in the shifted pattern of nucleosomes 5c and 5cm3.

This could mean that GbLFY-40 is indeed able to specifically bind its target site on 5c DNA but might have a higher affinity for DNA in general (or for low-affinity binding sites) than AtLFY-40. Although the reason for this differential behavior between AtLFY-40 and GbLFY-40 remains unclear, this could explain why in the experimental conditions of EMSAs in Figure 43, we did not see any difference. The same experiment carried out with AtLFY-40 instead, would allow us to definitely conclude about this hypothesis. Also, another useful control would be to perform the same experiment with nucleosomes, instead of free DNA.

D. AtLFY and GbLFY structural differences

The difference in specificity between AtLFY-40 and GbLFY-40 could also be explained by the oligomerization ability and stability of these proteins. We looked in detail at differences in the oligomerization interface of the two proteins. The residues mediating contacts between the two monomers through ionic interactions and hydrogen bonds are not entirely conserved in the two proteins. Out of 9 interactions, three are lost in AtLFY dimers (Figure 51). Therefore, it is possible that GbLFY oligomerization is more stable than in AtLFY. A stronger interaction between the oligomers could lead to a stronger binding on low-affinity sites as a result of a cooperative binding. This phenomenon was already observed at a genome-wide scale as presented in the introduction (Sayou et al., 2016).

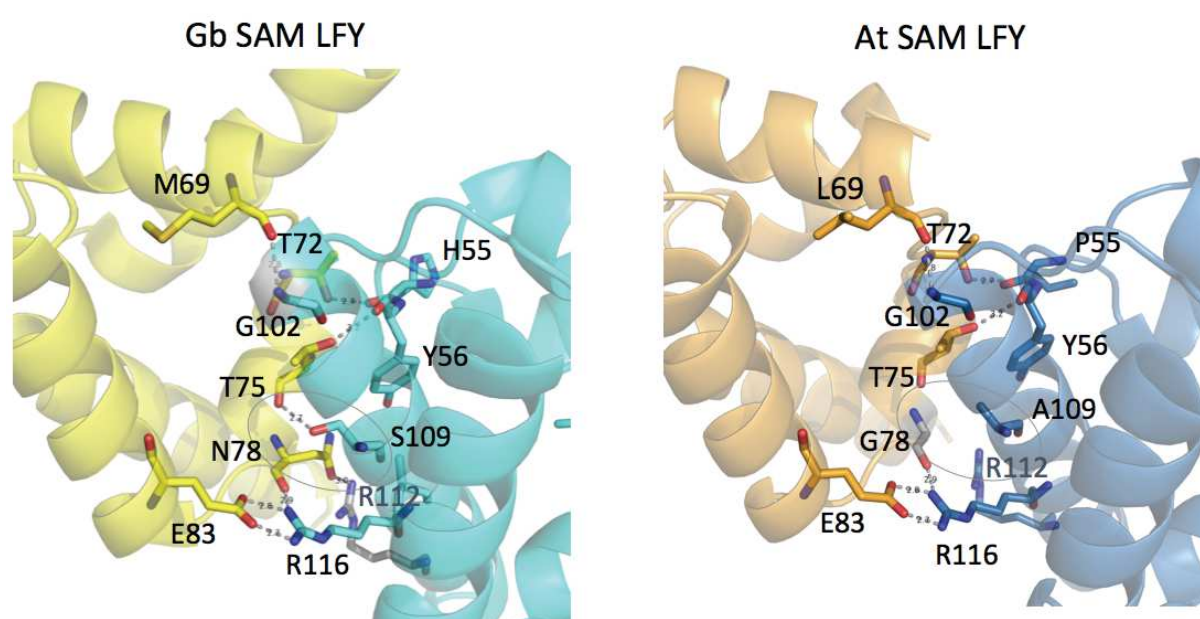


Figure 51. GbLFY and AtLFY oligomerization interface.

In the left, three-dimensional structure of GbLFY (code PDB 4UDE) and in the right, a three-dimensional model of AtLFY's structure generated by the program Phyre2 (www.sbg.bio.ic.ac.uk/phyre2). 3D structure images were done using PyMOL (www.pymol.org). Residues in GbLFY that are not conserved in AtLFY are: M69 > L69, N78 > G78, H55 > P55, S109 > A109. These changes cause three and potentially four less interactions in AtLFY compared to GbLFY, which can display nine interactions. In GbLFY, two interactions concerning N78 are lost in the AtLFY's G78. The interaction between T75 and S109 is lost in AtLFY. Furthermore, in the case of H55 > P55, even though the interacting chain does not change, proline, due to its conformational rigidity, the interaction with threonine could eventually be weaker or lost. Overall, oligomerization in GbLFY seems to be stabilized by more interactions than in AtLFY.

A third factor that could also be involved in this different specificity, could be a difference in the recognition of low affinity binding sites. Predictive position weight matrices were generated for both AtLFY and GbLFY. These matrices determine the DNA-binding preferences of the DBD of each protein using high-throughput Systematic Evolution of Ligands by Exponential enrichment (SELEX). This technique consists of a progressive selection, starting from a large library of double-stranded DNA oligonucleotides of fragments with increased affinity and specificity at each round of selection. Generated logos represent the information contained in the matrix. The matrices are similar for the two LFY proteins, yet they are not the same (Figure 52).

The matrix used for choosing the DNA fragments for nucleosome reconstitution was the AtLFY matrix. Possibly, in the DNA fragments there might be low-affinity binding sites for GbLFY that were not detected for AtLFY. These sites could help GbLFY-40 binding to DNA. It was observed from *in vivo* genome-wide data that oligomerization facilitates binding to regions of low-affinity binding sites (Sayou et al., 2016). This, in addition to the fact that GbLFY can most probably oligomerizes more efficiently than AtLFY, could explain why the specificity of this protein is difficult to observe in our experimental conditions.

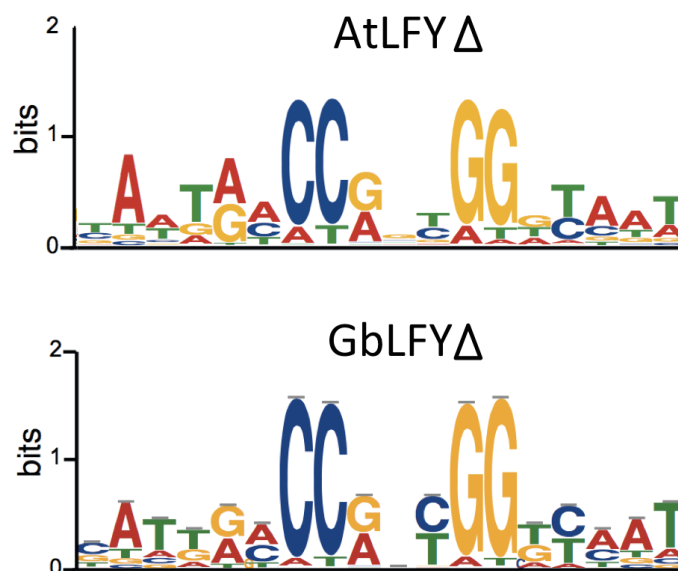


Figure 52. Logos constructed from a biophysical model, built on biochemical data and taking into account the three-dimensional structure of LFY DBD in complex with DNA.

Using this matrix with any 19-bp DNA sequence, scores can be calculated to indicate the affinity of LFY's DBD for the fragment. Logos from both proteins show a similar DNA sequence preference, but existing differences could make a difference in EMSAs (Moyroud et al., 2011).

Besides discussing these practical differences between the two proteins, it is also important to try to understand the meaning of these results at the living-organism level. What could it mean on a biological context, that GbLFY-40 could be able to bind nucleosomes much more strongly than AtLFY? Both proteins have different functions in their respective organisms, AtLFY as explained in the introduction, exerts a function in the specification of floral meristems as well as in flower patterning and this is true for all angiosperms. On the other hand, GbLFY, coming from a gymnosperm, has another function. Even though its role in non-flowering plants is not yet completely understood, some insights exist. Biophysical assays coupled with expression studies, show that central links, necessary for the control of floral development, may already have existed before the appearance of flowers (Moyroud et al., 2017). However, in gymnosperms, LFY alone is not sufficient to confer a reproductive identity (Moyroud et al., 2011).

It is therefore possible that GbLFY could participate in chromatin opening by binding to nucleosomes and without exerting a pioneer function in reproduction. In flowering plants, where LFY acquired this key role, it can also be possible that it requires a partner to bind nucleosomes and therefore, act as a pioneer factor only in concert with other TFs, what Zaret & Mango, (2016), call "a cocktail of TFs that elicit cell fate changes". When studying epigenetic regulation it is today still very difficult to know the chronological order of events that ultimately allow the establishment of a developmental program.

2. OTHER TFs BINDING TO NUCLEOSOMES

GbLFY-40 was shown to be able to bind nucleosomes 2 and 5c *in vitro* through EMSAs (Section III.1.B.). But, are other general TFs able to bind our nucleosome core particles in our experimental conditions? To answer this question, we looked at the ability of unrelated TFs to bind nucleosomes, in order to compare it with GbLFY-40's ability. We tested two other transcription factors: REGULATOR OF AXILLARY MERISTEMS 1 (RAX1) a Myb domain-

containing protein important for meristem formation (Denay et al., 2017), and ULTRAPETALA INTERACTING FACTOR 1 (UIF1), also a Myb domain-containing protein that contributes to the control of floral meristem activity in Arabidopsis (Moreau et al., 2016). We did not observed interaction of these two transcription factors to nucleosome 5cm3 (Figure 53). This is consistent with our hypothesis that not any general TF can directly bind nucleosomes.

In order to test if these two proteins are able to bind their respective DNA target sites on free DNA probes, an EMSA was performed in the same conditions. We used fragments from *CLAVATA-1* (*CLV-1*) and *WUSCHEL* (*WUS*) promoters as the target probes of each of RAX1 and UIF1, respectively. As shown in Figure 54, both proteins can bind their corresponding target.

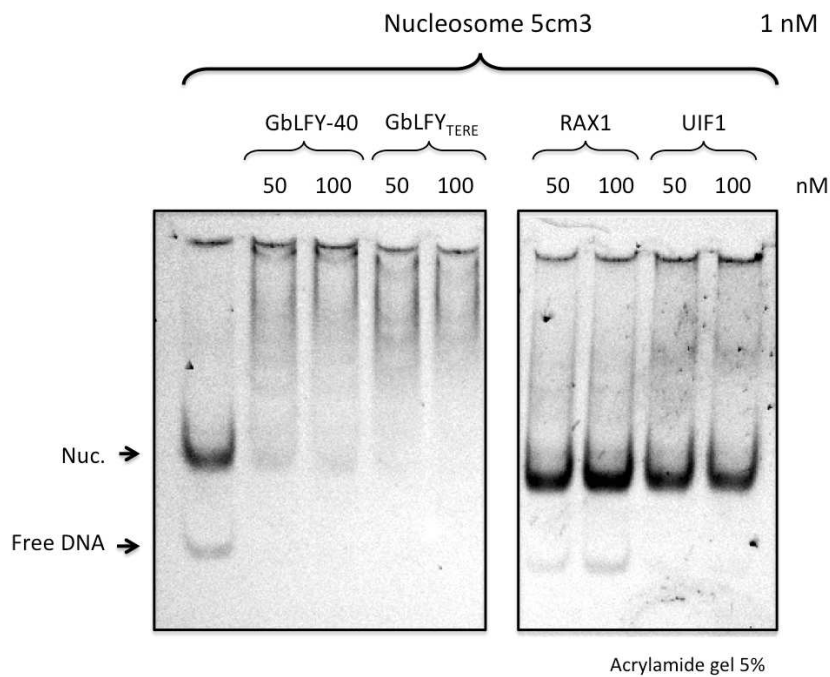


Figure 53. EMSA of nucleosome 5c against four TFs.

Neither RAX1 nor UIF1 seem to bind nucleosome 5c, since no higher order complex is formed. On the contrary, with GbLFY-40 and GbLFY_{TERE} higher order complexes are formed as a ladder-like pattern and they are able to completely shift the nucleosome band.

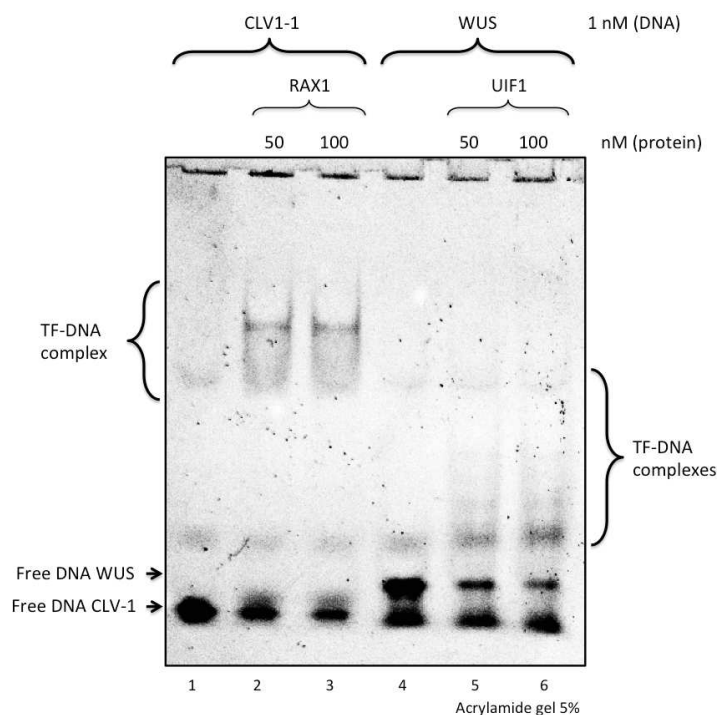


Figure 54. EMSA with *CLV-1* and *WUS* probes, the respective specific DNA targets for RAX1 and UIF1 proteins. This experiment was performed as a control to test that both of these TFs are functional and able to bind their cognate DNA sequence in our experimental conditions. Compared to lanes 1 and 4, where no protein was added, most of DNA is shifted at 100 nM, and TF-DNA complexes are formed (indicated on the gel) showing that both TFs are able to bind its target DNA in the EMSA conditions.

In conclusion, both RAX1 and UIF1 TFs do not seem to bind nucleosomes, which confirm our hypothesis that general TFs may not have a pioneer function and should not have the capacity to interact with nucleosome particles. However, this result is preliminary and more experiments need to be done in order to come to a clear conclusion. For instance, comparing binding of RAX1, UIF1 and GbLFY-40 to nucleosome 601. Also, having a nucleosome positioning sequence containing a specific binding site for RAX1 and UIF1 would allow to confirm their inability to bind nucleosomes.

3. EFFECT OF LFY'S BINDING ON NUCLEOSOMES' STRUCTURE

We tried to study the effect of LFY's binding on nucleosomes in their stability. What is the consequence of LFY's interaction for the nucleosome structure? Does LFY provoke an

ejection of the histone octamer? Does the nucleosome stay in its original form? To tackle these questions, core histones were labeled with a Cy3 dye. This allows tracking histone migration in addition to DNA migration in EMSAs. When Cy3 and Cy5 nucleosomes were incubated with GbLFY-40, we could observe a common band between the Cy5 and Cy3 channels (yellow arrow in Figure 55A right and 55B). As expected this band was not observed when histones were not present (Figure 55C). This indicates that a nucleosome-LFY complex can be formed without nucleosome disassembly.

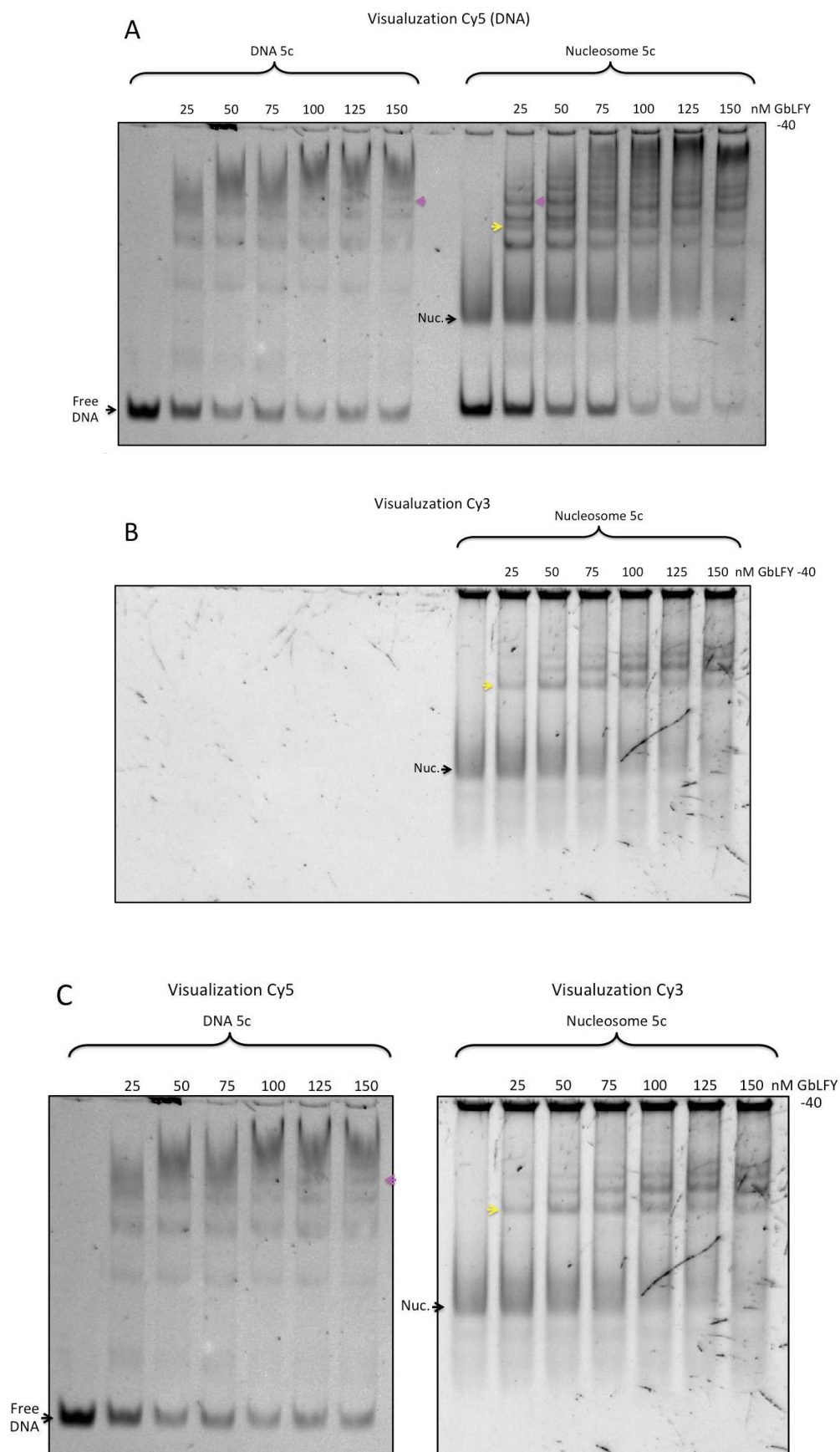


Figure 55. EMSAs of Cy5-labeled DNA 5c and nucleosome 5c (containing Cy3-labeled histones) with increasing amounts of GbLFY-40.

A | Cy5 visualization to detect labeled free DNA and labeled nucleosomes. The yellow arrow indicates a band that can be also found in the Cy3 visualization channel (see B). The magenta arrows indicate a band that can be found in both LFY-free DNA as well as LFY-nucleosome interaction reactions. B | Cy3 visualization of the gel, allowing to see labeled core histones. The yellow arrow indicates a band that can be also found with the Cy5 visualization channel (see A). C | Comparison of the free-DNA-containing samples visualizing Cy5-labeled DNA against nucleosome-containing samples visualized Cy3-labeled histones. This comparison allows to clearly see that LFY-nucleosome complexes migrating at a specific size are not found in the LFY-DNA complexes, indicating LFY binding to nucleosomes with no histone octamer ejection.

4. ROLE OF LFY'S OLIGOMERIZATION

Another important aspect of our study was the effect of the protein's oligomerization ability on nucleosome binding. In Sayou et al., (2016), we discovered that the SAM oligomerization domain enables LFY to bind to closed regions of chromatin in Arabidopsis. The DNA sequences that we selected for nucleosome reconstitution such as DNA 2 or 5c were chosen for being bound by LFY but not LFY_{TERE} (Sayou et al., 2016). A question that emerged from this observation was: can we also observe a loss of LFY_{TERE} binding *in vitro* at the scale of a single nucleosome?

For this, we used the mutant GbLFY_{TERE} protein. We compared binding of GbLFY-40 and GbLFY_{TERE} to nucleosome 5c (Figure 56A). A slight binding decrease of GbLFY_{TERE} compared to GbLFY-40 was observed, suggesting that LFY's oligomerization domain is also involved in the binding to nucleosome. We compared the shifted patterns to an EMSA done by Sayou et al., (2016) with a short DNA fragment containing two LFY-binding sites (Figure 56B). Schematic representations of the molecules are shown in Figure 56C.

