



Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

The University of Dublin, Trinity College



Specification of Floral Reproductive Organs: The Role of *AGAMOUS*

Diarmuid Seosamh Ó'Maoiléidigh

**A thesis submitted to
The University of Dublin
for the Degree of**

Doctor of Philosophy

**Supervised by Dr Frank Wellmer
Department of Genetics
Trinity College
University of Dublin
March, 2012**



Thesis 9751

Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university. Apart from the advice, assistance, and joints efforts mentioned in the acknowledgements and in the text, it is entirely my own work.

I agree to deposit this thesis in the University's open access institutional repository or allow the library to do so on my behalf, subject to Irish copyright legislation and Trinity College Library conditions of use and acknowledgement.

Signature of Author...

Two large, empty rectangular boxes with rounded corners, positioned side-by-side. The left box is for the author's signature, and the right box is for a stamp or seal.

Acknowledgements

I would like to thank Dr. Frank Wellmer for all of his guidance and tuition throughout the course of my Ph.D. I would especially like to thank Dr. Emmanuelle Graciet for all her patience in showing me techniques in the lab and for helpful discussions. The four years I have spent in the lab have been very enjoyable and I would like to thank you both for making the atmosphere in the lab so open and welcoming. Thanks to Brenda, Dave and Sue for helping to hold off the bureaucracy. Thanks to the Smurfit Institute, for hosting my research and to Science Foundation Ireland, for funding this project. Thanks to our collaborators in UCD, Dr. Amanda Lohan and Prof. Brendan Loftus and in Barcelona, Dr. Jiao Jin and Prof. Jose Luis Riechmann.

I would like to thank my current and former fellow lab members Andrea, Benoit, Chris, Fra, Kamila, Kasia and Liina for celebrating every success and commiserating every failure with me in the lab. I would especially like to thank Fra, Kamila, Kasia and Liina for their help in recent times, I wouldn't have gotten over the line without you. Last, but not least, I would like to thank Sam for keeping me in good spirits and for helping me to push myself to understand more everyday (nearly) for the last 4 years.

I would also like to thank all of my friends especially Áine and Maria for keeping me sane over the last year, we drink too much tea. I would like to thank Mick, Cathal, Mitch and Paul for helping me to forget about my stresses and have a good time, Aisling for listening to me moan and buying me coffee, Eimear for always cheering me up and Niall for always being there to talk to and have the craic. Thanks to Ronan for all the football he got me involved in, a major stress reliever (and sometimes stress maker), thanks to Ella for helping me to keep things in perspective and thanks to the other Ronan for all the craic we had while living together. Thanks to Kevin for all the pints and organising the walks, now I get to go outside again. A special thanks goes to Paddy who's always encouraged me and entertained me, in equal parts.

I would to thank Ciarán, Colin and Daibhid for paving the way for me so that I could have an easy ride and all their help during school, college and during my Ph.D. I'm lucky to have such kind brothers. And finally, I would like to thank my parents, John and Suzanne, for always believing in me and always encouraging me. I couldn't have asked for better parents, I love you both very much.

Summary

Homeosis is referred