It seems that between 10 and 75 nM of GbLFY_{TERE} we observe the same complexes, while in the case of GbLFY-40, lower-size complex quickly disappear with increasing protein concentration, to favor the formation of higher-order complexes

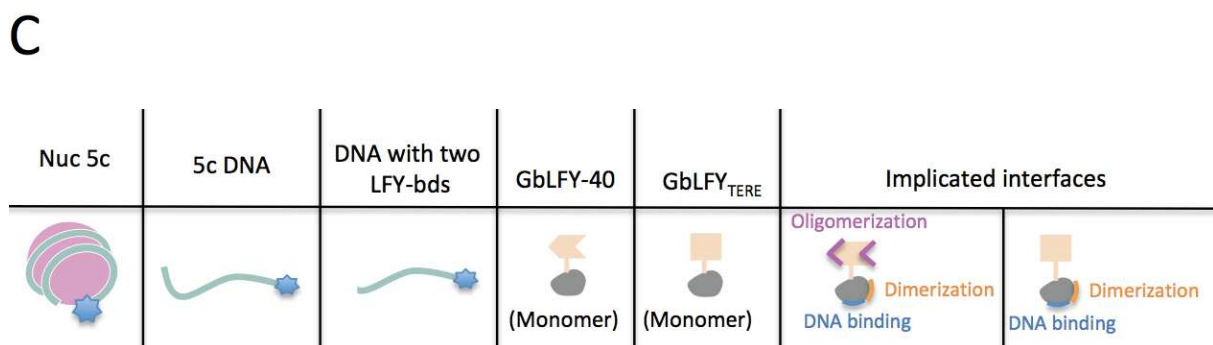
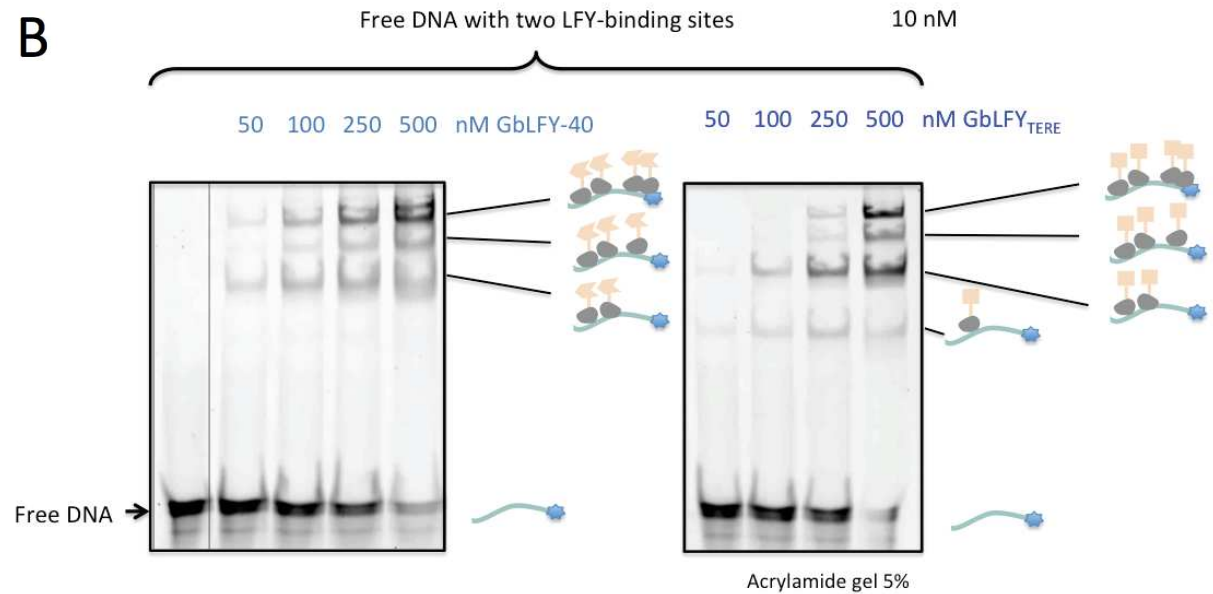
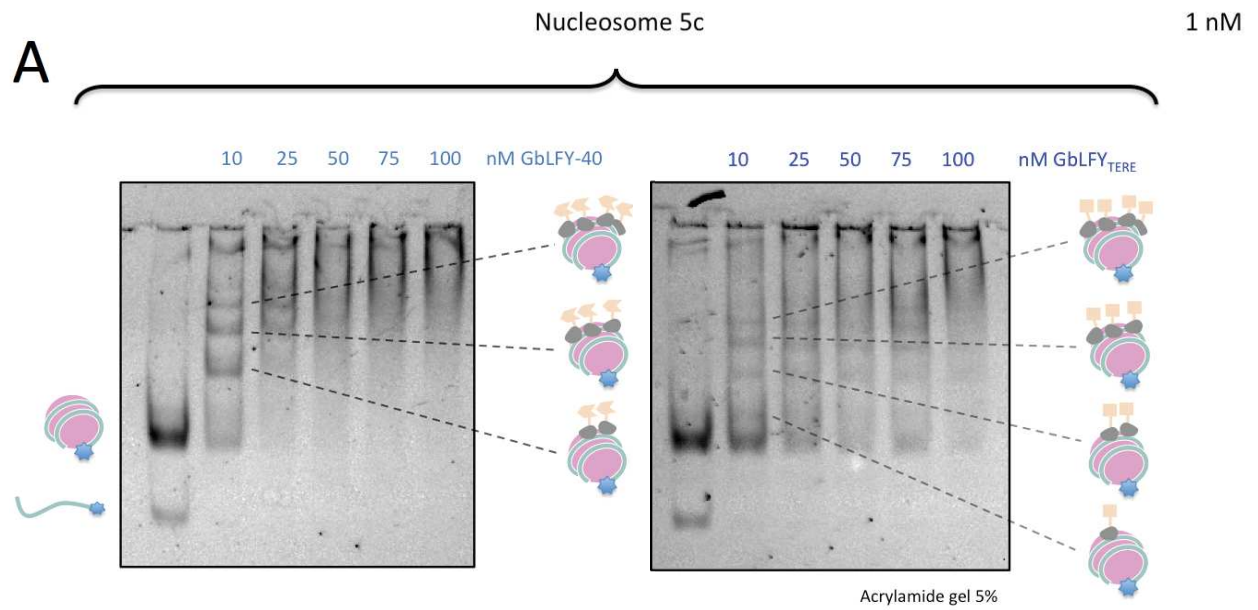


Figure 56. EMSA of GbLFY-40 and GbLFY_{TERE} binding comparison.

A | Nucleosome 5c in the presence of increasing concentrations of both proteins. The formation of higher-order complexes is favored with GbLFY-40 because of its oligomerization capacity. GbLFY-40 seems to form higher-order complexes at a concentration of 50 nM, while with the oligomerization mutant (GbLFY_{TERE}) form these types of complexes only at a concentration of 100 nM. On the right side of the gel, representation of LFY protein binding on nucleosomes as a monomer, dimer, a combination of dimer and monomer, and as a tetramer. B | EMSA from Sayou et al., (2016) of a DNA fragment containing two LFY-binding sites in the presence of increasing concentrations of both proteins. On the right side of the gel, representation of the same kind of complexes that the ones shown in Figure 56A, since a similar shifted pattern is observed. C | Diagram to indicate representations of nucleosomes, DNA and proteins. These EMSAs could suggest some behavior similarities between GbLFY-40 and GbLFY_{TERE} with respect to nucleosomes (A) free DNA probes (B), since both seem to share shifting patterns indicating a cooperative binding.

The complexes formed with nucleosomes and DNA reflects LFY capacity to have a cooperative DNA binding. This has already been proposed for proteins containing a SAM domain (Green et al., 2010). The EMSAs from Figure 56 showed that GbLFY_{TERE} was not as efficient as GbLFY-40 at forming the tetrameric or higher-order complexes, which are favored only at the highest concentration (100 nM) of GbLFY_{TERE}. We concluded that LFY cooperatively binds on multiple sites, via oligomerization of its SAM domain on nucleosomes as it does it on free DNA.

IV. DISCUSSION

During this first chapter I presented several studies from the literature made essentially with EMSAs to assess PTFs. Through different examples, we can retain some important pioneer characteristics of a TF. For instance, the TF have to bind on one side of the DNA. Also the flanking sequence of a TF target site positioned on a nucleosome can have an effect on its binding affinity. Moreover, Lone et al., (2013), showed that the position of the target site within a nucleosome can also have an impact on TF's binding, since not every position is equally accessible. Another important aspect, as observed with the paradigm PTF FoxA1, is that binding to nucleosomes can be composed of specific and non-specific interactions.

We decided to study a plant TF: LFY, whose potential pioneer molecular mechanism especially in relation to chromatin is a fascinating open question. Regarding the *in vitro* approach, one of the first observations was that both proteins AtLFY-40 and GbLFY-40 can bind free DNAs that contain LFY target sites but AtLFY-40 shows a higher specificity than GbLFY-40 in conditions of high protein concentration in relation to DNA concentration (Figure 50). Surprisingly, EMSAs where DNA was more abundant than protein, showed that, when testing GbLFY-40 against free DNAs 5c and 5cm3, a predominant band in the case of 5c DNA is observed, meaning that GbLFY-40 is actually able to bind its target site on 5c DNA in a specific way but might have a higher affinity for DNA in general than AtLFY-40.

These differential DNA affinities between AtLFY-40 and GbLFY-40 could be explained by their structural differences and more specifically, to the improved oligomerization capacity that GbLFY-40 could have with respect to AtLFY-40. A stronger oligomerization ability could confer GbLFY-40 a facilitated binding to DNA and nucleosomes, especially concerning regions with several low-affinity binding sites where cooperative binding plays an important role. Another possible explanation for this, is the fact that GbLFY might display slightly different affinities than AtLFY for a given binding site on a DNA fragment. Since the DNA sequences for nucleosome reconstitution were designed based on AtLFY matrix, there might be some "hidden" low-affinity binding sites for GbLFY within them.

Regarding nucleosome binding, it was observed that AtLFY-40 is able to bind nucleosome 2 but not nucleosome 5c. This could suggest that LFY binds to the DNA overhangs of nucleosome 2 and does not directly bind to a nucleosome core particle, such as nucleosome 5c (DNA fragment of 157 bp) and 601 (DNA fragment of 147 bp). Another explanation might be that binding sites on nucleosome 5c are not as accessible for LFY as they are in nucleosome 2. Further experiments with a shorter sequence 2 with no DNA overhangs could be useful to shed light on this question.

Concerning the binding of AtLFY-40 and AtLFY-40 DBD to nucleosome 2. We could hypothesize that other domains of the proteins apart from the DBD are implicated in nucleosome binding. Possibly, the disordered intermediate region between the N-terminal domain and the DBD, which has an acidic isoelectric point, could be implicated in core histone binding. For instance, the pioneer transcription factor EBF1, has been shown to induce DNase hypersensitivity on chromatin through its unstructured C-terminal domain (Boller et al., 2016). Also, the PTF HNF3, contains a core histone binding motif, which also lacks an ordered structure and contains few acidic residues (Cirillo et al., 2002). In the case of LFY, it would be useful to explore this line of research and see whether this protein contains a histone-binding motif. For example, *in vitro* biophysical techniques, such as HTRF, could allow mapping a potential histone-binding domain in LFY.

Additionally, we wanted to know if other TFs not suspected to have a pioneer activity, were able to bind nucleosomes. To do this, we compared the non-specific binding component observed with GbLFY-40, to the RAX1 and UIF1 TFs. Interestingly, our preliminary results show that, neither of them, seem to have the ability to target nucleosomes in a non-specific way, while they are able to specifically bind their target sites on free DNA probes.

Finally, when assessing the effect of oligomerization on LFY binding on nucleosomes, we discovered a behavior similar to the one seen with free DNA. GbLFY-40 was able to quickly form higher-order complexes compared to GbLFY_{TERE}. We concluded that LFY cooperatively binds on multiple sites, via oligomerization of its SAM domain on nucleosomes as it does on free DNA.

Undeniably, EMSA is a technique that presents some drawbacks. Dissociation can occur during electrophoresis, therefore affecting the interpretation of detection. Additionally, this kind of assay does not provide an accurate measure of the size or entities of the DNA-protein complexes since mobility in gels is influenced by several factors other than molecular weight, such as its three-dimensional conformation. Also, EMSA does not directly provide information on the nucleic acid sequence the proteins are bound to. Because of these drawbacks, it will be interesting to continue studying the LFY-nucleosome interaction by other techniques such as DNase footprinting assays or single-molecule techniques such as AFM.

Globally, for this kind of EMSAs, proteins require to be highly pure. For our AtLFY-40 protein, which is really our protein of interest, purification always results in samples still containing some bacterial DNA. In our conditions, it would be of interest to use a positive control with a confirmed PTF and a nucleosome with an inserted target site for this PTF next to LFY binding sites. This would allow us to draw conclusions more clearly.

CHAPTER II. TWO LFY PARTNERS: CHROMATIN REMODELLING COMPLEXES SPLAYED AND BRAHMA

This chapter specifically concerns the study of the synergy between LFY and the two SWI/SNF ATPase subunits SPLAYED (SYD) and BRAHMA (BRM). As mentioned in the introduction, these two ATPases are required for flower development and show overlapping functions during Arabidopsis development. They are recruited to the regulatory regions of the floral homeotic regulator genes *AG* and *AP3* and are necessary for their expression. These regulatory regions have been observed to be subjected to polycomb repression at earlier developmental stages (Goodrich et al 1997). The direct interaction of these SWI/SNF ATPases with LFY has been assessed through genetic and biophysical approaches. Furthermore, their association to *AG* and *AP3* loci overlaps in time and space with LFY binding. Also, it has been shown that SYD binding to these regions is strongly affected in *lfy* mutants; while the reverse is not true. These main elements support the model of recruitment of SYD and BRM by LFY (Wagner & Meyerowitz, 2002; Wu et al., 2012).

The information presented above, fits into our hypothesis that LFY could act as a pioneer transcription factor at least through the recruitment of CRCs to closed regions of chromatin in order to open them and initiate developmental programs. During my thesis project I worked to further elucidate the molecular basis of the physical interaction between LFY and the two ATPase subunits: BRM and SYD. The first part of this section will be about the *in vitro* aspect of the project, which has the objective of finely map the interacting domains of, on one hand, LFY and on the other hand, SYD and BRM. In the second part, I will present the work done in order to better understand the interactions of these partners at the genetic level. For this, we decided to work with the Arabidopsis 35S:LFY line, in which LFY is ectopically expressed to analyze the effect in *SYD* and *BRM* mutant backgrounds to see whether LFY was still able to function. Our main questions were, are we able to confirm and more finely map the interacting regions of these partners? What is the effect of *syd* and *brm* mutations on the phenotype of plants over-expressing LFY?

I. GENETIC, BIOCHEMICAL AND BIOPHYSICAL CONTEXT FOR A COMPLEX INTERPLAY

- Background information concerning LFY interactions with SYD and BRM.

In previous work, SYD and BRM were shown to interact with LFY (Wu et al., 2012). Y2H assays showed a weak interaction between the N-terminal part of SYD and BRM with LFY. This assay is based on the reconstitution of a functional TF when two proteins come into proximity. For this, genetically modified yeast strains are used, in which the transcription of a reporter gene, when activated, allows the growth on a selective medium. These interactions were confirmed in pull down assays (Figure 57).

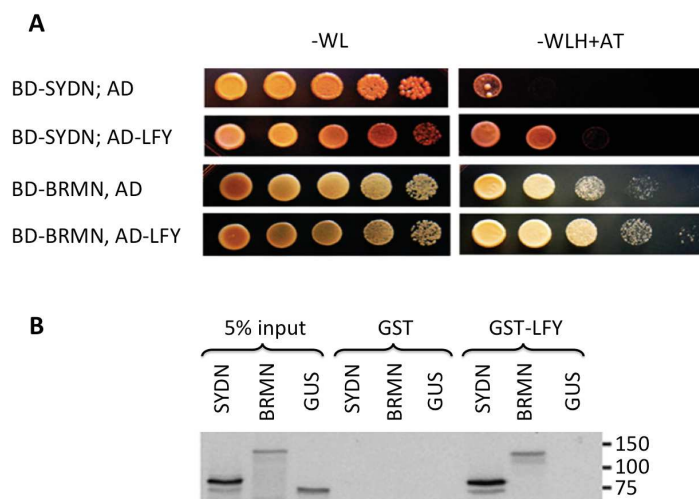


Figure 57. Y2H and pull down assay showing interactions of SYD and BRM with LFY.

A | The N-terminal regions of SYD or BRM (SYDN or BRMN) interact weakly with LFY in yeast two-hybrid assays. On the left, growth on permissive medium lacking tryptophan and leucine (-WL; not selecting for interaction). On the right, growth on selective medium that selects for interaction by further omitting histidine (-WLH and adding 0.1 mM 3-AT for SYD or 0.5 mM 3-AT for BRM). BD: binding domain, AD: activation domain. **B** | GST-LFY pull-down assay showing physical interaction between LFY and the [35S]methionine-labeled SYDN and BRMN. On the right, approximate molecular mass in kilodaltons (Wu et al., 2012). Both assays show an interaction between the N-terminal of BRM and SYD with LFY complete protein. However, on Y2H assays interactions are weak and background signal is observed.

The authors propose that this binding occurs via the helicase/SANT-associated (HSA) and the HSA-adjacent (A-HSA) domain, based on further pull-down assays carried out with different deletions of the SYD and BRM proteins, where they classified the binding strength of several BRM and SYD constructions against LFY (Figure 58).

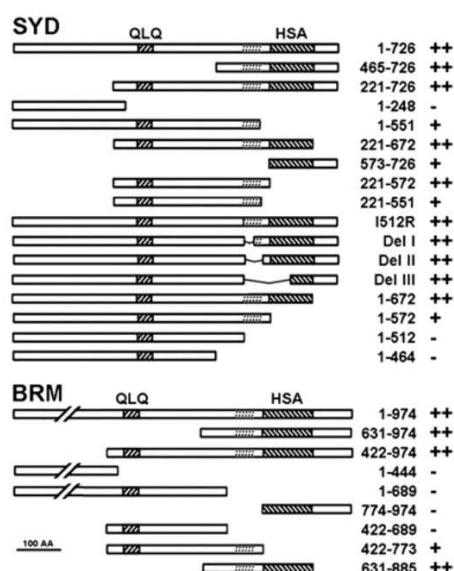


Figure 58. Schematic representation that summarizes pull down assays between LFY, SYD-N and BRM-N.

Interactions for each construct were classified as « strong », « weak » or no interaction with ++, + or -, respectively (Wu et al., 2012). For both proteins, the A-HSA and HSA domains seem to be the most important for an interaction with LFY.

Based on this, one of the objectives of my thesis project was to confirm, and to further finely map which region of LFY is involved in the interaction with BRM and SYD, and vice versa, which is the minimal domain of BRM and SYD that can bind LFY. Ultimately, this could lead to the crystal structure of a complex between a TF and an ATPase sub unit of a CRC, a structure that has not been observed yet for any other transcription factor.

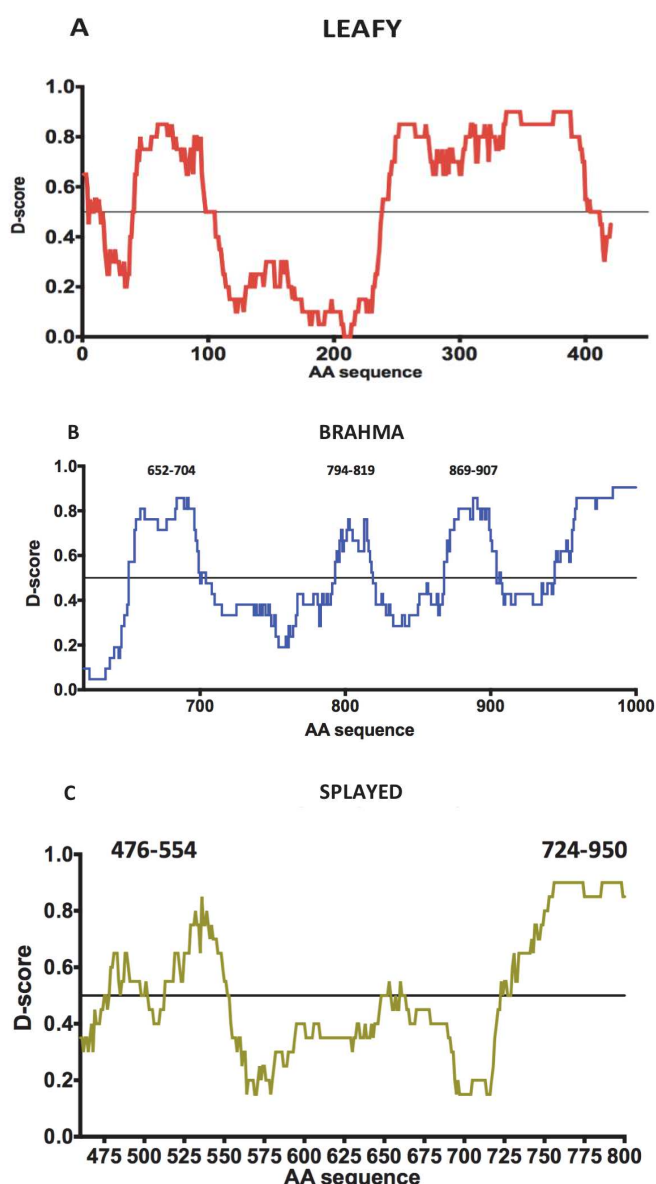
1. PRODUCTION OF RECOMBINANT PROTEINS FOR INTERACTION ASSAYS

In order to answer our question of the minimal interacting domains of LFY, BRM and SYD, we designed different protein constructs. How could we design adequate constructs for our interaction tests? Two different strategies were taken into account to design constructs.

A. Construct design of SYD, BRM and LFY deletions

Regarding LFY, constructs were mainly designed to present deletions from the C-terminal and carry a GST tag fused to the N-terminal side, taking into consideration the ordered and disordered domains of the protein (Figure 59A). For this, we used the sequence analysis made by Francine Gérard and Jean-Marie Bourhis (GRAL collaborators). They performed the analysis for the three proteins: LFY, BRM and SYD, assigning a D-score for each amino acid. The D-score represents a “disorder” value. A region with a D-score above 0.5 means that it is predicted as ordered. This is possible thanks to a meta-analysis: the amino acid sequence is entered into a server that uses different algorithms accessible through WEB servers (Foldindex, PONDR, NORSp, IUPred, DISOPRED, DisEMBL, RONN, PreLink97 and DRAWHCA (Gérard et al., 2009)) in order to generate a D-score. Some algorithms take into account chemical properties of the amino acids. Thus, the D-score can give a global idea of disordered and ordered regions of a protein (Figure 59).

For BRM and SYD, the D-score analysis and ProtParam online tool, which is a web-based service for protein structure prediction (<https://web.expasy.org/protparam/>) helped to verify the likelihood that constructs would be soluble. Constructs were mainly selected around the A-HSA and HSA domains (Figure 59B and C).



The calculated D-score is displayed on the y-axis and amino acid sequence on the x-axis. The horizontal line at 0.5 represents the boundary between ordered and disordered regions in BRM and SYD from Arabidopsis.

axis and amino acid sequence on the x-axis. The horizontal line at 0.5 represents the boundary between ordered and disordered regions. A | LFY (1-420 aa), ordered regions correspond to the conserved SAM domain and the DBD domain. B | BRM (1 - 1000 aa), the predicted ordered region between aa 794 to 819 is contained in the HSA domain which goes from 713 to 885 aa. C | SYD (450 – 800 aa), A-HSA domain and HSA domains go from aa 512 to 672. Only a small part, corresponding to the A-HSA domain is predicted as being ordered region in SYD. Predicted ordered regions of the proteins, were taken into account in the design of different constructs, in concert with ProtParam, a web-based service for protein structure prediction.

B. Protein production of SYD, BRM and LFY constructs

In collaboration with Emmanuel Thévenon (from our team), we tried to clone, produce and purify the designed constructs in different buffers to find the optimal conditions for protein stability. Figure 4.7, shows designed protein constructs. Also in collaboration with Emmanuel Thévenon, Thermal Stability Analysis (TSA) was performed. TSA is a useful technique that consists in measuring the temperature at which a protein unfolds in a given buffer, called melting temperature. The buffer, in which the protein has the highest melting temperature, would be considered as the optimal buffer for the protein's stability. As shown

in Figure 60, the production of many constructs was successful and we worked with the ones that were produced.

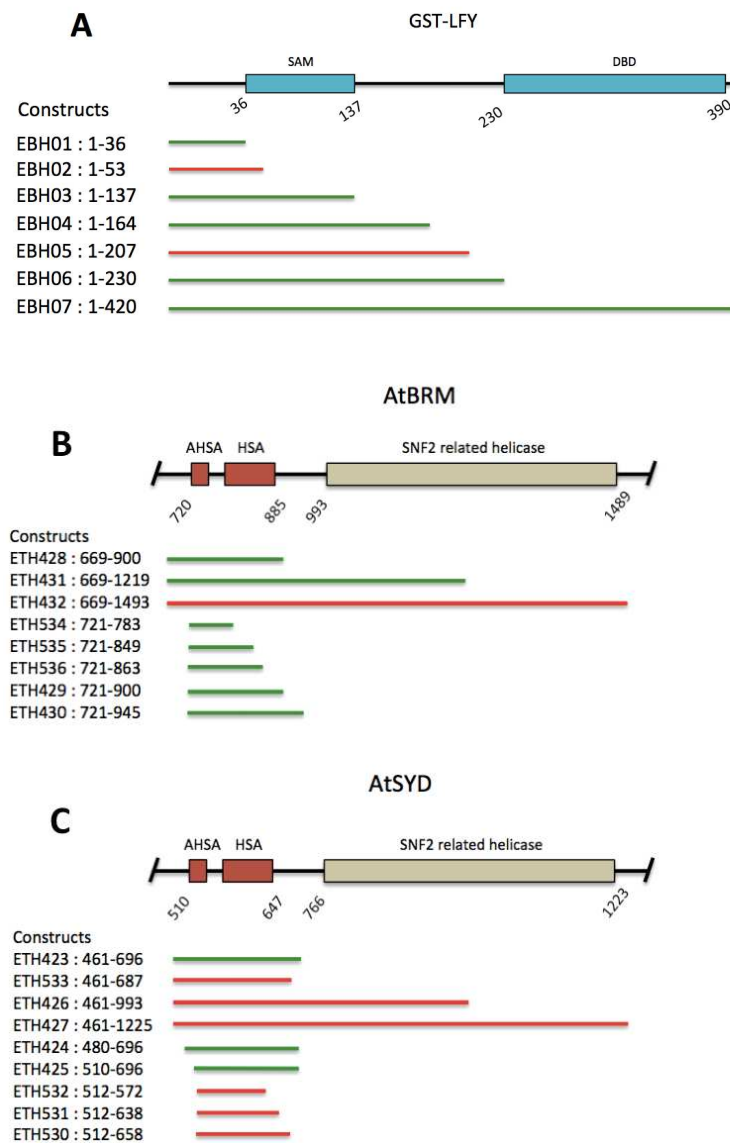


Figure 60. Constructs designed for LFY, BRM and SYD.

For BRM and SYD, constructs were designed and around A-HSA and HSA domains based on D-score analysis. In green, constructs that we were able to clone and produce, in red constructs that were not produced. A | Representation of the GST-LFY constructs. LFY construct 1 (EBH01) consists of a disordered region before the SAM domain. LFY construct 3 (EBH03), goes up to the end of the SAM domain. LFY construct 4 (EBH04), goes from up to the middle disordered region between the two LFY conserved domains. In LFY construct 6 (EBH06), the DBD is absent. LFY construct 7 (EBH07), the full-length protein. B | Constructs for BRM. Several proteins were designed to cover the A-HSA and HSA domains at different regions. Most of them were successfully cloned with a Histidine tag at the N-terminal. C | Constructs for SYD. Similarly, constructs covering the A-HSA and HSA domains were designed carrying the same tag as BRM constructs.

C. Protein purification of SYD, BRM and LFY constructs

Four truncated version of LFY as well as the complete LFY protein were purified (Figure 61). However, the full-length LFY-GST fusion (construct 7) displayed degradation during the purification on the Glutathione-sepharose column. BRM constructs ETH428 and ETH429, were obtained as a soluble form after trying several buffers (Figure 62A and B). In a similar way, some SYD constructs were obtained as a soluble form (Figure 62C). We focused our work in testing first interactions between LFY and BRM, since BRM is easier for manipulation than SYD.

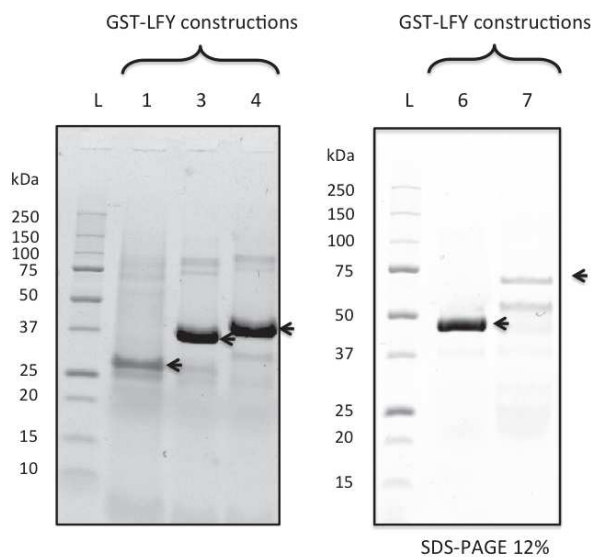


Figure 61. Affinity tag purification of GST-LFY proteins. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) of the purified LFY proteins containing a GST-tag, bound to a Glutathione-Sepharose column. Protein's migration corresponds to their expected sizes. For LFY full-length, some degradation is observed (size expected for the complete GST-LFY, 73.9 kDa).

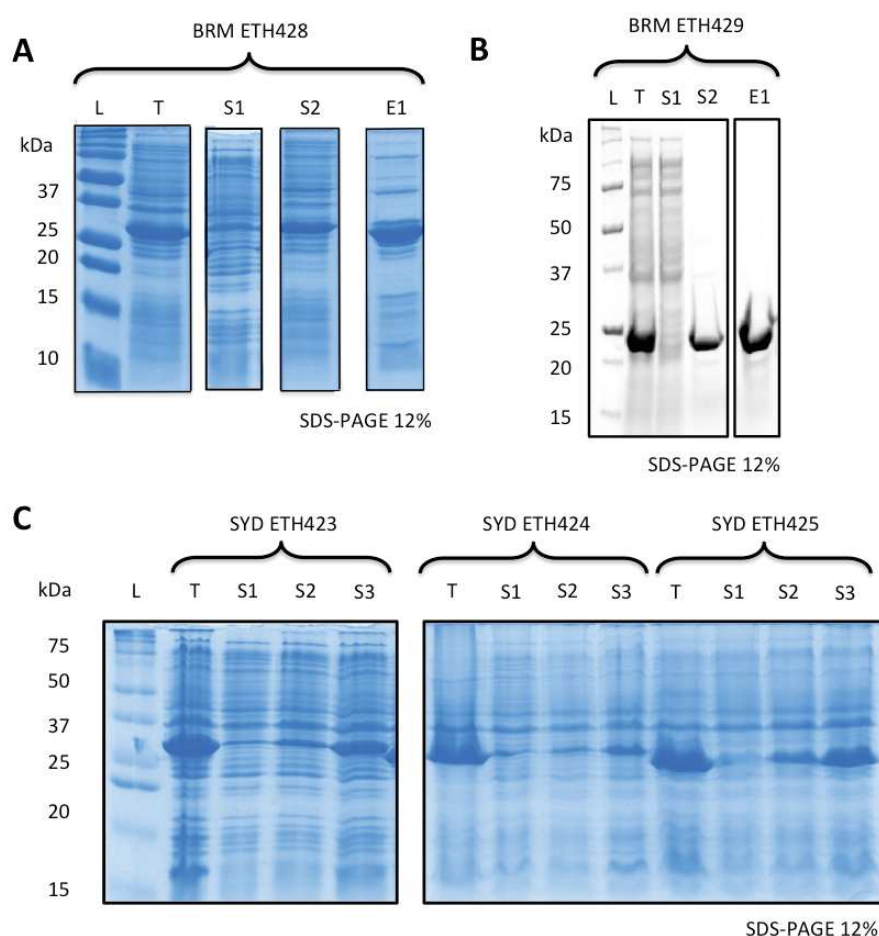


Figure 62. Solubilisation and purification of some of the BRM and SYD constructs produced.

All gels are SDS-PAGE 12% stained by coomassie blue. For all of the five proteins shown, the first buffer tested for solubilisation was Tris-HCl 20mM pH 8, TCEP 1mM. The total fraction is a sample of protein after sonication in this buffer. The total fraction is only presented for the first solubilisation buffer as a reference. The total fractions after sonication with the other tested buffers are not shown. A | Purification on a Nickel-Sepharose column of the BRM construct ETH428 containing a histidine-tag. L: Protein ladder. T: total fraction. S1: Soluble fraction resuspended with Tris-HCl 20mM pH 8, TCEP 1mM. S2: Soluble fraction resuspended with CAPS 200 mM pH 10.5, TCEP 1 mM, NaCl 1 M. E1: elution at 300 mM imidazole. The protein is not soluble in the first buffer but it is in CAPS buffer. B | Purification of BRM protein ETH429 by precipitation and next by a Nickel-sepharose column. L: Protein ladder. S1: Soluble fraction resuspended in Tris-HCl 20 mM, TCEP 1mM. S2: Soluble fraction resuspended in Urea 2 M, NaCl 0.5 M, TCEP 1mM. E1: Elution at 300 mM imidazole. The protein ETH429 is not soluble in Tris-HCl 20 mM, TCEP 1mM but it is in Urea 2 M, NaCl 0.5 M, TCEP 1 mM. C | Solubilisation of SYD proteins ETH423, ETH424 and ETH425 containing a histidine tag. L: Protein ladder. T: total fraction. S1: Soluble fraction resuspended in Tris-HCl 20 mM, TCEP 1 mM. S2: Soluble fraction resuspended in CAPS 100 mM, TCEP 1 mM. S3: Soluble fraction resuspended in CAPS 200 mM, TCEP 1 mM. The three proteins are soluble in CAPS 200 mM, TCEP 1 mM. In order to be able to have a purified, soluble and not aggregated protein for interaction assays, several solubilisation tests were performed.

In the previous sections I presented the work performed to design, produce and purify soluble constructs of LFY, BRM and SYD. With the presented protein material, were we able to map LFY's region interacting with the HSA domain? To answer this question, we next performed biophysical and *in vivo* (in a heterologous system) interaction assays.

2. INTERACTION ASSAYS

A. Pull down assay

Pull-down assays were carried out with all the LFY constructs and BRM construct ETH428 shown in Figure 63. In this assay, GST-tagged LFY proteins are separately bound to a Glutathione-Sepharose resin. The potentially interacting protein BRM is then added to each of the columns. After incubation, the column is washed in order to remove unbound BRM proteins while LFY-BRM complexes should remain bound to the columns. Complexes are finally eluted with reduced glutathione (Figure 64).

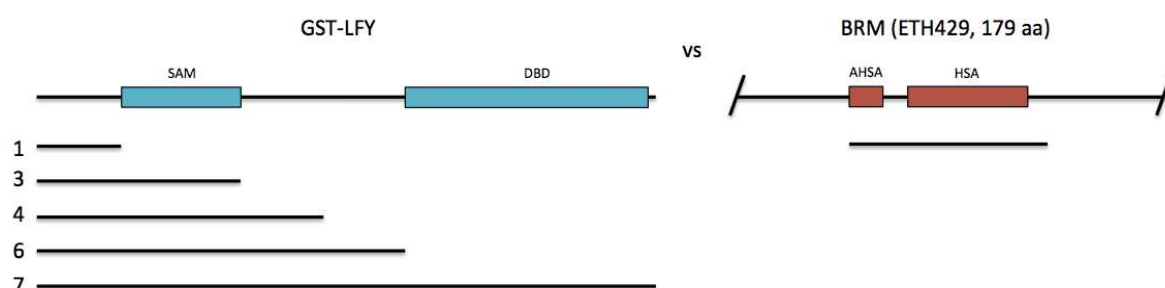


Figure 63. Representation of GST-LFY and BRM constructs used for pull down assays.

LFY construct 1, from 1-36 aa. 3, 1-137 aa. 4, 1-230 aa. 7, 1-420, complete protein. BRM ETH429 is a soluble construct that comprises the two essential domains (A-HSA and HSA) for interaction with LFY as shown by (Wu et al., 2012).

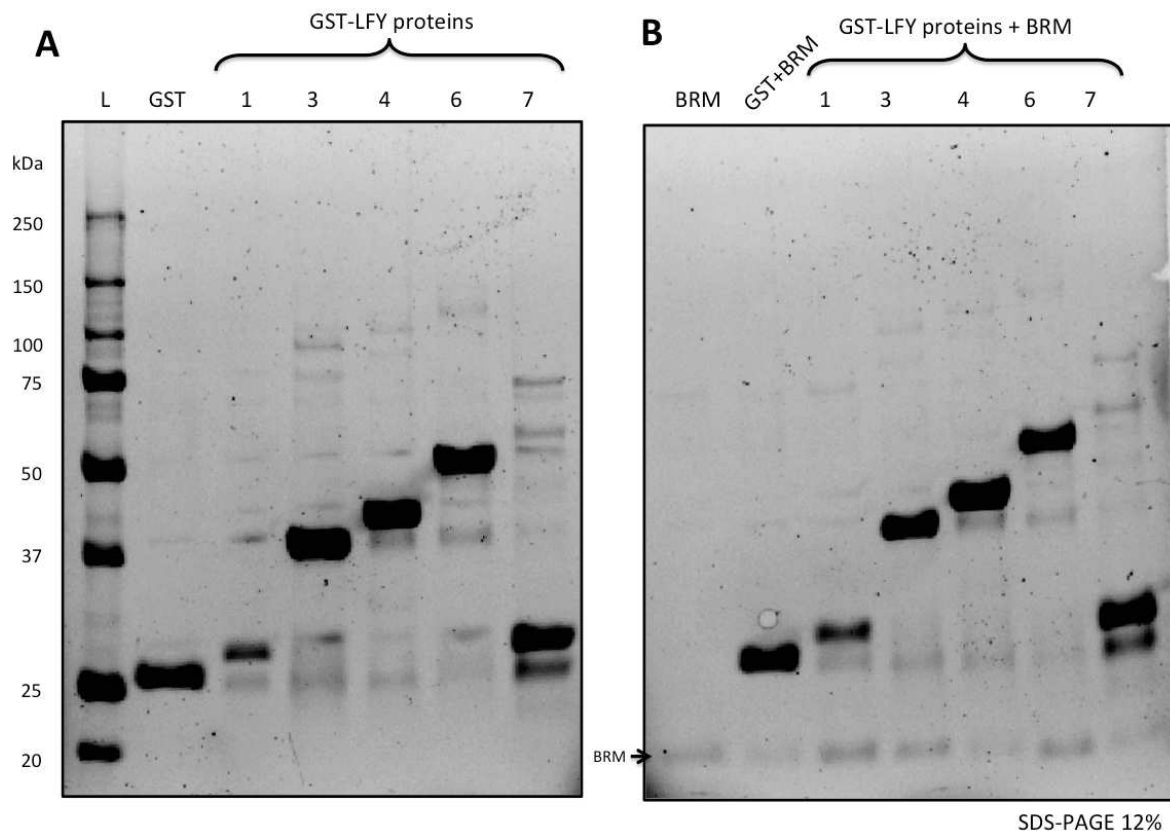


Figure 64. Stain-free PAGE gel of the pull down assay between GST-LFY and BRM.

A | GST-tagged proteins alone. B | Firstly, BRM protein alone (expected size, 21 kDa) after interaction test between the tagged proteins and BRM. GST-LFY proteins were bound to the column through their GST-tag. BRM protein without any tag was passed through the column. We can observe a faint band corresponding to BRM, however, it was expected to be removed during the column washing, which is not the case. Therefore, BRM protein presents a non-specific binding to the Glutathione-Sepharose column as well as to the GST protein alone.

In Figure 64B, interactions between all GST-LFY proteins and BRM can be observed. However, BRM was, to a certain extent, binding to the column non-specifically, being found in the elution fraction (First lane of Figure 64B). This was most probably due to issues with buffer compatibility between the different proteins and the assay conditions, which did not allow us to perform an experiment with all the required controls. In order to eliminate this non-specific binding, we tried adding to the purification buffer 5% Bovine Serum Albumin (BSA) and 0.5% triton, in order to eliminate the non-specific binding. Indeed, the interaction of BRM to the column was eliminated but non-specific binding of BRM to the GST protein alone was still observed. Unfortunately, we could not draw any conclusion from these tests (Figure 64).

B. HTRF assay

We next used a different method to study these interactions: Homogeneous time resolved fluorescence (HTRF). This is an assay based on the transfer of energy between two fluorophores, a donor and an acceptor, when in close proximity. These fluorophores are coupled to antibodies that recognize specific tags fused to the proteins of interest. HTRF combines standard Fluorescence Resonance Energy Transfer (FRET) technology with time-resolved measurement of fluorescence, allowing elimination of short-lived background fluorescence (Figure 65).

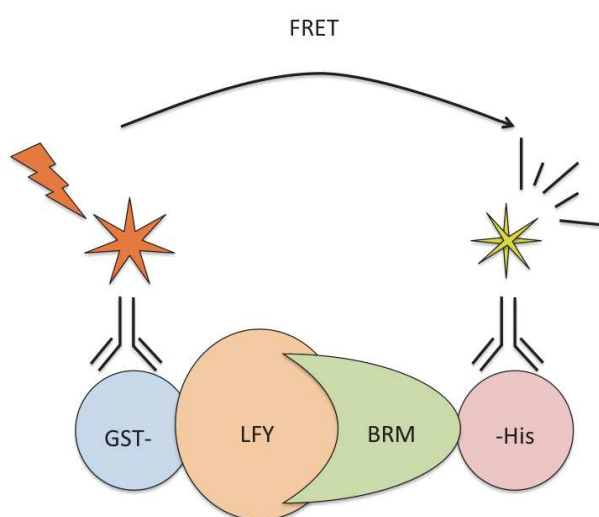


Figure 65. Schematic representation of FRET.

The donor and acceptor are covalently attached onto two different antibodies recognizing the GST and His tag fused to the protein of interest. The donor is excited by an energy source and when the two partners come close enough to each other, the excitation of the donor triggers an energy transfer to the acceptor, which in turn emits specific fluorescence at a given wavelength.

We made additional complementary constructs for LFY with deletions starting at the N-terminal end and tested a total of nine different constructs showed in Figure 66. For the HTRF assay, we worked with purified BRM ETH429 construct and with crude bacterial extracts expressing LFY proteins to increase LFY stability and functionality. All LFY constructs containing the DBD showed binding to BRM in HTRF test as shown in the Figure 67, leading to the conclusion that it is most probably the DBD of LFY that is responsible for the interaction with BRM ETH429. With the aim of confirming this result, several other biophysical techniques were explored.

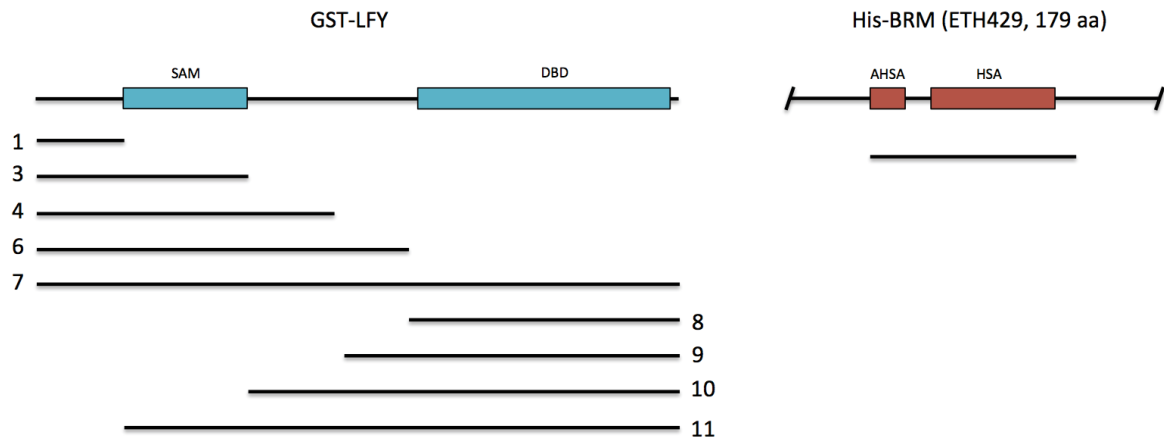


Figure 66. Representation of GST-LFY constructs starting from the C-terminal and His-BRM used for HTRF assays. Both GST and 6-His tags are fused to the N-terminal side of LFY and BRM, respectively.

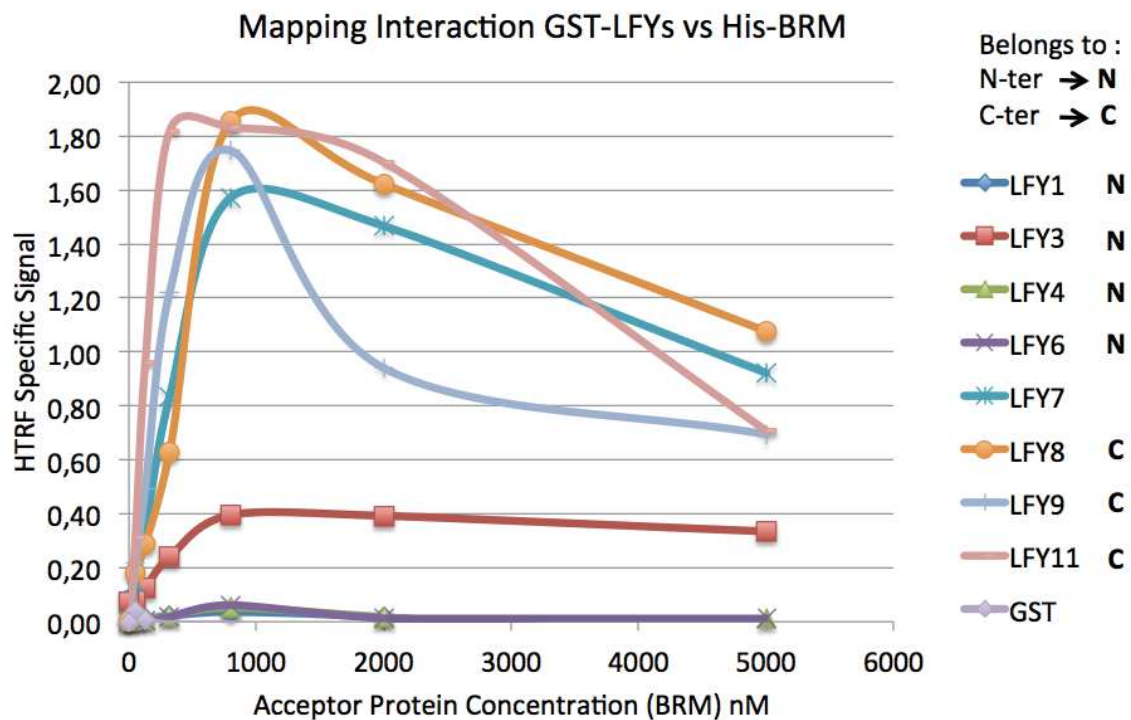


Figure 67. HTRF at six different concentrations of GST-LFY constructs against a fixed concentration of His-BRM ETH428.

Two different fluorophores are coupled with specific antibodies against the desired tagged proteins: LFY proteins 1 to 11 with a GST-tag and BRM ETH428 with a His-tag. The FRET signals were generated once these proteins interact with each other to bring the labeled antibodies into proximity. Measurements were done after one day. GST-LFY construct 8, which corresponds to the DBD, is able to interact with BRM ETH429, as well as all the other GT-LFY constructs containing this domain. Contrary to this, GST-LFY constructs lacking the DBD are not able to interact with BRM ETH429.

One of the used techniques to further test the interaction between LFY- C terminal and BRM (ETH429) was gel filtration, simply by observing first, the individual purification curve and specifically the elution volume of the each of the proteins of interest's peak. Then, by performing a co-purification, in which the purification curve of the supposed LFY-BRM complex is analyzed. The complex peak should have a different elution volume because of its larger size.

Another of the biophysical techniques is the BLITZ system, based on bio-layer interferometry technology. One of the proteins of interest is attached to a biosensor tip surface equipped with an optical layer. The BLITZ system (ForteBio) emits white light down the biosensor, and then collects any light reflected back. Reflected wavelengths are affected by the thickness of the coating on the optical layer. Changes in wavelengths are captured by a spectrometer. Any change in the number of molecules bound to the biosensor causes a shift in the interference pattern that is measured in real time. Finally, a surface plasmon resonance (SPR)-based analytical instrument called Biacore was also tested. Here one binding partner is attached to a sensor chip and the other is injected in a continuous flow of solution. Biacore utilizes the phenomenon of surface plasmon resonance (SPR) to detect interactions as they happen. SPR causes a reduction in the intensity of light reflected at a specific angle. This angle is proportional to the mass of material bound. The tests performed with the mentioned techniques were unsuccessful. Proteins might have been aggregated, and protein aggregation is a main barrier hindering structural and functional studies.

To summarize, it was observed through HTRF assays that GST-LFY construct 8 which comprises the DBD, is able to interact with BRM ETH429, as well as all the other LFY constructs containing the DBD. This suggests that this domain would mediate interaction to BRM A-HSA and HSA domains. In order to confirm this result, the *in vivo* (in a heterologous system) assay Y2H, was performed.

C. Y2H assay

To further verify LFY interacting domains with BRM, we performed yeast-two-hybrid (Y2H) assays. Constructions shown in Figure 68 were tested against the BRM ETH429

construct containing the A-HSA and HSA domains. The assay was carried out on selective media lacking tryptophan, leucine and histidine (–WLH) and in the presence of 0.5 mM of 3-Amino-1,2,4-triazole (3-AT), a competitive inhibitor of the *HIS3* gene product used as a reporter gene. No interaction was observed with this assay, probably because of their reduced size compared to Wu et al., (2012), where the BRM construct used was a longer N-terminal domain of the protein up to the amino acid 976.

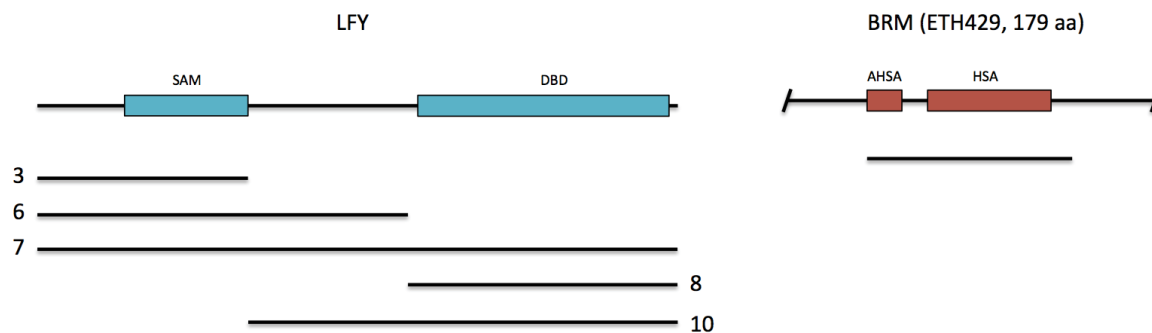


Figure 68. Representation of the constructs of LFY and BRM proteins used for Y2H assays.

Since the constructs used by Wu et al., (2012), were much larger for BRM, in collaboration with Dr. Catarina Da Silva (LPCV Grenoble), we repeated the assay including the complete BRM protein as well as larger domains such as the N-terminal, a middle region and the C-terminal (Figure 69). Unluckily, the assays again did not show any interaction. The experimental conditions of this assay were more stringent than the Y2H assay performed with BRM ETH429 protein as we used a yeast strain with two growth-reporter genes (*HIS3* and *ADE2*) instead of one (*HIS3*) because we tried to avoid the background activity shown in Wu et al., (2012) (Figure 57), when using only *HIS3* as a reporter.

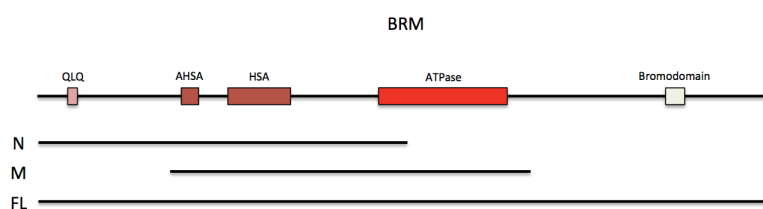


Figure 69. Representation of the new, larger constructs of BRM protein used in Y2H assays.

BRM N (1-967 aa), BRM M (670-1219 aa) and BRM FL (1-2193 aa).

To summarize these experiments, data from Wu et al., (2012) showed an interaction between LFY and the N-terminal domain of BRM (aa 1 - 976) and SYD (aa 1 - 657) and LFY in pull down assays, in BiFC assays as well in Y2H. Further mapping of the regions of BRM and

SYD interacting with LFY suggested that the most important domains for this interaction are the A-HSA and HSA domains. We tried to continue this mapping and we designed several BRM constructs near the two previously mentioned domains. We observed by HTRF that the C-terminal domain of LFY is able to interact with the A-HSA and HSA domains of BRM. We tried to confirm this interaction without success by other biophysical assays (Pull-down, gel filtration, Blitz, Biacore) most probably because protein aggregation prevents potential interactions with another protein of interest. Previous work from Wu et al., (2012) showed a weak interaction between LFY and BRM in Y2H assays that we were not able to reproduce. However, work from Catarina Da Silva showed that BRM is indeed able to interact with the SEP3 transcription factor, which has been suggested to be able to recruit BRM and SYD to AG and AP3 regulatory region Wu et al., (2012). This result indicates that in Y2H assays, the BRM construct ETH429 is correctly produced and functional for interaction.

To conclude, the work of LFY and BRM or SYD interaction, distinct constructs were designed with the help of informatic tools to calculate a D-score for a given protein sequence and ProtParm, in order to predict protein solubility. These constructs were designed with the aim of mapping their interacting domains. For BRM and SYD, constructs were specifically designed around the A-HSA and HSA domains in relation to precedent experiments that indicated these domains as essential for interaction with LFY. We produced and purified these constructs using several buffers. BRM and SYD are proteins that easily aggregate and are not always suitable for biophysical interaction assays. Nevertheless, an interaction was observed in HTRF between LFY DBD and BRM A-HSA and HSA domains. The weak interaction observed by Y2H between LFY and BRM by Wu et al., (2012) was not confirmed by our assays.

II. CHROMATIN REMODELING COMPLEXES AND LFY FUNCTION *IN PLANTA*

1. PHENOTYPIC ANALYSIS OF LFY OVER-EXPRESSION IN SYD AND BRM MUTANT BACKGROUNDS

The main question that we asked *in planta* was: is LFY still able to carry out its function when ectopically expressed in absence of SYD and BRM functions? For this, we worked with three different T-DNA insertion mutants. The *syd-5* null mutant that presents homeotic defects, such as reduced petal and stamen identity as well as carpel defects (Wu et al., 2012). The null mutant *brm-1* (Hurtado et al., 2006; Kwon, 2006) and the intermediate mutant *brm-3* (Tang et al., 2008). These mutations are represented in Figure 11, in section II of the introduction. The three mutants were crossed to the 35S:LFY line.

A major phenotype of 35S:LFY lines is the ectopic formation of flowers at the axil of rosette leaves, which is due to the LFY-dependent conversion of axillary inflorescence meristems to flower meristems. In F1 plants resulting from these three crosses, this phenotype was lost. Subsequent generations also showed a lack of ectopic flowers in homozygous *syd-5*, *brm-1* and *brm-3* backgrounds. This striking observation suggests that there is an important genetic interaction between LFY and SYD and BRM remodelers. Figure 70, shows the phenotypes of these plants.

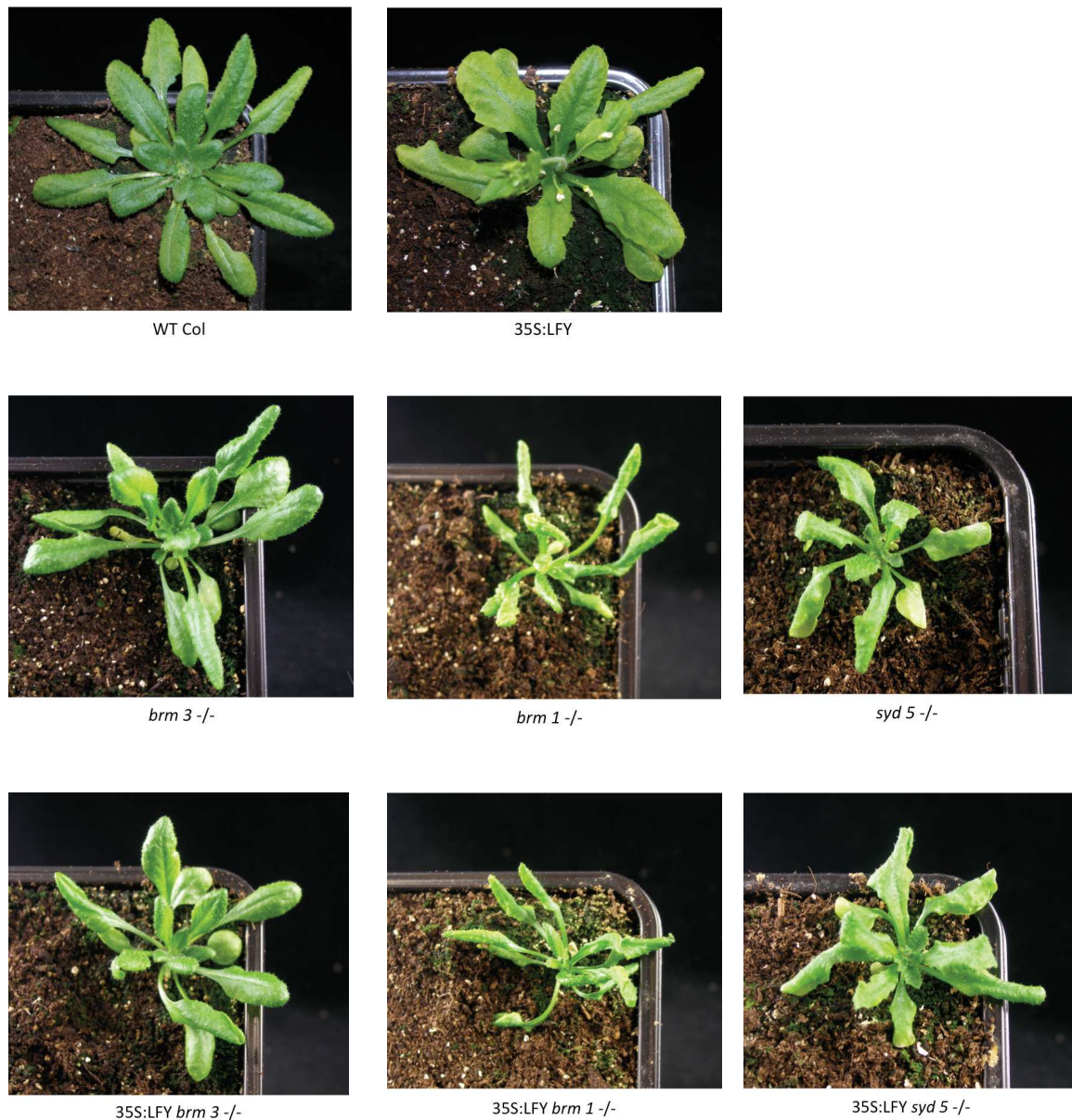


Figure 70. Phenotypes of 5-week old plants of the Columbia ecotype.

The genotype is indicated below the pictures. A loss of the 35S:LFY phenotype can be observed in each of the three heterozygous mutant backgrounds, 35S:LFY *brm*-3, 35S:LFY *brm*-1 and 35S:LFY *syd*-5, since the phenotype displayed by the F1 of the crosses is the same as the individual *syd* and *brm* mutants.

2. T-DNA INSERTIONS

The results showed above are surprising, especially because the *syd* and *brm* mutants are acting as dominant mutations to suppress 35S:LFY phenotype. T-DNA insertions in SALK lines are prone to inducing transcriptional silencing of transgenes expressed from the 35S promoter (Mlotshwa et al., 2010). Potential triggers of this transcriptional silencing are promoter sequences that are common to both the T-DNA insertions and the unlinked

transgenes that are silenced by the T-DNA. The T-DNA insertions in SALK, mutant lines all carry a copy of the cauliflower mosaic virus (CaMV) 35S promoter, and a study of SALK lines found that a high proportion of those lines induced transcriptional silencing of transgenes expressed from the 35S promoter (Daxinger et al., 2008).

In consequence, we carried out two different experiments with the aim of verifying if the suppression of the 35S:LFY phenotype could be due to transcriptional silencing. The first experiment consisted in looking at LFY protein level in 15 day-old seedlings grown under long days for the following genotypes: wild type, 35S:LFY and 35S:LFY *brm-3* $-/-$ and 35S:LFY *syd-5* $-/-$. LFY was detected at similar levels in 35S:LFY control line and in mutant backgrounds (Figure 71).

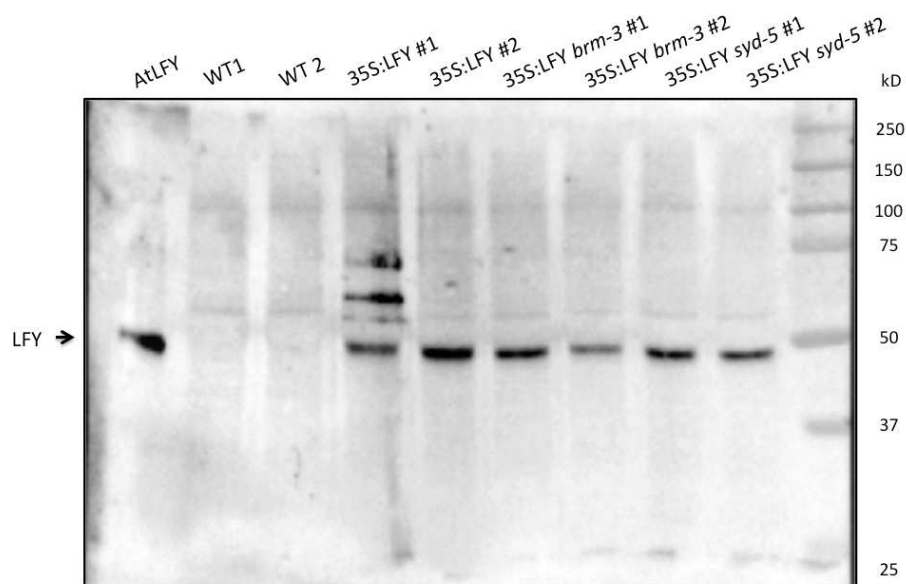


Figure 71. Western blot comparing LFY expression in WT, 35S:LFY, 35S:LFY *brm-3* and 35S:LFY *syd-5* plants.

To verify whether LFY expression in *syd* and *brm* mutants is altered compared to the 35S:LFY line, protein extraction of 15 day-old seedlings grown under long day conditions was performed. Three leaves were pooled from different individual plants for each genotype. Purified AtLFY protein was loaded as a reference in the first lane. Protein expression does not seem to be altered in the mutant backgrounds.

The second experiment consisted in making different crosses of the 35S:LFY line with four SALK lines randomly selected. The SALK mutant lines used were *ult1-3*, *rax-1*, *ala-10* and *alis-1*. Mutants *ult1-3* and *rax-1*, are both related to meristem development but less directly

than *BRM* and *SYD*, whereas *ala-10* and *alis-1* are not and therefore should not affect the 35S:LFY phenotype. The first mutation is in the *ULTRAPETALA1 (ULT1)* gene. This protein acts as a negative regulator of meristem size in inflorescence and floral meristems. *ult1* mutations cause inflorescence meristem enlargement, as well as a decrease in floral meristem determinacy. *ULT1* and *LFY* pathways function independently in specifying identity at the floral meristem (Engelhorn et al., 2014). The second mutant affects the *RAX1* gene, already mentioned as important for meristem formation (Chahtane et al., 2013; Denay et al., 2017). The third SALK line, *ala-10*, is a mutant of the phospholipid-transporting ATPase 10 (*ALA10*), an ATPase involved in the transport of phospholipids (Botella et al., 2016). Finally, the *alis-1* mutant has a T-DNA insertion in the *ALA-interacting sub unit 1 (ALIS1)* gene, which encodes a protein potentially necessary for the function and localization of *ALA10* (Botella et al., 2016).

The F1 plants from the crosses between these lines and the overexpressing *LFY* line showed the expected 35S:LFY phenotype. A cross was also made between the 35S:LFY line and the weak allele *brm-106* (*Ler* background) which has a substitution in the ATPase domain (P1140=>S), and is not a SALK but an Ethyl methanesulfonate (EMS) mutant. F1 plants from this cross, showed the 35S:LFY phenotype. This is possibly because *brm-106* is a weak allele that allows *BRM* and *LFY* to genetically interact (Figure 72).

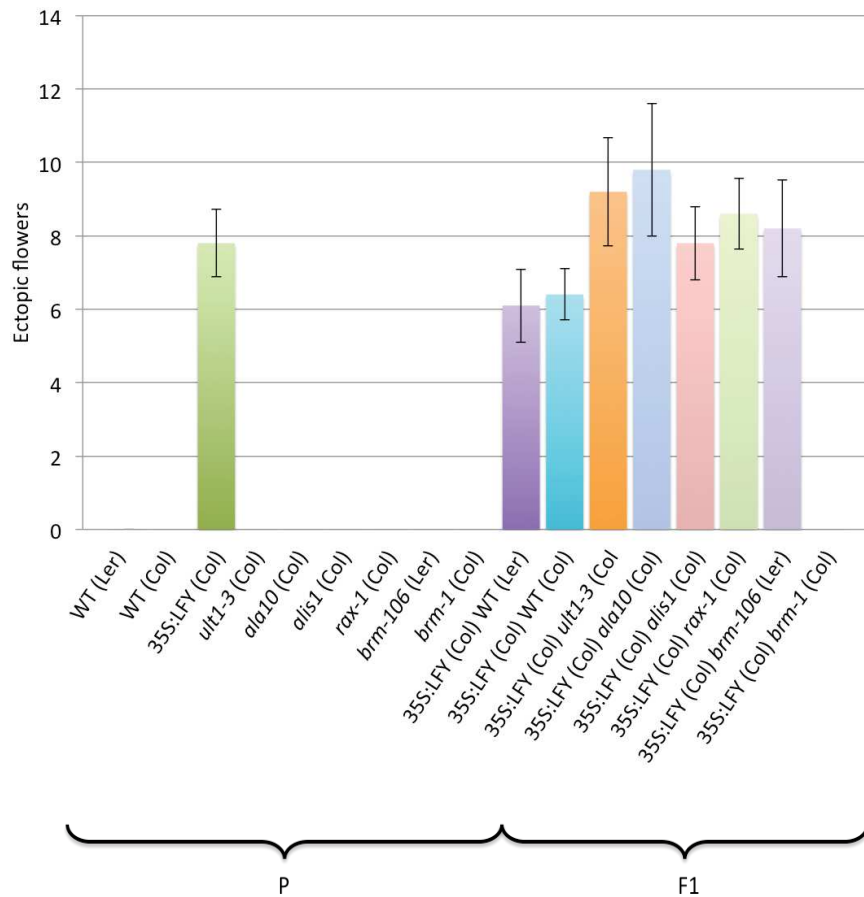


Figure 72. Number of ectopic flowers in the parents and the plants from crosses between SALK lines.

10 plants for each genotype were taken into account for ectopic flower counting. The 35S:LFY phenotype is not suppressed in « random » SALK background: *ult1-3*, *rax-1*, *ala-10* and *alis-1*, as expected, since the formation of ectopic flowers is present. The 35S:LFY phenotype is not suppressed in the weak *brm-106* background.

Overall, the F1 plants resulting from the crosses between the LFY overexpressing line and the three mutants *syd-5*, *brm-1* and *brm-3*, did not show the 35S:LFY phenotype. This suggests an important genetic interaction between LFY and SYD and BRM remodelers.

The two different experiments performed with the aim of verifying if the suppression of the 35S:LFY phenotype could be due to transcriptional silencing indicated that there is a real genetic interaction. Through western blots, the measured amount of LFY in the three crosses was similar to that of the 35S:LFY line alone, suggesting that the protein is also well produced in the mutants. Moreover, crosses of the 35S:LFY line to non-related T-DNA

mutants showed the expected 35S:LFY phenotype. These results confirm our main observation of a genetic interaction between LFY and both ATPase subunits, BRM and SYD.

III. DISCUSSION

The chromatin-remodeling ATPases BRM and SYD are fundamental catalytic subunits of large protein complexes and are among the best characterized SWI2/SNF2 ATPases in Arabidopsis. They act as TrxG proteins countering the effect of polycomb repression at the *AP3* and *AG* target regulatory regions during flower development. In previous work, SYD and BRM were shown to interact physically with LFY and a model of recruitment of SYD to *AP3* and *AG* dependent on LFY and SEP3 was proposed (Wu et al., 2012). The objective of this part of my project was to finely map the interacting domains of, on one hand, LFY and on the other hand, SYD and BRM and also, to better understand the interactions of these partners at the genetic level.

Wu et al., (2012) proposed direct binding between LFY and the two ATPases occurs via the helicase/SANT-associated (HSA) and the HSA-adjacent (A-HSA) domain. We designed, produced and purified construct of these proteins for interaction tests. A major difficulty encountered in the biochemical studies of this project was certainly the fact that we worked with proteins or domains that present low solubility in certain buffers and aggregation in several biophysical assays. We finally found that LFY DBD interacts with BRM A-HSA and HSA domains. However, using the biophysical techniques of pull-down, gel filtration, Blitz and Biacore, we were not able to observe an interaction.

An interesting alternative to protein production and purification from bacteria is the use of insect cells instead. They offer the possibility, through baculovirus infection, to co-produce and co-purify proteins with posttranslational modification since most of the posttranslational modification pathways present in mammalian cells also occur in insect cells (Kost, Condreay, & Jarvis, 2005). This is a technique worth to explore that could lead to the identification of the minimal interacting regions of SYD and BRM with LFY. Identifying a SYD-LFY or BRM-LFY complex with interacting domains, open the way to do crystallization experiments and would give valuable information, and clues to find a common mechanism

that could extend to other transcription factors and other ATPases as well. We however find that LFY DBD interacts BRM A-HSA and HSA domains.

Concerning the results obtained from Y2H assays, we could notice that selective medium –WLAH is likely too stringent for the interaction between LFY and BRM. Probably for this reason, Wu et al., (2012) used in their Y2H assays used –WLH as a selective medium and further added a precise concentration of the 3AT inhibitor. Our Y2H assay in the latter conditions did not showed any interaction. To repeat our assay in the conditions of Wu et al., (2012) using not only BRM ETH429 but also all the other larger constructs designed (M, N, and the full-length protein), would be very interesting.

Furthermore, Y2H results from Catarina Da Silva confirmed that BRM interacts with the SEP3 transcription factor, which has been suggested to be able to recruit BRM and SYD to AG and AP3 regulatory region (Wu et al., 2012). This is a promising result that could be further studied to map minimal interacting domains between a TF and an ATPases such as BRM and SYD.

To conclude, our results show that BRM is able to interact with SEP3, but not with LFY in Y2H assays. Concerning LFY and BRM, we observed an interaction in HTRF but no interaction with the other previously mentioned biophysical techniques. Altogether, these results lead us to question the interaction results between LFY and BRM observed by Wu et al., (2012), which seem, at least in Y2H assays, particularly weak.

Regarding the *in vivo* experiments, we obtained an exiting result of the genetic interaction between LFY and SYD and BRM. This result was confirmed by verifying the LFY level in all plants by western blot as well as additional crosses with non-related SALK mutant lines. An additional control could be done by crossing the 35S:LFY line with the mutants *brm-101*, and *syd-2* (Figure 11 in Introduction section II), which are both strong alleles with no SALK T-DNA insertion. In conclusion, these results strongly suggest that LFY function involves BRM and SYD functions in order to activate its target genes. Whether a physical interaction between LFY and SYD or BRM takes place in plant cells remains to be determined.

One perspective in this research line is to assess to what extent LFY's binding on chromatin is affected by *brm* or *syd* mutations. This could be done by performing ChIP-seq assays to compare the binding of LFY in the 35S:LFY line to 35S:LFY *syd-5*, for example. Performing MNase-seq as well as DNaseI assays with this biological material would also allow comparing the chromatin landscape and having a larger scope to assess the genetic interaction between LFY, and BRM or SYD.

A complementary experimental set up that will allow the study of LFY's ability to open closed regions of chromatin is to use the Glucocorticoid (GR) inducible system. The mammalian nuclear steroidal receptor Glucocorticoid (GR) is activated by the presence of the steroid, dexamethasone. The GR protein is not only a receptor but also a transcription factor whose nuclear localization and transcriptional regulatory function is also regulated by steroids, therefore it can be fused to a protein of interest in order to detect the immediate effect of this protein activation (Corrado & Karali, 2009). Therefore, transformed plants with a constitutively expressed translational fusion of LFY to the glucocorticoid receptor (35S:LFY-GR line) could allow to induce expression with a dexamethasone treatment, comparing it to a treatment made with a solution without hormone (mock-treatment). Carrying out genome-wide assays from these two samples would allow to see if LFY induction results in a change of the chromatin landscape from closed to open, especially in LFY target regulatory regions.

For instance, in our team, a new method to map chromatin accessibility at the genomic scale is being explored. The Assay for Transposase Accessible Chromatin with high-throughput sequencing (ATAC-seq), probes DNA accessibility with the highly active transposase Tn5, which inserts sequencing adapters into accessible regions of chromatin. Sequencing reads can then be used to map accessible regions of chromatin. This method is a fast and sensitive alternative to DNase-Seq for assaying chromatin accessibility genome-wide, or to MNase-seq for assaying nucleosome positions in accessible regions of the genome (Buenrostro et al., 2015) .

Recently published results from Archacki et al., (2016) on BRM genome-wide localization in Arabidopsis, show that BRM occupies thousands of sites in the genome, most of which located within or close to genes. Furthermore, Torres & Deal, (2018),

comprehensively characterized the antagonistic and cooperative contributions of H2A.Z and BRM to transcriptional regulation, and illuminated their interrelated roles in chromatin organization. Similarly, the analysis of genome-wide datasets, such as ChIP-seq showing co-localization between LFY and BRM, is a powerful resource that will allow us to identify genes that show relationships between these partners and how they act in concert for transcription.

MATERIALS AND METHODS

I. MATERIALS

1. PLANT MATERIAL

A. thaliana plants used in this work belong to the Columbia-0 ecotype unless otherwise specified. The *syd-5* (SALK_023209), *brm-1* (SALK_0300469) and *brm-3* (SALK_088462) mutants were ordered from The Nottingham Arabidopsis Stock Centre (NASC). The SALK mutants lines *ala-10* and *alis-1* were kindly provided by Cesar Botella (Botella et al., 2016), *ult1-3* by Cristel Carles (Carles & Fletcher, 2009) and *rax-1*, (Keller, 2006; Muller, 2006) by Grégoire Denay. Also, the mutant *brm-106* of the *Landsberg erecta* (*Ler*) ecotype, (P1140=>S) described in (Efroni et al., 2013) was provided by Doris Wagner.

2. BACTERIAL MATERIAL

Cloning was done in *E. coli* DH5 α strain (Invitrogen). Recombinant protein expression was done in *E. coli* Rosetta2 strain (Novagen).

3. YEAST MATERIAL

Saccharomyces cerevisiae yeast strains Y187 and AH10, were used for Y2H assays and transformed with constructs in pGBKT7 and pGADT7 vectors (Clontech), designed to express proteins fused to the GAL4 DNA binding domain and the GAL4 activation domain, respectively. For cloning PCR fragments, the zero blunt PCR Cloning kit was used (Invitrogen).

II. METHODS

1. PLANT BIOLOGY

- *A. thaliana* growth conditions

Plants were grown in long-day conditions (at 22°C, with 16 h of light exposure and 8 h of darkness). Seeds were first treated by exposure at -80°C for 2 h to kill potential insect eggs. Then, they were sterilized by a 10-min treatment in a solution of 0.4% chlorine, 0.02% SDS and 86% ethanol. Seeds were then rinsed with a solution of 95% ethanol and dried under a sterile hood. Seeds were then sown on a sterile media 0.5x Murashige and Skoog (MS) and stratified for 2 days at 4°C in darkness, before being transferred to long-day conditions. After 15 days, seedlings were transferred to soil (2/3 potting soil, 1/3 vermiculite) treated with a biological insecticide containing *Bacillus thuringiensis* (Vectobac 12-AS, Biobest).

- Selection of seeds having received the transgene

The selection of seeds carrying the 35S:LFY promoter was carried out using a cassette allowing the production of the Green Fluorescent Protein (GFP) under control of the At2S3 promoter as selection method (Bensmihen et al., 2004). Thereby, after ultra-violet (UV) excitation (between 300 and 400 nm), transgenic seeds emit a green light, around 530 nm that allows their selection (Figure 73).

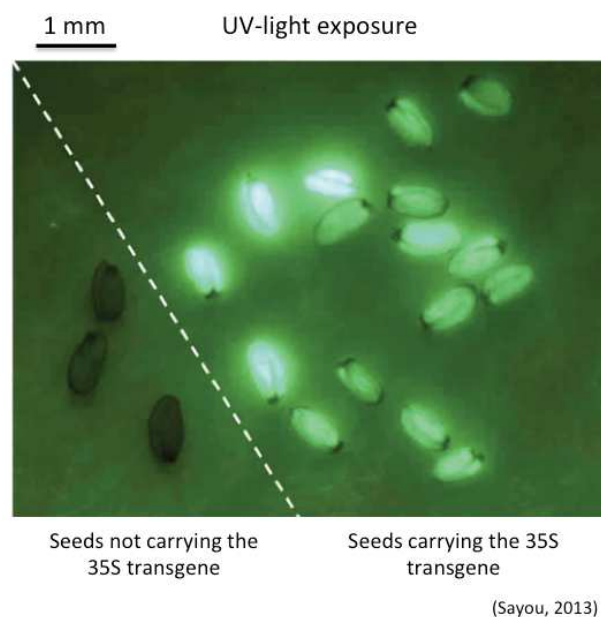


Figure 73. UV-light exposure to seeds carrying the transgene 35S:LFY, producing GFP.

- DNA extraction

Plant genomic DNA was extracted from a fragment of a young rosette leaf (the equivalent of 25 mm²) flash frozen in liquid nitrogen and mechanically grinded (Grinder MM30, Retsch/Qiagen). The rest of the extraction took place at room temperature. The grinded leaf was resuspended in 500 µL of Edwards buffer (200 mM Tris-HCl pH 7.5, 250 mM NaCl, 25 mM EDTA, 0.5% SDS) and then centrifuged 1 min at 13000 rpm. 400 µL of supernatant were then mixed with 400 µL of isopropanol and incubated 2 min at room temperature. Afterwards, they were centrifuged 5 min at 13000 rpm. The pellet was washed with 500 µL a solution of 70% ethanol and centrifuged 5 min at 13000 rpm. The supernatant was eliminated and after complete ethanol evaporation, the pellet was resuspended in 30 µL of water. DNA can then be conserved at -20°C during several weeks.

- Genotyping by PCR

Genotyping of the *syd-5* allele was performed with primers GID1C-TDNA and oGV149. For the wild type allele, oGV148 and oGV149 were used. Genotyping of the *brm-1* allele was performed with primers GID1C-TDNA and oGV143. For the wild type allele, oGV142 and oGV143 were used. Finally, genotyping of the *brm-3* allele was performed with primers GID1C-TDNA and oGV147. For the wild type allele, oGV146 and oGV147. The list and sequences of primers used for genotyping is indicated in Table 4.

Table 4. Primers used for genotyping *syd-5*, *brm-1* and *brm-3*.

Primers allowed the identification of the presence of T-DNA insertion.

Primer name	Mutation	Sequence 5' > 3'	Allele
GID1C-TDNA	<i>syd-5</i>	TGGTTCACGTAGTGGGCCATCG	mutant
oGV149	<i>syd-5</i>	CTTCTCACGGTGAAGTCGTTTC	mutant
oGV148	SYD	CTGCGAAGTTAGCTGTTTTGG	wild type
oGV149	SYD	CTTCTCACGGTGAAGTCGTTTC	wild type
GID1C-TDNA	<i>brm-1</i>	TGGTTCACGTAGTGGGCCATCG	mutant

oGV143	<i>brm-1</i>	AGTTTATACCGTTGCATCCCC	mutant
oGV142	BRM	AATATACGCTTGCTGCATTGG	wild type
oGV143	BRM	AGTTTATACCGTTGCATCCCC	wild type
GID1C-TDNA	<i>brm-3</i>	TGGTTCACGTAGTGGGCCATCG	mutant
oGV147	<i>brm-3</i>	GACCTTCCTTGTCGATTCTCC	mutant
oGV146	BRM	CTGGGAATGAAGAAGAGGGAG	wild type
oGV147	BRM	GACCTTCCTTGTCGATTCTCC	wild type

- Protein samples from plant tissues

2 mL-eppendorf tubes were individually weighed with two grinding beads. Three rosette leaves were collected for each genotype, put into the eppendorf tubes and weighed in order to calculate the amount of biologic material. Tubes were then frozen with liquid nitrogen and grinded with the grinder TissueLyser during 2 min at 30 beats/s. For 10 mg of biologic material, 80 µL of bromophenol blue 1x. Samples were then heated at 95°C for 5 min and finally stored at -20°C.

- Western blot

Analysis of protein levels from each of the mutants described above, was performed with the iBlot 2 Dry blotting system (Biorad), using as transfer program: 20 V during 1 min, 23 V during 4 min and 25 V during 2 min. The purified anti-LFY antibody (4029b) was used at a dilution 1/5000. The anti-KARI antibody was used at a dilution 1/3000 to detect KARI as a loading control. The secondary antibody anti-rabbit was used at a dilution 1/3000.

2. MOLECULAR BIOLOGY AND PROTEIN BIOCHEMISTRY

- Histone octamer co-production and co-purification

The co-production and co-purification of the four core histones was done following the protocol of (Shim et al., 2012). Briefly, *Xenopus laevis* core histones were expressed from a pET29a polycistronic plasmid containing the four core histones. Histone H2A, was tagged with a N-terminal hexahistidine (His6)-tag followed a thrombin cleavage site. Histone H4 was tagged with a C-terminal His6-tag preceded by a thrombin cleavage site. The plasmid encoding all four core histones was transformed into BL21(DE3)pLysS cells and plated on to an LB agar plate containing kanamycin (50 mg/mL) and chloramphenicol (25 mg/mL). The plate was incubated overnight (~16 h) at 37 °C. For a 1 L scale preparation, one colony was inoculated into 10 mL 2xTY media containing kanamycin (50 mg/mL) and chloramphenicol (25 mg/mL). This starter culture was shaken at 150 rpm for ~4-5 h at 37 °C until slightly cloudy. Subsequently, the culture was transferred into 1 L of the same media and was grown for another 6-7 h at 37 °C. When the OD600 reaches ~0.4, histone co-expression was induced by adding 0.4 mM isopropyl-beta-D-1-thiogalactopyranoside (IPTG). The culture was further shaken at 150 rpm at 37 °C overnight.

Cells were harvested by centrifugation at 4,500 x g for 30 min at 4 °C. Cell pellets were processed immediately or stored at -80 °C for future purification. Afterwards, cell pellets were resuspended in 50 mL of lysis buffer (20 mM Tris-HCl pH 8.0, 2 M NaCl, 0.5 mM tris(2-carboxyethyl)phosphine (TCEP). Resuspended cells were lysed by sonication (Sonifier 250 ultrasonic cell disruptor) during 15 min and clarified by centrifugation at 18,000 x g at 4 °C for 30 min. The supernatant was collected and injected on a 3 mL-resin column Nickel-Sepharose pre-equilibrated with 20 volumes of lysis buffer. The column was washed with 30 mL of lysis buffer containing 30 mM imidazole. Then, bound proteins were eluted by increasing the imidazole concentration from 50 to 300 mM. Each fraction was analyzed by 18% SDS-PAGE. 10 µL of sample was mixed with 3.33 µL of SDS loading buffer 4x, incubated at 95 °C for 10 min, then loaded on the gel. Electrophoresis was run at 200 V for 1 h. The SDS-PAGE gel system and protein visualization used is indicated further in the text in section: *SDS-PAGE polyacrylamide gels*.

Thrombin digestion was carried out by adding purified thrombin (Sigma) in 25:1 mass ratio and incubating the samples at room temperature for ~ 4 h. The digestion was confirmed by SDS-PAGE.

Thrombin-digested histones were then concentrated up to 3.3 mg/mL by centrifugation using Amicon membrane (MWCO 50 kDa) at 4 °C (Millipore) at 4,000 x g during 10 min intervals, with gentle mixing between each centrifugation interval. The concentrated sample was then injected onto a Superdex 200 10/300GL column in lysis buffer. The histone octamer peak was eluted at an elution volume of 12.8 mL. The peak fractions were pooled and concentrated, aliquoted and flash-frozen in the presence of 50% glycerol for storage of maximum 7 months.

- Amplification of DNA fragments for nucleosome reconstitution

The 601 sequence and the 601_{LFY} sequence with a LFY-binding site were present in the vectors pET601 and pETH 500, respectively. The 182-bp « 2 » DNA fragment, and its mutant without LFY-binding sites, « 2me » correspond to the genomic location: chr1-30130416-30130778. The 157-bp « 5c » DNA fragment, and its mutant without LFY-binding sites, « 5cm3 » correspond to the genomic location: chr3-21507065-21507222. Mutated bases are underlined.

The corresponding sequences are:

Sequence 2

GACAAACAGAGCGGAAGTGCAAGAAAACAGTACATAGGAATGGTGACCATGCTCGACGTTGTGGCT
CACATCGCCGGCGATGACGGCGAGAGTGGGCTTGATAAGAAAATGGCTGCCCCTGTCTCTTCTATCAT
TGGTCATTGCCCTGAGGGTCTTAGCCTCTGGTCTCTCAATCCTAATAC

Sequence 2m3

GACAAACAGAGCGGAAGTGCAAGAAAACAGTACATAGGAATGGTGACCATGCTCGACGTTGTGGCT
CACATCGCCGGCGATGACGGCGAGAGTGGGCTTGATAAGAAAATGGCTGCCCCTGTCTCGCCCTGAG

TGTCTTAGCCTCTGAGGGTGCCTGAGAGTCTTAGCCTCTCCTAATAC

Sequence 5c

GACCTAGCCGCGGTTAACGGACCATTGGTCCACGACGGGTAATGACGGTAACTGTAAGGGTTATTGT
CAGGATGATAAAAGTAAGAGGAAACACCGCGTTTGGCAAGCTGACGTTTCGCGTTTGTGAGCGTTTTG
GTGTCCACCTAGGGCTTGTGAAG

Sequence 5cm3

GACCTAGCCGCGGTTGCCCTGATAGTCTTAGTCTCCTAGTAACGACGGTAACTGTAAGGGTGCCCTGA
GATTCTTAGCCTCTAAGAGGAAACACCGCGTTTGGCAAGCTGACGTTTCGCGTTTGTGAGCGTTTTGGT
GTCCACCTAGATCTTGTGAAG

The labeled DNA sequences were created by PCR with end-labeled primers (below in Table 5). The fluorescent-tagged DNA fragments were gel extracted and further purified using the DNA PCR clean-up kit (Macherey-Nagel). Sequences from 1 to 6 including 5c were inserted in the vector backbone pMA-T (invitrogen thermo Fisher scientific). Sequences 2m3 and 5cm3 were inserted in the vector backbone pMA-RQ (invitrogen thermo Fisher scientific).

Table 5. Primers used for amplification of DNA fragments for nucleosome reconstitution.

Primers were labeled with the Cy5 fluorescent dye to allow detection on EMSA.

Primer name	DNA fragment	Sequence 5' > 3'	Label
oJLM001	601	CAGGATGTATATATCTGACA	Cy5
oJLM002	601	TGGAGAATCCCGGTGCCGAG	-
oJLM001	601 _{LFY}	CAGGATGTATATATCTGACA	Cy5
oJLM002	601 _{LFY}	TGGAGAATCCCGGTGCCGAG	-
oEBH040	2	ATAGCGGCCGCGACAAACAGAGCGGAAGT	Cy5
oEBH023	2	GTATTAGGATTGAGAGACCAGAGGCTAAGACCCTCA	-
oEBH040	2m3	ATAGCGGCCGCGACAAACAGAGCGGAAGT	Cy5
oEBH037	2m3	GTATTAGGAGAGGCTAAGACTCTCA	-
oEBH038	5c	TATAGCGGCCGCGACCTAGCCGCGG	Cy5

oEBH042	5c	CTTCACAAGCCCTAGGTGGACACCAAAAC	-
oEBH038	5cm3	TATAGCGGCCGCGACCTAGCCGCGG	Cy5
oEBH039	5cm3	CTTCACAAGATCTAGGTGGACACCA	-

- Nucleosome reconstitution

Nucleosomes were reconstituted by one of the two following methods:

Salt dialysis method

Core histones and DNA were mixed in the molar ratio of histone octamer to DNA at a 1.1:1 ratio and dialyzed sequentially against TE buffers (10 mM Tris-HCl pH 8.0 and 1 mM TCEP) containing first 1 M, then 600 mM and finally 0 mM NaCl each, during 2h. The dialysis was carried out in eppendorf lids that were cutted off and covered with a piece of 3K MWCO membrane (Thomas Scientific), as shown in Figure 74. Results were verified on 5% native polyacrylamide gels run in 0.25x TBE at 130 V for 1h at 4 °C.

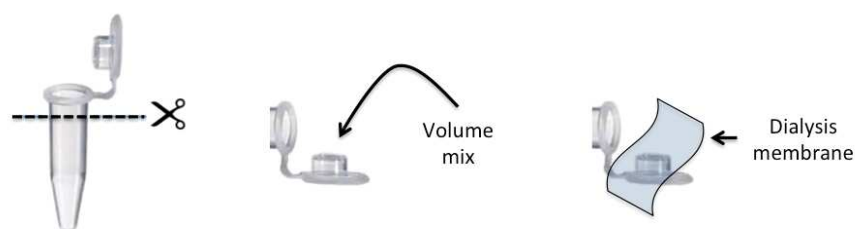


Figure 74. General steps for making dialysis buttons.

The lid (upper part) of a 200 µL-eppendorf tube was cut off. The sample was deposited in the lid and closed after inserting a piece of 3K MWCO dialysis membrane. The side of lid covered by the membrane was then putted in contact with the dialysis buffers.

Salt dilution method

NCPs were assembled with the salt dilution method as described in (Okuwaki et al., 2005). Briefly, an equal amount of recombinant histones and DNA fragment of interest were mixed in 10 µl of 10 mM Tris-HCl, pH 7.4, 1 mM ethylenediaminetetraacetic acid (EDTA), 0.1 mg/mL bovine serum albumin (BSA), and 1 mM DL-Dithiothreitol (DTT) in the presence of 2 M

NaCl and incubated at 37°C for 10 min. The reaction was serially diluted to 1.5, 1, 0.8, 0.7, 0.6, 0.5, 0.4, 0.25, and 0.2 M NaCl in 50 mM hydroxyethyl piperazineethane sulfonic acid (HEPES-NaOH) pH 7.5, 1 mM EDTA, and 1 mM DTT with 15-min incubations at 30°C for each dilution step. The salt concentration was brought to 0.1 M by adding 100 µl of 10 mM Tris-HCl (pH 7.5), 1 mM EDTA, and 1 mM DTT, 10% glycerol, and 0.1 mg/mL bovine serum albumin and incubated for 15 min at 30°C. The reconstitutions were confirmed by native gel analysis.

- Electrophoretic Mobility Shift Assay (EMSA)

Nucleosomes were incubated with recombinant proteins in DNA-binding buffer (10 mM Tris-HCl pH7.5, 1 mM MgCl₂, 1 mM DTT, 10 mM KCl, 0.5 mg/mL BSA, 5% Glycerol) at room temperature for 60 min. Free and bound DNA were separated on 5% non-denaturing polyacrylamide gels run in 0.2X Tris–borate–EDTA (TBE) buffer. DNA-protein complexes were visualized on the gels with Cy5-exposition filter (Biorad ChemiDoc MP Imaging System). As competitor DNA, fish DNA or competitive non-specific oligonucleotides shown in Table 6 were used. LFY proteins and their corresponding buffers used in EMSAs are shown in Table 7.

Table 6. Oligonucleotides for use as non-specific DNA competitor in EMSA.

Oligonucleotides were specifically designed with no LFY-binding site using the MORPHEUS web tool (<http://biodev.cea.fr/morpheus/>).

Oligonucleotide name	Sequence 5' > 3'
oEBH34	AGCCCGTACTATACAGCGTAGCCCGTA
oEBH35	TCGGGCATGATATGTCGCATCGGGCAT

Table 7. Purified LFY proteins used in EMSA with their respective buffer.

For proteins in buffers containing salt, a dialysis step was performed before the EMSA.

Protein	Expression vector name	Purification buffer
AtLFY	pETH94	20 mM Tris–HCl pH 8.5, 1 mM TCEP
AtLFY-DBD	pCA28	20 mM Tris–HCl pH 8.5, 1 mM TCEP
AtLFY-40	pETH79	20 mM Tris–HCl pH 8.5, 1 mM TCEP

GbLFY-40	pETH164	1 M NaCl, 20 mM Tris-HCl pH 8.5, 1 mM TCEP
GbLFY _{TERE}	pCA22	1 M NaCl, 20 mM Tris-HCl pH 8.5, 1 mM TCEP

- Annealing of competitor non-specific oligonucleotides

The lyophilized oligonucleotides from Table 6, were resuspended in autoclaved water to a concentration of 100 μ M. Annealing of both strands was done at 10 μ M as final concentration in Tris-HCl 50 mM, NaCl 150 mM (annealing buffer). The mix was incubated at 95°C for 5 minutes and then left over night at room temperature.

- Core histone labeling with Cy3 dye.

For core histone labeling, a mix containing 100 μ g of core histones, 1 mg of Cy3-NHS ester in a buffer of 2 M NaCl, 25 mM HEPES, 1 mM TCEP was incubated during 3 h at room temperature, and quenched with 1 mL Tris-HCl, 2 M NaCl. The sample was then purified in a PD-10 column (GE Healthcare), which was first equilibrated with 4 column volumes of buffer (2 M NaCl, 25 mM HEPES, 1 mM TCEP). The sample was added to the PD-10 column followed by 10 mL of buffer. Eluted histones were then collected in successive fractions.

- Site directed mutagenesis

LFY constructions for Pull down and HTRF assays were designed using site directed mutagenesis. For constructions 1 to 7 a stop codon was inserted. These constructs correspond to LFY proteins starting from the N-terminal, up to the complete protein. For constructions 8 to 10, which in contrast, start from the C-terminal and are complementary to some of the first constructs starting in the N-terminal, an amplification was performed and cloning was done by inserting sites for the restriction enzymes BamHI et XhoI. PCR reactions were carried out in a volume of 50 μ L, 40 ng of plasmid pCH2, 125 ng of each oligonucleotide, 0.2 μ M dNTP and Phusion polymerase in Phusion buffer (New England Biolabs).

PCR program:

Step 1: 98°C 30 s

Step 2: 98°C 30 s

Step 3: 55°C 1 min

Step 4: 72°C 7 min, back to step 2 x16

Step 5: hold at 16°C

Afterwards, the PCR product was digested with DpnI enzyme during 1 h at 37 °C, in order to digest methylated bacterial plasmid, and *E. coli* DH5α bacteria were transformed with the amplification product. Plasmids extracted from the transformed colonies were sent for sequencing (Eurofins) to verify the presence of the desired mutation. Oligonucleotides used for each construct are shown in Table 8.

Table 8. Primers used for site-directed mutagenesis and amplification of GST-LFY proteins.

Constructs 1 to 6 start in the N-terminal. Constructs 8 to 11, start from the C-terminal and are complementary to some of the constructs from 1 to 6. Construct 7 corresponds to the complete protein.

Primer name	LFY construction	Sequence 5' > 3'
oEBH01	1	AGCCGTCTGCTATGTACCGGCTGTTGCTGC
oEBH02	1	GCAGCAACAGCCGGTGACATAGCAGACGGCT
oEBH03	2	GTGTAGAAACGTATACCGTATGGCTAGAATAGTCCCTCTAAACCACCAA
oEBH04	2	TTGGTGGTTTAGAGGGACTATTCTAGCCATACGGTATACGTTTCTACAC
oEBH05	3	GAGTACCGGAATCACCAGCCTAGGAGAGTAGCAAATGACGG
oEBH06	3	CCGTCATTTGCTACTCTCCTAGGCTGGTGATTCCGGTACTC
oEBH07	4	GTTATTCCCCGCCGCATCCTACTGGTCTTGTTGCTGCAC
oEBH08	4	GTGCAGCAACAAGACCAGTAGGATGCGGCGGGGAATAAC
oEBH09	5	CCTCACCTTCGTTGACCTATTCGTCGGTTTCCACTG
oEBH10	5	CAGTGGAACCGACGAATAGGTCAACGAAGGTGAGG
oEBH11	6	GATGCTCCCTCTGTCACTCTGTCCCAAACC
oEBH12	6	GGTTTGGGGACAGAGTGACAGAGGGAGCATC

oEBH13	7	ATGGTGATGGCCACTCTAGAAACGCAAGTCGTC
oEBH14	7	GACGACTTGCGTTTCTAGAGTGGCCATCACCAT
oEBH16	8	GGATCCATGGAGAGACAGAGGGAGCATCCGTTT
oEBH15	8	CTCGAGTCAGAAACGCAAGTCGTCGCCGCCGCA
oEBH17	9	CGGCGACCACTAAGGCCATGAGTAGTGCGAGAA
oEBH15	9	CTCGAGTCAGAAACGCAAGTCGTCGCCGCCGCA
oEBH18	10	CGACGAAAACCTACGCTGAACCAACAAATCTC
oEBH15	10	CTCGAGTCAGAAACGCAAGTCGTCGCCGCCGCA
oEBH19	11	TGACTACGCCGCCCTTATTGCCGCCTCCTCT
oEBH15	11	CTCGAGTCAGAAACGCAAGTCGTCGCCGCCGCA

- Protein expression and purification

LFY constructs

The various LFY proteins were expressed in *E. coli* Rosetta 2 strain. 1L-bacterial cultures were grown at 37 °C until reaching an OD_{600nm} of approximately 0.7. Protein expression was induced by IPTG at a final concentration of 400 µM at 18°C overnight. Cultures were then centrifuged at 4000 x g during 30 min at 4°C and pellets were resuspended in 20 mL of 20 mM Tris-HCl pH 8 buffer. Bacterial cells were then lysed by sonication. After sonication with Sonifier 250 ultrasonic cell disruptor during 15 min, at power level of 6. The sample was then centrifuged for 30 min at 18,000 x g. The supernatant was passed through Glutathion-sepharose (GE Healthcare) column previously washed with phosphate-buffered saline (PBS), 10 mM DTT buffer. Bound proteins were eluted with three buffers with a decreasing concentration of glutathione: 50 mM Tris-HCl pH 8, 10 mM, then 50 mM and finally 100 mM glutathion buffer. Proteins were analyzed by SDS-PAGE and finally, aliquoted, frozen in liquid nitrogen with 10 % glycerol and stored at -80 °C.

BRM constructs

ETH428: (SATANA to KITATK)

SATANANNNLTLAYDIKDLICEEGAFLSKKRTDSLKKINGLLAKNLERKRIRPDLVLRQLQIEEKKRLSLDLQSR
VREEVDRQQQEIMSPDRPYRKFVRLCERQRLEMNRQVLANQKAVREKQLKTIFQWRKKLLEAHWAIR

DARTARNRGVAKYHEKMLREFSKRKDDGRNKRMEALKNNDVERYREMLLEQQTNMPGDAAERYAVLSS
FLTQTEDYLHKLGGKITATK

Number of amino acids: 231

Molecular weight: 27388.5

ETH429 : (RPDLVL to KITATK)

RPDLVLRRLQIEEKKLRSLDLQSRVREEVDRQQQEIMSPDRPYRKFVRLCERQRLEMNRQVLANQKAVRE
KQLKTIFQWRKKLLEAHWAIRDARTARNRGVAKYHEKMLREFSKRKDDGRNKRMEALKNNDVERYREML
LEQQTNMPGDAAERYAVLSSFLTQTEDYLHKLGGKITATK

Number of amino acids: 179

Molecular weight: 21599.9

Purification of BRM proteins

BRM ETH428:

Bacterial pellets were resuspended in buffer BRM: 200 mM CAPS pH 10.5, 1 mM TCEP, 1 M NaCl. Sonication and centrifugation steps were performed as for LFY constructs. Afterwards, the supernatant was passed through Nickel-sepharose (GE Healthcare) column previously equilibrated with the buffer BRM. The column was washed with 10-column volumes of BRM buffer with 50 mM imidazole. Bound proteins were eluted with a higher imidazole concentration of 300 mM.

BRM ETH429:

The same initial procedure used for BRM ETH428 was used for this construct. Afterwards, the supernatant was passed through Nickel-sepharose (GE Healthcare) column previously equilibrated with the buffer 2 M urea, 0.5 M NaCl, 1 mM TCEP. The column was washed with 10-column volumes of BRM buffer with 20 mM imidazole followed by another wash with 50 mM imidazole. Bound proteins were eluted with a higher imidazole concentration of 300 mM.

SDS-PAGE polyacrylamide gels

Protein fractions from purification steps (total extract, soluble extract, wash and elution) were analyzed by SDS-PAGE in order to verify the protein purification yield. The gel

composition was: Stacking gel 5 % acrylamide/bisacrylamide 37.5:1 (Merck), 62.5 mM Tris-HCl pH 6.8, 0.1 % (p/v) SDS, 0.1 % (p/v) ammonium persulfate (APS), 0.01 % Tetramethylethylenediamine (TEMED). Separation gel: 12% acrylamide/bisacrylamide 37.5:1, 375 mM Tris-HCl pH 8, 0.1 % (p/v) SDS, 0.1 % (p/v) APS, 0.01 % TEMED. Around 5 µg of pure proteins and around 150 µg for crude extracts were loaded on gels in denaturing buffer (50 mM Tris-HCl pH 6.8, 10 % p/v glycerol, 1% (p/v) SDS, 0.0025% (p/v) bromophenol blue, 0.4 % (p/v) DTT after incubation 5 min at 95°C. Gels were electrophoresed in Laemmli buffer (25 mM Tris-HCl, 192 mM glycine, 0.1 % (p/v) SDS). A protein ladder was also loaded (Precision Plus Protein Standards Dual Color, Bio-Rad). Electrophoresis was carried out until bromophenol blue went out of the gel at 200 V during approximately 1 h. Finally, gels were colored in a Coomassie blue solution (0.25 % (w/v) blue R 250, 45 % ethanol, 10 % acetic acid) and de-stained in a decolorizing solution (10 % ethanol, 10 % acetic acid). Stain-free SDS-page polyacrylamide gels (Biorad) were also used for the detection of Tryptophan-containing proteins. Proteins were visualized after UV-light gel activation (Biorad ChemiDoc MP Imaging System).

Protein quantification

Proteins were quantified by measuring their absorbance at 280 nm. Concentration is calculated following the Beer-Lambert law: $A = \epsilon L C$, where A is absorbance at 280 nm, ϵ is the protein's mass extinction coefficient, L is the measure of the optical distance and C , protein concentration. Absorbance measurements were done in the spectrophotometer NanoDrop-2000 (Thermo Scientific).

SDS-PAGE was also used for protein quantification. Increasing volumes of proteins are loaded on a gel and then bands intensity is compared to the intensity of a reference band for which the amount is known. This was performed with the software ImageJ (<http://rsbweb.nih.gov/ij/>).

- HTRF assay

For this assay, LFY proteins contained a GST tag and BRM protein contained a 6His tag. Two antibodies were added to the reaction. The anti-GST antibody contained the donor

fluorophore Tb³⁺. The anti-6His antibody contained the acceptor fluorophore d2. To measure the interaction between the two proteins of interest, the donor is excited at a wavelength of 337 nm. The donor then emits at a wavelength of 620 nm, and the acceptor at 665 nm (FRET). HTRF experiments were performed in Greiner 384 Flat-bottom-wells plate. After a 2 h-incubation at room temperature in the dark, emission measurements were taken with a Tecan infinite M1000PRO.

Calculations:

Ratiometric data reflecting the « raw » results:

Ratio = Fluorescence Intensity (665 nm) / Fluorescence Intensity (620 nm)

Delta ratio reflecting the « HTRF specific signal »:

DR = Ratio (sample) – Ratio (background)

- Yeast-two-hybrid assay

Preparation of yeast pre-culture

Yeast strains Y187 and AH109 were streaked on Yeast Extract Peptone Dextrose (YPD) agar plates supplemented with Adenine (0.02 g/L) and incubated 48 h at 30°C. A colony of each strain was inoculated in a pre-culture YPD medium of 5 mL supplemented with adenine and incubated 24 h at 30°C. Next, 7 drops of 60 µL of yeast pre-culture were put on YPD agar plates supplemented with adenine (yeast strains Y187 and AH109 are in different plates).

Transformation

After 24 h of incubation, two Yeast Nitrogen Base Dextrose (YNB) medium plates were poured, one for each strain. They were supplemented with the amino acid rich solution CSM –WT, glucose 2%. Then according to the vector to be transformed, leucine or tryptophan were added. From the plates containing the yeast drops, 3 different drops were scraped and resuspended in 1 mL of sterile water. This step was repeated as many times as the number of vectors to be transformed. Resuspended yeast drops were then centrifuged for 10 s at 13000 rpm. The supernatant was removed. After re-resuspension in water, the centrifugation was repeated. Finally, cells were resuspended in 500 µL of YNB solution. In a tube, 10 µL of fish DNA at 10 mg/mL and 1 µg of vector to transform were mixed gently. To this mix, the 500 µL

of resuspended cells were added and incubated at room temperature for 15 min. 60 μ L of DMSO were added to the mix and mixed gently. A heat shock was then performed at 42°C during 15 min. Cells were then centrifuged for 5 s at 13000 rpm and the supernatant removed. 200 μ L of sterile water were added to the cells in order to resuspend them gently. Cells were then spread on the respective selective plates containing CSM media (Thermo Fisher Scientific) media lacking Leucine (L) (for pGADT7-derived vectors) or lacking Tryptophan (W) (for pGBKT7-derived vectors). They were then incubated at 30 °C for 48 h.

Mating

From each transformed YNB agar petri dish, a line of colonies was scraped and diluted in 100 μ L of water. Over a squared YPD agar plate, a 5 μ L drop was made for each of the pGBKT7 vectors to be mate. After drying and on top of the drops, a 5 μ L drop of each pGADT7 vector to be mated was done. The plate was incubated at 30°C for 24 h. After this period of time, streaks and resuspensions for drops were made in order to let incubate at 30°C for 24 h. Finally, two YNB agar plates were poured, one supplemented with medium CSM –WL (lacking Tryptophan and Leucine) glucose 2% (not selecting for interaction) and another one with CSM –WLAH (lacking Tryptophan, Leucine, Adenine and Histidine) glucose 2% (selecting for interaction). Drops on both types of square plates were done for all clones and then inspected for interaction every 24 h.

- AFM

The microscope used for AFM imaging was a Multimode 8 (Bruker) run by the software Nanoscope 9.2. The imaging mode was dry ScanAsyst with a tip of Silicone nitride (ScanAsyst-AIR-HR). The scan speed was set between 1 and 1.5 Hertz with a force of around 500pN. Images were treated by the Gwyddion software, a visualization and SPM data analysis tool.

CONCLUSIONS AND PERSPECTIVES

My work consisted in studying LFY's molecular mechanism using a structural integrative biology approach. LFY is a key transcription factor in plant development, and especially in flower development in angiosperms. It has an important role, first, in the establishment of floral meristems and later, in the specification of their floral organ identities. This activity could implicate major chromatin rearrangements in cell nucleus, as target loci need to pass from a closed to an opened state.

In the last two decades, a subset of transcription factors, called Pioneer Transcription Factors (PTFs) have been and are still extensively studied because they have been shown to bind their target sites on nucleosomal DNA. PTFs are able to overcome the steric constraints that nucleosomes provide and establish a "competent" state in a particular genomic region, so it can be further regulated by other partners (Iwafuchi-Doi & Zaret, 2014). Since PTFs present these characteristics, they are key regulators for a given developmental program.

LFY has been demonstrated to physically and genetically interact with two ATPases belonging to SWI/SNF ATP-dependent chromatin remodeling complexes, SYD and BRM (Wu et al 2012). Besides, a combination of structural, biochemical and genome-wide data analyses strongly suggest that its N-terminal oligomerization domain, confers access to closed regions of chromatin (Sayou et al., 2016). In this way, LFY presents common features with PTFs. Until now, no pioneer transcription factor has been described in plants. During my thesis, I used different approaches in order to better understand LFY's mode of action in relation to SYD and BRM as well as with chromatin. This understanding could shed light to the hypothesis of LFY as a plant pioneer transcription factor.

In Chapter I, through *in vitro* experiments, LFY interaction with the basic unit of chromatin compaction, the nucleosome, was assessed. For nucleosome reconstitution, we identified nucleosome-enriched regions in Arabidopsis genome, efficiently targeted by LFY, having at least one strong LFY-binding site but having also other LFY-binding sites of lower affinity. Regions were also selected based on the loss of ChIP-seq signal from p35S:LFY_{TERE}

plants, thus, selecting regions for which oligomerization is important for LFY binding. These regions were selected using genome-wide data from ChIP-seq of LFY in overexpressing lines. Also, DNase-seq as well as MNase-seq data was useful to analyze chromatin landscape (Zhang et al., 2015; Zhang et al., 2012). A strong but non-specific binding of LFY from the gymnosperm *Ginkgo biloba* (GbLFY) to nucleosomes was observed. However, LFY from *Arabidopsis thaliana* (AtLFY), showed a weak and non-specific binding to nucleosomes. This difference in nucleosome affinity could be explained by the structural differences between these two proteins and more specifically, to the improved oligomerization capacity of GbLFY compared to AtLFY. A stronger oligomerization ability of GbLFY could confer a facilitated binding to DNA and nucleosomes, especially to genomic regions with multiple low-affinity binding sites where cooperative binding plays an important role.

Moreover, LFY's evolution has been studied in our team. Sayou et al 2014, showed that LFY underwent modifications in its DNA binding motif and structure. Like AtLFY, most LFY proteins from land plants, including angiosperms, gymnosperms, ferns, and liverworts, bind the same DNA motif (called type I). However, LFY from the moss *Physcomitrella patens* (PpLFY1) binds to a different motif (called type II), even though it contains the same 15 amino acids involved in DNA binding as AtLFY (Fig. 1B). Interestingly, the newly identified hornwort and algal LFY proteins from *Nothoceros aenigmaticus* (NaLFY) and *Klebsormidium subtile* (KsLFY) respectively, bind to a third motif (called type III) that resembles motif II, but without the central 3-bp spacer (Figure 75).

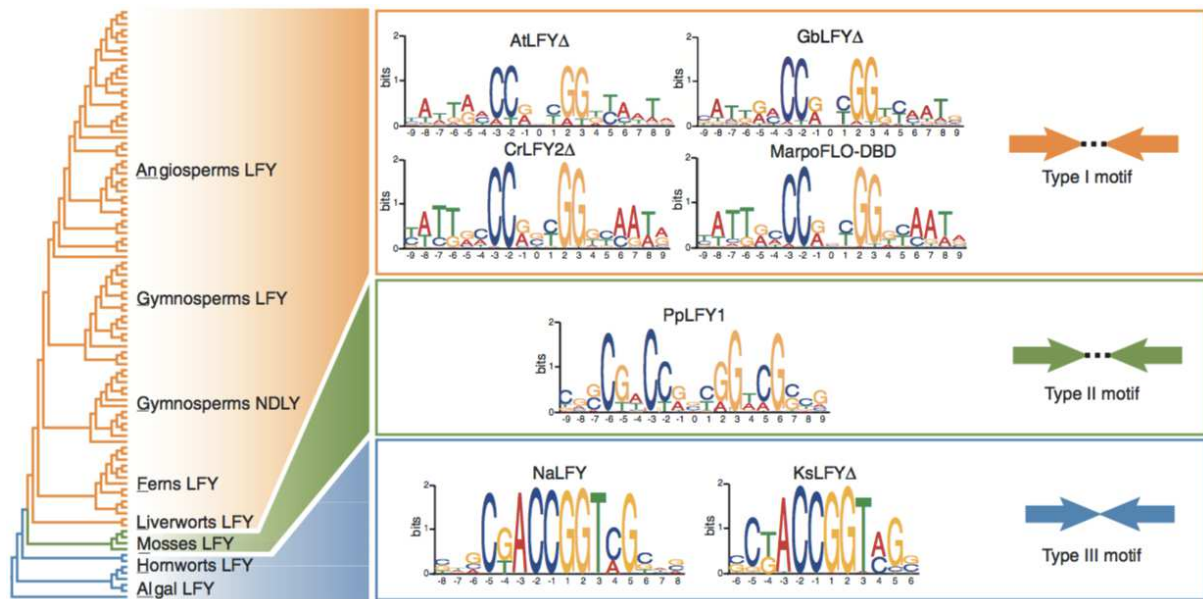


Figure 75. Simplified LEAFY phylogeny.

DNA binding specificities are color coded: type I, orange; type II, green; or type III, blue. Transcriptomes from algal species allowed the identification of six LFY homologs. (Sayou et al., 2014).

Upon modeling of LFY DBD on these motifs, it appeared by modeling LFY DBD motifs, that the interaction between helices $\alpha 1$ and $\alpha 7$, responsible for the stabilization of dimeric AtLFY and PpLFY1 DBD positioning, could no longer exist for motif III (Figure 76).

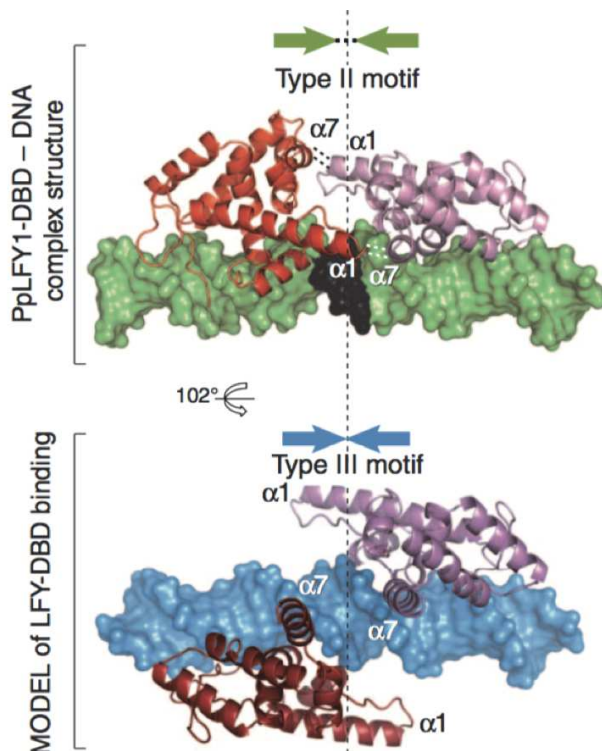


Figure 76. Type II and III LFY DBD motifs.

On top, PpLFY1-DBD dimer (in red and pink) bound to DNA (in green, except the black 3-bp spacer). Interactions between monomers (involving a helices $\alpha 1$ and $\alpha 7$) are shown with dashed lines. On the bottom, modeled type III binding with DNA shown in blue. The dashed vertical line denotes the center of the pseudopalindromic DNA sequence. Binding type III would not prone to bind nucleosomal DNA since it need both sides of the DNA chain, which in nucleosomes, one is in contact with core histones. (Sayou et al., 2014).

In the context of studying LFY's pioneer function through evolution, a possible perspective would be to use the available material of the different purified LFY DBDs and compare nucleosome binding for LFY proteins binding motif I, II and III. Since binding of motif III might be not compatible for nucleosome binding, it could be interesting to study this through EMSAs.

In Chapter II, the aim was to map the minimal necessary interacting domains of LFY and the ATPases SYD and BRM. Using the HTRF technique, LFY C-terminal domain was shown to interact with BRM HSA domain. However, the used biophysical techniques such as pull-down, gel filtration, Blitz and Biacore, were without success because proteins were not always suitable and easy to work with, for this kind of assays.

However, our Y2H assays, did not confirmed the weak interaction saw by Wu et al., (2012), between BRM and LFY, but an interaction between the BRM HSA domain and the SEP3 MADS domain protein was observed. Meaning that, in Y2H we did not see the same results published in Wu et al., (2012) for the LFY-BRM interaction but we did, for SEP3-BRM interaction.

An alternative to protein production and purification from bacteria, that would allow us to have non-aggregated proteins suitable for the mentioned biophysical tests, is the use of insect cells, which allows expressing protein with post-translational modification similar to mammalian cells. This is a technique worth to explore that could lead to the identification of the minimal interacting regions of SYD and BRM with LFY.

In addition, through an *in vivo* approach, we observed the loss of the 35S:LFY phenotype in F1 plants: 35S:LFY *syd-5/+*, 35S:LFY *brm-1/+* and 35S:LFY *brm-3/+*. This suggests a genetic interaction, and indicates that when BRM or SYD are not functional, LFY function is impaired. This result was further confirmed by looking at LFY level in these plants by western blot as well as additional crosses with non-related SALK mutant lines.

One perspective in this research line is to further study to what extent LFY binding on chromatin is affected by *brm* or *syd* mutations. Performing ChIP-seq assays to compare

binding of the 35S:LFY line alone to 35S:LFY *syd-5* or 35S:LFY *brm-3*, coupled with MNase-seq and DNaseI assays with this biological material would allow comparing LFY binding to chromatin and having a larger scope to assess the genetic interaction between LFY, and BRM or SYD.

The hypothesis of LFY acting as a pioneer transcription factor is still an open question. In a more general way, it fits into the larger question of the position that occupies the different regulatory layers for transcriptional control identified up to date, which ultimately lead to the expression of developmental genes in a specific spatiotemporal way.

REFERENCES

- Aichinger, E., Villar, C. B. R., Farrona, S., Reyes, J. C., Hennig, L., & Köhler, C. (2009). CHD3 proteins and polycomb group proteins antagonistically determine cell identity in Arabidopsis. *PLoS Genetics*. 5, 1-12. <https://doi.org/10.1371/journal.pgen.1000605>
- Alves, C., & Cunha, C. (2012). Electrophoretic Mobility Shift Assay: Analyzing Protein – Nucleic Acid Interactions. InTech, Available from: <http://www.intechopen.com/books/gel-electrophoresis-advanced-techniques/electrophoretic-mobility-shift-assay-analyzing-protein-nucleic-acid-interactions>. 205-228.
- Archacki, R., Yatusovich, R., Buszewicz, D., Krzyczmonik, K., Patryn, J., Iwanicka-Nowicka, R., Biecek, P., Wilczynski, B., Koblowska, J., Jerzmanowski, A., & Swiezewski, S. (2016). Arabidopsis SWI/SNF chromatin remodeling complex binds both promoters and terminators to regulate gene expression. *Nucleic Acids Research*. 45, 3116-3129. <https://doi.org/10.1093/nar/gkw1273>
- Bancaud, A., Wagner, G., Conde e Silva, N., Lavelle, C., Wong, H., Mozziconacci, J., Barbi, M., Sivolob, A., Le Cam, E., Mouawad, L., Viovy, J., Victor, J. & Prunell, A. (2007). Nucleosome Chiral Transition under Positive Torsional Stress in Single Chromatin Fibers. *Molecular Cell*. 27, 135-147. <https://doi.org/10.1016/j.molcel.2007.05.037>
- Becker, J. S., Nicetto, D., & Zaret, K. S. (2016). H3K9me3-Dependent Heterochromatin: Barrier to Cell Fate Changes. *Trends in Genetics*. 32, 29-41. <https://doi.org/10.1016/j.tig.2015.11.001>
- Becker, P. B., Luger, K., Rechsteiner, T. J., & Richmond, T. J. (1997). Chromatin Protocols. *Methods in Molecular Biology*. Humana Press. 119, 1-16.
- Benhamed, M., Bertrand, C., Servet, C., & Zhou, D. X. (2006). Arabidopsis GCN5, HD1, and TAF1/HAF2 Interact to Regulate Histone Acetylation Required for Light-Responsive Gene Expression. *The Plant Cell Online*. 18, 2893-2903. <https://doi.org/10.1105/tpc.106.043489>
- Bensmihen, S., To, A., Lambert, G., Kroj, T., Giraudat, J., & Parcy, F. (2004). Analysis of an activated ABI5 allele using a new selection method for transgenic Arabidopsis seeds. *FEBS Letters*. 561, 127-131. [https://doi.org/10.1016/S0014-5793\(04\)00148-6](https://doi.org/10.1016/S0014-5793(04)00148-6)
- Bernstein, E., Hake, S. B., Bernstein, E., & Hake, S. (2006). The nucleosome: a little variation goes a long way. Published on the NRC Research Press Web Site at Biochem. Cell Biol. 84, 505-517. <https://doi.org/10.1139/O06-085>
- Bezhani, S., Winter, C., Hershman, S., Wagner, J. D., Kennedy, J. F., Kwon, C. S., Pfluger, J., Su, Y., & Wagner, D. (2007). Unique, Shared, and Redundant Roles for the Arabidopsis SWI/SNF Chromatin Remodeling ATPases BRAHMA and SPLAYED. *The Plant Cell Online*. 19, 403-416. <https://doi.org/10.1105/tpc.106.048272>
- Bintu, L., Ishibashi, T., Dangkulwanich, M., Wu, Y., Lubkowska, L., Kashlev, M., & Bustamante, C. (2012). Nucleosomal elements that control the topography of the barrier to transcription. *Cell*. 151, 738-749. <https://doi.org/10.1016/j.cell.2012.10.009>
- Blázquez, M. A., Ferrándiz, C., Madueño, F., & Parcy, F. (2006). How floral meristems are built. *Plant Molecular Biology*. 60, 855-870. <https://doi.org/10.1007/s11103-006-0013-z>
- Boller, S., Ramamoorthy, S., Akbas, D., Nechanitzky, R., Burger, L., Murr, R., Schübeler, D., & Grosschedl, R. (2016). Pioneering Activity of the C-Terminal Domain of EBF1 Shapes the Chromatin Landscape for B Cell Programming. *Immunity*. 44, 527-541. <https://doi.org/10.1016/j.immuni.2016.02.021>

- Botella, C., Sautron, E., Boudiere, L., Michaud, M., Dubots, E., Yamaryo-Botté, Y., Albrieux, C., Maréchal, E., Block, M. A., & Jouhet, J. (2016). ALA10, a Phospholipid Flippase, Controls FAD2/FAD3 Desaturation of Phosphatidylcholine in the ER and Affects Chloroplast Lipid Composition in *Arabidopsis thaliana*. *Plant Physiology*. 170, 1300-1314. <https://doi.org/10.1104/pp.15.01557>
- Buenrostro, J. D., Wu, B., Chang, H. Y., & Greenleaf, W. J. (2015). ATAC-seq: A method for assaying chromatin accessibility genome-wide. *Current Protocols in Molecular Biology*. 109, 1-10. <https://doi.org/10.1002/0471142727.mb2129s109>
- Campos-Rivero, G., Osorio-Montalvo, P., Sánchez-Borges, R., Us-Camas, R., Duarte-Aké, F., & De-la-Peña, C. (2017). Plant hormone signaling in flowering: An epigenetic point of view. *Journal of Plant Physiology*. 214, 16-27. <https://doi.org/10.1016/j.jplph.2017.03.018>
- Carles, C., & Fletcher, J. C. (2009). The SAND domain protein ULTRAPETALA1 acts as a trithorax group factor to regulate cell fate in plants. *Genes and Development*. 23, 2723-2728. <https://doi.org/10.1101/gad.1812609>
- Chahtane, H., Vachon, G., Le Masson, M., Thévenon, E., Pérignon, S., Mihajlovic, N., Kalinina, A., Michard, R., Moyroud, E., Monniaux, M., Sayou, C., Grbic, V., Parcy, F., & Tichtinsky, G. (2013). A variant of LEAFY reveals its capacity to stimulate meristem development by inducing RAX1. *The Plant Journal: For Cell And Molecular Biology*, 74, 678–689. <https://doi.org/10.1111/tpj.12156>
- Chanvivatana, Y., Bishopp, A., Schubert, D., Stock, C., & Moon, Y., (2004). Interaction of Polycomb-group proteins controlling flowering in *Arabidopsis*. *Development*. 131, 5263-5276. <https://doi.org/10.1242/dev.01400>
- Cirillo, L. A., Lin, F. R., Cuesta, I., Friedman, D., Jarnik, M., & Zaret, K. S. (2002). Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Molecular Cell*. 9, 279-289. [https://doi.org/10.1016/S1097-2765\(02\)00459-8](https://doi.org/10.1016/S1097-2765(02)00459-8)
- Cirillo, L. A., McPherson, C. E., Bossard, P., Stevens, K., Cherian, S., Shim, E. Y., Clark, K. L., Burley, S. K., & Zaret, K. S. (1998). Binding of the winged-helix transcription factor HNF3 to a linker histone site on the nucleosome. *EMBO Journal*. 17, 244-254. <https://doi.org/10.1093/emboj/17.1.244>
- Clapier, C. R., & Cairns, B. R. (2009). The Biology of Chromatin Remodeling Complexes. *Annual Review of Biochemistry*. 78, 273-304. <https://doi.org/10.1146/annurev.biochem.77.062706.153223>
- Corrado, G., & Karali, M. (2009). Inducible gene expression systems and plant biotechnology. *Biotechnology Advances*. 27, 733-743. <https://doi.org/10.1016/j.biotechadv.2009.05.006>
- Daxinger, L., Hunter, B., Sheikh, M., Jauvion, V., Gascioli, V., Vaucheret, H., Matzke, M., & Furner, I. (2008). Unexpected silencing effects from T-DNA tags in *Arabidopsis*. *Trends in Plant Science*. 13, 3-6. <https://doi.org/10.1016/j.tplants.2007.10.007>
- Denay, G., Chahtane, H., Tichtinsky, G., & Parcy, F. (2017). A flower is born: an update on *Arabidopsis* floral meristem formation. *Current Opinion in Plant Biology*. 35, 15-22. <https://doi.org/10.1016/j.pbi.2016.09.003>
- Dhalluin, C., Carlson, J. E., Zeng, L., He, C., Aggrwal, A. K., & Zhou, M. (1999). Structure and ligand of a histone acetyltransferase bromodomain. *Nature*, 399, 491–496.
- Efroni, I., Han, S. K., Kim, H. J., Wu, M. F., Steiner, E., Birnbaum, K. D., Hong, J. C., Eshed, Y., & Wagner, D. (2013). Regulation of leaf maturation by chromatin-mediated modulation of cytokinin responses. *Developmental Cell*. 24, 438, 445. <https://doi.org/10.1016/j.devcel.2013.01.019>
- Engelhorn, J., Blanvillain, R., & Carles, C. (2014). Gene activation and cell fate control in plants: A chromatin

- perspective. *Cellular and Molecular Life Sciences*. 71, 3119–3137. <https://doi.org/10.1007/s00018-014-1609-0>
- Engelhorn, J., Blanvillain, R., Kröner, C., Parrinello, H., Rohmer, M., Pose, D., Ott, F., Schmid, M., & Carles, C. (2017). Dynamics of H3K4me3 chromatin marks take the lead over H3K27me3 for gene regulation during flower morphogenesis in *Arabidopsis thaliana*. *Epigenomes*. 1, 1-8.
- Esch, D., Vahokoski, J., Groves, M. R., Pogenberg, V., Cojocaru, V., Vom Bruch, H., Han, D., Drexler, H. C. A., Araúzo-Bravo, M. J., Ng, C. K. L., Jauch, R., Wilmanns, M., & Schöler, H. R. (2013). A unique Oct4 interface is crucial for reprogramming to pluripotency. *Nature Cell Biology*. 15, 295-302. <https://doi.org/10.1038/ncb2680>
- Fan, H. F., Liu, Z. N., Chow, S. Y., Lu, Y. H., & Li, H. (2015). Histone chaperone-mediated nucleosome assembly process. *PLoS ONE*. 10, 1-16. <https://doi.org/10.1371/journal.pone.0115007>
- Farrona, S., Coupland, G., & Turck, F. (2008). The impact of chromatin regulation on the floral transition. *Seminars in Cell and Developmental Biology*. 19, 560-573. <https://doi.org/10.1016/j.semcdb.2008.07.015>
- Foster, C. S. P. (2016). The evolutionary history of flowering plants. *Journal & Proceedings of the Royal Society of New South Wales*, 149, 65–82.
- Friedman, W. E., & Floyd, S. K. (2001). Perspective: The origin of flowering plants and their and their reproductive biology - A tale of two phylogenies. *Evolution*, 55 (552), 217–231.
- Gallois, J. L., Nora, F. R., Mizukami, Y., & Sablowski, R. (2004). WUSCHEL induces shoot stem cell activity and developmental plasticity in the root meristem. *Genes and Development*. 18, 375-380.
- Gentry, M., & Hennig, L. (2014). Remodelling chromatin to shape development of plants. *Experimental Cell Research*. 321, 40-46. <https://doi.org/10.1016/j.yexcr.2013.11.010>
- Gerard, F. C. A., Ribeiro, E. de A., Leyrat, C., Ivanov, I., Blondel, D., Longhi, S., Ruigrok, R. W. H., & Jamin, M. (2009). Modular Organization of Rabies Virus Phosphoprotein. *Journal of Molecular Biology*. 388, 978-996. <https://doi.org/10.1016/j.jmb.2009.03.061>
- Gilbert S. F., *Developmental Biology*. 6th edition. Sunderland (2000). The Vegetative-to-Reproductive Transition. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK10122/>
- Goodrich, J., Puangsomelee, P., Martin, M., Long, D., Meyerowitz, E., & Coupland, G. (1997). A polycomb-group gene regulates homeotic gene expression in *Arabidopsis*. *Nature*, 386 (6), 44–51.
- Green, S. M., Coyne, H. J., McIntosh, L. P., & Graves, B. J. (2010). DNA binding by the ETS protein TEL (ETV6) is regulated by autoinhibition and self-association. *Journal of Biological Chemistry*. 285, 18496-18504. <https://doi.org/10.1074/jbc.M109.096958>
- Guo, C., & Morris, S. A. (2017). Engineering cell identity: establishing new gene regulatory and chromatin landscapes. *Current Opinion in Genetics and Development*. 46, 50-57. <https://doi.org/10.1016/j.gde.2017.06.011>
- Hamès, C., Ptchelkine, D., Grimm, C., Thevenon, E., Moyroud, E., Gérard, F., Martiel, J., Benlloch, R., Parcy, F., & Müller, C. W. (2008). Structural basis for LEAFY floral switch function and similarity with helix-turn-helix proteins. *The EMBO Journal*. 27, 2628-2637. <https://doi.org/10.1038/emboj.2008.184>
- Han, S. K., Wu, M. F., Cui, S., & Wagner, D. (2015). Roles and activities of chromatin remodeling ATPases in plants. *Plant Journal*. 83, 62-67. <https://doi.org/10.1111/tpj.12877>
- Hargreaves, D. C., & Crabtree, G. R. (2011). ATP-dependent chromatin remodeling: genetics, genomics and

- mechanisms. *Cell Research*. 21, 396-420. <https://doi.org/10.1038/cr.2011.32>
- Havas, K., Flaus, A., Phelan, M., Kingston, R., Wade, P. A., Lilley, D. M. J., & Owen-Hughes, T. (2000). Generation of superhelical torsion by ATP-dependent chromatin remodeling activities. *Cell*. 103, 1133-1142. [https://doi.org/10.1016/S0092-8674\(00\)00215-4](https://doi.org/10.1016/S0092-8674(00)00215-4)
- Henikoff, S., & Ahmad, K. (2005). Assembly of variant histones into chromatin. *Annual Review of Cell and Developmental Biology*. 21, 133-153. <https://doi.org/10.1146/annurev.cellbio.21.012704.133518>
- Ho, J. W. K., Jung, Y. L., Liu, T., Alver, B. H., Lee, S., Ikegami, K., Sohn, K., Minoda, A., Tolstorukov, M. Y., Appert, A., ... & Park, P. J. (2014). Comparative analysis of metazoan chromatin organization. *Nature*. 512, 449-452. <https://doi.org/10.1038/nature13415>
- Hurtado, L., Farrona, S., & Reyes, J. C. (2006). The putative SWI/SNF complex subunit BRAHMA activates flower homeotic genes in *Arabidopsis thaliana*. *Plant Molecular Biology*. 62, 291-304. <https://doi.org/10.1007/s11103-006-9021-2>
- Ingouff, M., & Berger, F. (2010). Histone 3 variants in plants. *Chromosoma*. 119, 27-33. <https://doi.org/10.1007/s00412-009-0237-1>
- Irish, V. F. (2010). The flowering of *Arabidopsis* flower development. *Plant Journal*. 61, 1014-1028. <https://doi.org/10.1111/j.1365-313X.2009.04065>
- Iwafuchi-Doi, M., Donahue, G., Kakumanu, A., Watts, J. A., Mahony, S., Pugh, B. F., Lee, D., Kaestner, K. H., & Zaret, K. S. (2016). The Pioneer Transcription Factor FoxA Maintains an Accessible Nucleosome Configuration at Enhancers for Tissue-Specific Gene Activation. *Molecular Cell*. 62, 79-91. <https://doi.org/10.1016/j.molcel.2016.03.001>
- Iwafuchi-Doi, M., & Zaret, K. S. (2014). Pioneer transcription factors in cell reprogramming. *Genes and Development*. 28, 2679-2692. <https://doi.org/10.1101/gad.253443.114>
- Kærn, M., Elston, T. C., Blake, W. J., & Collins, J. J. (2005). Stochasticity in gene expression: from theories to phenotypes. *Nature Reviews Genetics*. 6, 451-464. <https://doi.org/10.1038/nrg1615>
- Keller, T. (2006). *Arabidopsis* REGULATOR OF AXILLARY MERISTEMS1 Controls a Leaf Axil Stem Cell Niche and Modulates Vegetative Development. *The Plant Cell Online*. 18, 598-611. <https://doi.org/10.1105/tpc.105.038588>
- Kost, T. A., Condreay, J. P., & Jarvis, D. L. (2005). Baculovirus as versatile vectors for protein expression in insect and mammalian cells. *Nature Biotechnology*. 23, 567-575. <https://doi.org/10.1038/nbt1095>
- Krizek, B. A., & Fletcher, J. C. (2005). Molecular mechanisms of flower development: an armchair guide. *Nature Reviews Genetics*. 6, 688-698. <https://doi.org/10.1038/nrg1675>
- Kwon, C. S. (2006). A role for chromatin remodeling in regulation of CUC gene expression in the *Arabidopsis* cotyledon boundary. *Development*. 133, 3223-3230. <https://doi.org/10.1242/dev.02508>
- Lavelle, C., & Prunell, A. (2017). Chromatin Polymorphism and the Nucleosome Superfamily: A Genealogy. *Cell Cycle*. 6 (17), 2113–2119. <https://doi.org/10.4161/cc.6.17.4631>
- Lee, Jung-Youn, Yoo, Byung-Chun, Rojas, Maria R., Gomez-Ospina, Natalia, Staehelin, Andrew, & Lucas, W. J. (2003). Selective Trafficking of Non-Cell-Autonomous Proteins Mediated by NtNCAPP1. *Science*. 299 (5605), 392–396. <https://doi.org/10.1126/science.1077813>
- Lee, Y. T., Gibbons, G., Lee, S. Y., Nikolovska-Coleska, Z., & Dou, Y. (2015). One-pot refolding of core histones from bacterial inclusion bodies allows rapid reconstitution of histone octamer. *Protein Expression and*

- Purification. *110*, 89-94. <https://doi.org/10.1016/j.pep.2015.02.007>
- Li, B., Carey, M., & Workman, J. L. (2007). The Role of Chromatin during Transcription. *Cell*. *128*, 707-719. <https://doi.org/10.1016/j.cell.2007.01.015>
- Li, C., Chen, C., Gao, L., Yang, S., Nguyen, V., Shi, X., & Cui, Y. (2015). The Arabidopsis SWI2/SNF2 Chromatin Remodeler BRAHMA Regulates Polycomb Function during Vegetative Development and Directly Activates the Flowering Repressor Gene SVP. *PLoS Genetics*. *11*, 1-25. <https://doi.org/10.1371/journal.pgen.1004944>
- Li Qiao, & Wrangé Orjan. (1997). Methods: A Companion to Methods in Assays for Transcription Factors Access to Nucleosomal DNA. *Enzymology*. *12*, 96-104.
- Lone, I. N., Shukla, M. S., Charles Richard, J. L., Peshev, Z. Y., Dimitrov, S., & Angelov, D. (2013). Binding of NF- κ B to Nucleosomes: Effect of Translational Positioning, Nucleosome Remodeling and Linker Histone H1. *PLoS Genetics*. *9*, 1-12. <https://doi.org/10.1371/journal.pgen.1003830>
- Lowary, P., & Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *Journal of Molecular Biology*. *276*, 19-42. <https://doi.org/10.1006/jmbi.1997.1494>
- Machida, Y., Fukaki, H., & Araki, T. (2013). Plant meristems and organogenesis: The new era of plant developmental research. *Plant and Cell Physiology*. *54*, 295-301. <https://doi.org/10.1093/pcp/pct034>
- Malik, H. S., & Henikoff, S. (2003). Phylogenomics of the nucleosome. *Nature Structural Biology*. *10*, 882-891. <https://doi.org/10.1038/nsb996>
- Mariño-Ramírez, L., Hsu, B., Baxevanis, A. D., & Landsman, D. (2006). The histone database: A comprehensive resource for histones and histone fold-containing proteins. *Proteins: Structure, Function and Genetics*. *48*, 1-7. <https://doi.org/10.1002/prot.20814>
- Mlotshwa, S., Pruss, G. J., Gao, Z., Mgtshini, N. L., Li, J., Chen, X., & Vance, V. (2010). Transcriptional silencing induced by Arabidopsis T-DNA mutants is associated with 35S promoter siRNAs and requires genes involved in siRNA-mediated chromatin silencing. *Plant Journal*. *64*, 699-704. <https://doi.org/10.1111/j.1365-3113.2010.04358.x>
- Moreau, F., Theenon, E., Blanvillain, R., Lopez-Vidriero, I., Franco-Zorrilla, J. M., Dumas, R., Parcy, F., Morel, P., Trehin, C., & Carles, C. (2016). The Myb-domain protein ULTRAPETALA1 INTERACTING FACTOR 1 controls floral meristem activities in Arabidopsis. *Development*. *144*, 1497-1505. <https://doi.org/10.1242/dev.127365>
- Moshe, A., & Kaplan, T. (2017). Genome-wide search for Zelda-like chromatin signatures identifies GAF as a pioneer factor in early fly development. *Epigenetics & Chromatin*. *10*, 1-14. <https://doi.org/10.1186/s13072-017-0141-5>
- Moyroud, E., Minguet, E. G., Ott, F., Yant, L., Posé, D., Monniaux, M., & Parcy, F. (2011). Prediction of Regulatory Interactions from Genome Sequences Using a Biophysical Model for the Arabidopsis LEAFY Transcription Factor. *The Plant Cell*. *23*, 1293-1306. <https://doi.org/10.1105/tpc.111.083329>
- Moyroud, E., Monniaux, M., Thévenon, E., Dumas, R., Scutt, C. P., Frohlich, M. W., & Parcy, F. (2017). A link between LEAFY and B-gene homologues in *Welwitschia mirabilis* sheds light on ancestral mechanisms prefiguring floral development. *New Phytologist*. *216*, 469-481 <https://doi.org/10.1111/nph.14483>
- Muller, D., Schmitz, G., & Theres, K. (2006). Blind Homologous R2R3 Myb Genes Control the Pattern of Lateral Meristem Initiation in Arabidopsis. *The Plant Cell Online*. *18*, 586-597.

<https://doi.org/10.1105/tpc.105.038745>

- Ngo, T. M., Zhang, Q., Zhou, R., Yodh, J. G., & Ha, T. (2015). Asymmetric unwrapping of nucleosomes under tension directed by DNA local flexibility. *Cell*. *160*, 1135–1144. <https://doi.org/10.1016/j.cell.2015.02.001>
- Okuwaki, M., Kato, K., Shimahara, H., Tate, S. I., & Nagata, K. (2005). Assembly and Disassembly of Nucleosome Core Particles Containing Histone Variants by Human Nucleosome Assembly Protein I. *Molecular and Cellular Biology*. *25* (23), 10639–10651. [https://doi.org/10.1128/MCB.25.23.10639–10651.2005](https://doi.org/10.1128/MCB.25.23.10639-10651.2005)
- Olins, D. E. & Olins, A. L. (2003). Chromatin history: our view from the bridge. *Nature Reviews*. *4*, 809–814. Retrieved from www.nature.com/reviews/molcellbio
- Oliver, M. J., Tuba, Z., & Mishler, B. D. (2000). The evolution of vegetative desiccation tolerance in land plants. *Plant Ecology*. *151*, 85–100.
- Ordu, O., Lusser, A., & Dekker, N. H. (2016). Recent insights from in vitro single-molecule studies into nucleosome structure and dynamics. *Biophysical Reviews*. *8*, 33–49. <https://doi.org/10.1007/s12551-016-0212-z>
- Pajoro, A., Madrigal, P., Muiño, J. M., Matus, J., Jin, J., Mecchia, M. A., Debernardi, J. M., Palatnik, J. F., Balazadeh, S., Arif, M., O'Maoileidigh, D. S., Wellmer, F., Krajewski, P., Riechmann, J., Angenent, G. C., & Kaufmann, K. (2014). Dynamics of chromatin accessibility and gene regulation by MADS-domain transcription factors in flower development. *Genome Biology*. *1*–18. <https://doi.org/10.1186/gb-2014-15-3-r4115>
- Pelaz, S., Ditta, G. S., Baumann, E., Wisman, E., & Yanofsky, M. (2000). B and C floral organ identity functions require SEPALLATA MADS-box genes. *Nature*. *405*, 200–203.
- Peterson, C. L., & Herskowitz, I. (1992). Characterization of the yeast SWI1, SWI2, and SWI3 genes, which encode a global activator of transcription. *Cell*. *68*, 573–583. [https://doi.org/10.1016/0092-8674\(92\)90192-F](https://doi.org/10.1016/0092-8674(92)90192-F)
- Pfluger, J., & Wagner, D. (2007). Histone modifications and dynamic regulation of genome accessibility in plants. *Current Opinion in Plant Biology*. *10*, 645–652. <https://doi.org/10.1016/j.pbi.2007.07.013>
- Phelan, M. L., Schnitzler, G. R., & Kingston, R. E. (2000). Octamer Transfer and Creation of Stably Remodeled Nucleosomes by Human SWI-SNF and Its Isolated ATPase. *Molecular & Cellular Biology*. *20* (17), 6380–6389.
- Phillips, T. & Hoopes, L. (2008) Transcription factors and transcriptional control in eukaryotic cells. *Nature Education*. *1*, 110–119.
- Polo, S. E., & Almouzni, G. (2006). Chromatin assembly: A basic recipe with various flavours. *Current Opinion in Genetics and Development*. <https://doi.org/10.1016/j.gde.2006.02.011>
- Posé, D., Yant, L., & Schmid, M. (2012). The end of innocence: Flowering networks explode in complexity. *Current Opinion in Plant Biology*. *15*, 45–50. <https://doi.org/10.1016/j.pbi.2011.09.002>
- Pugh, B. F. (2010). A preoccupied position on nucleosomes. *Nature Structural & Molecular Biology*. *17*, 918–923. <https://doi.org/10.1038/nsmb0810-923>
- Risseuw, E., Venglat, P., Xiang, D., Komendant, K., Daskalchuk, T., Babic, V., & Datla, R. (2013). An activated form of UFO alters leaf development and produces ectopic floral and inflorescence meristems. *PLoS One*. *8*, 1–12. <https://doi.org/10.1371/journal.pone.0083807>
- Rosas-Guerrero, V., Aguilar, R., Martén-Rodríguez, S., Ashworth, L., Lopezaraiza-Mikel, M., Bastida, J. M., & Quesada, M. (2014). A quantitative review of pollination syndromes: Do floral traits predict effective

- pollinators? *Ecology Letters*. *17*, 388-400. <https://doi.org/10.1111/ele.12224>
- Saleh, A., Al-Abdallat, A., Ndamukong, I., Alvarez-Venegas, R., & Avramova, Z. (2007). The Arabidopsis homologs of trithorax (ATX1) and enhancer of zeste (CLF) establish “bivalent chromatin marks” at the silent AGAMOUS locus. *Nucleic Acids Research*. *35*, 6290-6296. <https://doi.org/10.1093/nar/gkm464>
- Sarnowski, T. J., Â wiez Ćewski, S. S., Pawlikowska, K., Kaczanowski, S., & Jerzmanowski, A. (2002). AtSWI3B, an Arabidopsis homolog of SWI3, a core subunit of yeast Swi/Snf chromatin remodeling complex, interacts with FCA, a regulator of flowering time. *Nucleic Acids Research*. *30*, 3412-4321.
- Sayou, C., Nanao, M. H., Jamin, M., Posé, D., Thévenon, E., Grégoire, L., Tichtinsky, G., Denay, G., Ott, F., Periaets Llobet, M., Schmid, M., Dumas, R., & Parcy, F. (2016). A SAM oligomerization domain shapes the genomic binding landscape of the LEAFY transcription factor. *Nature Communications*. *7*, 1-12. <https://doi.org/10.1038/ncomms11222>
- Schmitz, R. J., & Amasino, R. M. (2007). Vernalization: A model for investigating epigenetics and eukaryotic gene regulation in plants. *Biochimica et Biophysica Acta - Gene Structure and Expression*. *1769*, 269-285. <https://doi.org/10.1016/j.bbaexp.2007.02.003>
- Schubert, H. L., Wittmeyer, J., Kasten, M. M., Hinata, K., Rawling, D. C., Héroux, A., Cairns, B. R., & Hill, C. P. (2013). Structure of an actin-related subcomplex of the SWI/SNF chromatin remodeler. *110*, 3345-3350. <https://doi.org/10.1073/pnas.1215379110>
- Sekiya, T., Muthurajan, U. M., Luger, K., Tulin, A. V., & Zaret, K. S. (2009). Nucleosome-binding affinity as a primary determinant of the nuclear mobility of the pioneer transcription factor FoxA. *Genes and Development*. *23*, 804-809. <https://doi.org/10.1101/gad.1775509>
- Shaked, H., Avivi-Ragolsky, N., & Levy, A. (2006). Involvement of the arabidopsis SWI2/SNF2 chromatin remodeling gene family in DNA damage response and recombination. *Genetics*. *173*, 985-994. <https://doi.org/10.1534/genetics.105.051664>
- Shim, Y., Duan, M. R., Chen, X., Smerdon, M. J., & Min, J. H. (2012). Polycistronic coexpression and nondenaturing purification of histone octamers. *Analytical Biochemistry*. *427*, 190-192. <https://doi.org/10.1016/j.ab.2012.05.006>
- Silva, C. S., Melzer, R., Dornelas, M. C., Ruempler, F., Zubieta, C., Puranik, S., Hugouvieux, V. & 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. *6*, 1-18. <https://doi.org/10.3389/fpls.2015.01193>
- Simpson, G. G., & Dean, C. (2002). Arabidopsis, the Rosetta Stone of Flowering Time? *Science*. *296*, 295-289.
- Smaczniak, C., Immink, R. G. H., Muiño, J. M., Blanvillain, R., Busscher, M., Busscher-Lange, J., Dinh, Q. D., Liu, S., Westphal, A. H., Boeren, S., Parcy, F., Xu, L., Carles, C. C., Angenent, G. C., Kaufmann K. (2011). Characterization of MADS-domain transcription factor complexes in Arabidopsis flower development. *PNAS*. *109*, 1560-1565. <https://doi.org/10.1073/pnas.1112871109>
- Soufi, A., Garcia, M. F., Jaroszewicz, A., Osman, N., Pellegrini, M., & Zaret, K. S. (2015). Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell*. *161*, 555-569. <https://doi.org/10.1016/j.cell.2015.03.017>
- Sullivan, A. M., Arsovski, A. A., Lempe, J., Bubbs, K. L., Weirauch, M. T., Sabo, P. J., Sandstorm, R., Thruman, R. E., Neph, S., Reynolds, A. P., Stergachis, A. B., Vernot, B., Johnson, A. K., Haugen, E., Sullivan, S. T., Thompson, A., Neri III, F. V., Weaver, M., Diegel, M., Mnaimneh, S., Yang, A., Hughes T. R., Nemhauser, J. L., Quietsch,

- C., & Stamatoyannopoulos J. A. (2014). Mapping and dynamics of regulatory DNA and transcription factor networks in *A. thaliana*. *Cell Reports*. 8, 2015-2030. <https://doi.org/10.1016/j.celrep.2014.08.019>
- Sung, S., & Amasino, R. M. (2006). Molecular genetic studies of the memory of winter. *Journal of Experimental Botany*. 57, 3369-3377. <https://doi.org/10.1093/jxb/erl105>
- Takahashi, K., & Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*. 126, 663-676. <https://doi.org/10.1016/j.cell.2006.07.024>
- Tang, L., Nogales, E., & Ciferri, C. (2010). Structure and function of SWI/SNF chromatin remodeling complexes and mechanistic implications for transcription. *Progress in Biophysics and Molecular Biology*. 102, 122-128. <https://doi.org/10.1016/j.pbiomolbio.2010.05.001>
- Tang, X., Hou, A., Babu, M., Nguyen, V., Hurtado, L., Lu, Q., Reyes, J. C., Wang, A., Keller, W. A., Harada, J. J., Tsang, E. W. T., & Cui, Y. (2008). The Arabidopsis BRAHMA Chromatin-Remodeling ATPase Is Involved in Repression of Seed Maturation Genes in Leaves. *Plant Physiology*. 147, 1143-1157. <https://doi.org/10.1104/pp.108.121996>
- Theissen, G. (2001). Development of floral organ identity: Stories from the MADS house. *Current Opinion in Plant Biology*. 4, 75-85. [https://doi.org/10.1016/S1369-5266\(00\)00139-4](https://doi.org/10.1016/S1369-5266(00)00139-4)
- Torres, E. S., & Deal, R. B. (2018). The histone variant H2A.Z and chromatin remodeler BRAHMA act coordinately and antagonistically to regulate transcription and nucleosome dynamics in Arabidopsis. *BioRxiv*. 1-34. <https://doi.org/10.1101/243790>
- Treiber, N., Treiber, T., Zocher, G., & Grosschedl, R. (2010). Structure of an Ebf1:DNA complex reveals unusual DNA recognition and structural homology with Rel proteins. *Genes and Development*. 24, 2270-2275. <https://doi.org/10.1101/gad.1976610>
- Tsukiyama, T., & Wu, C. (1995). Purification and properties of an ATP-dependent nucleosome remodeling factor. 83, 1011-1020. *Cell*. [https://doi.org/10.1016/0092-8674\(95\)90216-3](https://doi.org/10.1016/0092-8674(95)90216-3)
- Van der Niet, T., & Johnson, S. D. (2012). Phylogenetic evidence for pollinator-driven diversification of angiosperms. *Trends in Ecology and Evolution*. 27, 353-361. <https://doi.org/10.1016/j.tree.2012.02.002>
- Vercruyssen, L., Verkest, A., Gonzalez, N., Heyndrickx, K. S., Eeckhout, D., Han, S. K., Jégu, T., Archaki, R., Van Leene, J., Andriankaja, M., De Bodt, S., Abeel, T., Coppens, F., Dhondt, S., De Milde, L., Vermeersch M., Malux, K., Gevaert, K., Jerzmanowski, A., Benhamed M., Wagner, D., Vandepoele, K., De Jaeger, G., & Inze, D. (2014). ANGUSTIFOLIA3 Binds to SWI/SNF Chromatin Remodeling Complexes to Regulate Transcription during Arabidopsis Leaf Development. *The Plant Cell*. 26, 210-229. <https://doi.org/10.1105/tpc.113.115907>
- Vignali, M., Hassan, A. H., Neely, K. E., & Workman, J. L. (2000). Minireview. ATP-Dependent Chromatin-Remodeling Complexes. *Molecular & Cellular Biology*. 20 (6), 1899–1910.
- Voss, T. C., & Hager, G. L. (2014). Dynamic regulation of transcriptional states by chromatin and transcription factors. *Nature Reviews. Genetics*. 15 (2), 69–81. <https://doi.org/10.1038/nrg3623>
- Wagner, D., & Meyerowitz, E. M. (2002). SPLAYED, a Novel SWI/SNFATPase Homolog, Controls Reproductive Development in Arabidopsis. *Current Biology*. 12, 85–94.
- Weigel, Detlef & Nilsson, O. (1995). A developmental switch sufficient for flower initiation in diverse plants. *Nature*. 377, 495–500.
- Weigel, D., Alvarez, J., Smyth, D. R., Yanofsky, M. F., & Meyerowitz, E. M. (1992). LEAFY controls floral meristem

- identity in Arabidopsis. *Cell*. 69, 843-859. [https://doi.org/10.1016/0092-8674\(92\)90295-N](https://doi.org/10.1016/0092-8674(92)90295-N)
- Winter, C. M., Austin, R. S., Blanvillain-Baufumé, S., Reback, M. A., Monniaux, M., Wu, M. F., Sang, Y., Yamaguchi, A., Yamaguchi M., Parker, J. E., Parcy, F., Jensen, S. T., Li, H., & Wagner, D. (2011). LEAFY Target Genes Reveal Floral Regulatory Logic, cis Motifs, and a Link to Biotic Stimulus Response. *Developmental Cell*. 20, 430-443. <https://doi.org/10.1016/j.devcel.2011.03.019>
- Woodcock, C. L. (2006). Chromatin architecture. *Current Opinion in Structural Biology*. 16, 213-220. <https://doi.org/10.1016/j.sbi.2006.02.005>
- Wu, M. F., Sang, Y., Bezhani, S., Yamaguchi, N., Han, S. K., Li, Z., & Jacobsen, S. E. (2012). SWI2/SNF2 chromatin remodeling ATPases overcome polycomb repression and control floral organ identity with the LEAFY and SEPALLATA3 transcription factors. *109*, 3576-3581. <https://doi.org/10.1073/pnas.1113409109>
- Wu, M. F., Yamaguchi, N., Xiao, J., Bargmann, B., Estelle, M., Sang, Y., & Wagner, D. (2015). Auxin-regulated chromatin switch directs acquisition of flower primordium founder fate. *eLife*. 4, 1-19. <https://doi.org/10.7554/eLife.09269>
- Zaret, K. S., & Mango, S. E. (2016). Pioneer transcription factors, chromatin dynamics, and cell fate control. *Current Opinion in Genetics & Development*. 37, 76-81. <https://doi.org/10.1016/j.gde.2015.12.003>
- Zhang, T., Cooper, S., & Brockdorff, N. (2015). The interplay of histone modifications – writers that read. *EMBO Reports*. 16, 1467–1481. <https://doi.org/10.15252/embo>
- Zhang, T., Zhang, W., & Jiang, J. (2015). Genome-Wide Nucleosome Occupancy and Positioning and Their Impact on Gene Expression and Evolution in Plants. *Plant Physiology*. 168, 1406-1416. <https://doi.org/10.1104/pp.15.00125>
- Zhang, W., Zhang, T., Wu, Y., & Jiang, J. (2012). Genome-Wide Identification of Regulatory DNA Elements and Protein-Binding Footprints Using Signatures of Open Chromatin in Arabidopsis. *The Plant Cell*. 24, 2719-2731. <https://doi.org/10.1105/tpc.112.098061>
- Zhang, Y., Smith, C. L., Saha, A., Grill, S. W., Mihadja, S., Smith, S. B., & Bustamante, C. (2006). DNA translocation and loop formation mechanism of chromatin remodeling by SWI/SNF and RSC. *Molecular Cell*. 24, 559-568. <https://doi.org/10.1016/j.molcel.2006.10.025>
- Zheng, B., & Chen, X. (2011). Dynamics of histone H3 lysine 27 trimethylation in plant development. *Current Opinion in Plant Biology*. 14, 123-129. <https://doi.org/10.1016/j.pbi.2011.01.001>
- Zook, E. C., Ramirez, K., Guo, X., van der Voort, G., Sigvardsson, M., Svensson, E. C., Fu, Y., & Kee, B. L. (2016). The ETS1 transcription factor is required for the development and cytokine-induced expansion of ILC2. *The Journal of Experimental Medicine*. 213, 687-696. <https://doi.org/10.1084/jem.20150851>

Abstract

LFY is a key transcription factor in plant development, and especially in flowering for angiosperms. It has an important role, first, in the establishment of floral meristems and later, in the specification of their floral organ identities. This activity implicates on cells' nucleus major chromatin rearrangements. Target loci need to pass from a closed to an opened state. In the last two decades, Pioneer Transcription Factors (PTFs) have been studied because they can bind their target sites at nucleosomal DNA, they are able to overcome the steric constraints of nucleosomes and establish a "competent state" in a particular region, so it can be further regulated by other partners (Iwafuchi-Doi & Zaret, 2014). LFY has been demonstrated to physically and genetically interact with two ATPases belonging to ATP-dependent chromatin remodeling complexes, SYD and BRM (Wu et al., 2012). Besides, genome-wide data analyses strongly suggest that its N-terminal oligomerization domain, confers LFY access to closed regions of chromatin (Sayou et al., 2016). In this way, LFY presents common features with PTFs. We worked in order to better understand LFY's mode of action in relation to the mentioned ATPases as well as with chromatin.

In Chapter I, through *in vitro* experiments, LFY's potential interaction with nucleosomes, was assessed. We performed reconstituted nucleosomes by identifying nucleosome-enriched regions in Arabidopsis genome, efficiently targeted by LFY. These regions were selected used genome-wide data from ChIP-seq of LFY in overexpressing lines and DNase-seq as well as MNase-seq data, which was useful to analyze chromatin landscape (Zhang, Zhang, & Jiang, 2015; Zhang, Zhang, Wu, & Jiang, 2012). A strong but non-specific binding of LFY from the gymnosperm *Ginkgo biloba* to nucleosomes was observed. However, LFY from *Arabidopsis thaliana*, showed a weak binding to nucleosomes.

In Chapter II, the aim was to map the minimal necessary interacting domains of LFY and the ATPases SYD and BRM. Using the HTRF technique, LFY's C-terminal domain was shown to interact with BRM's HSA domain. In addition, through an *in vivo* approach, we observed the loss of the 35S:LFY phenotype in the F1 plants from each of the three crosses: 35S:LFY *syd-5*, 35S:LFY *brm-1* and 35S:LFY *brm-3*. This suggested a strong interaction, meaning that when BRM or SYD are not functional, LFY's function is affected and no ectopic flowers are formed.

Résumé

LFY est un facteur de transcription clé dans le développement des plantes, et en particulier dans la floraison des angiospermes. Il a un rôle important, d'abord, dans l'établissement des méristèmes floraux et plus tard, dans la spécification de leurs identités d'organes floraux. Cette activité implique des réarrangements majeurs de la chromatine dans le noyau des cellules. Des loci cibles doivent passer d'un état fermé à un état ouvert. Au cours des deux dernières décennies, les Facteurs de Transcription Pionniers (PTF) ont été étudiés car ils peuvent lier leurs sites cibles à l'ADN nucléosomique, ils peuvent surmonter les contraintes stériques des nucléosomes et établir un état «compétent» dans une région particulière pour qu'il puisse être davantage régulé par d'autres partenaires (Iwafuchi-Doi & Zaret, 2014). Il a été démontré que LFY interagit physiquement et génétiquement avec deux ATPases appartenant à des complexes de remodelage de la chromatine ATP-dépendants, SYD et BRM (Wu et al., 2012). En outre, des analyses de données à l'échelle du génome suggèrent fortement que son domaine d'oligomérisation N-terminal, confère à LFY un accès à des régions fermées de la chromatine (Sayou et al., 2016). De cette manière, LFY présente des caractéristiques communes avec les PTF. Nous avons travaillé afin de mieux comprendre le mode d'action de LFY par rapport aux ATPases mentionnées ainsi qu'à la chromatine.

Au chapitre I, à travers des expériences *in vitro*, l'interaction potentielle de LFY avec les nucléosomes a été évaluée. Nous avons reconstitué des nucléosomes, en identifiant des régions enrichies en nucléosomes dans le génome d'*Arabidopsis*, ciblées efficacement par LFY. Ces régions ont été sélectionnées à partir des données génomiques de ChIP-seq de LFY dans les lignes de surexpression ainsi que des données de DNase-seq et de MNase-seq, qui ont été utiles pour analyser le paysage chromatinien (Zhang, Zhang, & Jiang, 2015; Zhang, Zhang, Wu, & Jiang, 2012). Une liaison forte mais non-spécifique de LFY de la gymnosperme *Ginkgo biloba* aux nucléosomes a été observée. Cependant, LFY d'*Arabidopsis thaliana* a montré une faible liaison aux nucléosomes.

Au chapitre II, l'objectif était de cartographier les domaines d'interaction minimale nécessaires de LFY et les ATPases SYD et BRM. En utilisant la technique HTRF, nous avons montré que le domaine C-terminal de LFY interagit avec le domaine HSA de BRM. De plus, grâce à une approche *in vivo*, nous avons observé la perte du phénotype 35S:LFY dans les plantes F1 de chacun des trois croisements: 35S:LFY *syd-5*, 35S:LFY *brm-1* et 35S:LFY *brm-3*. Cela suggère une interaction forte, ce qui signifie que lorsque BRM ou SYD ne sont pas fonctionnels, la fonction de LFY est affectée et aucune fleur ectopique n'est formée.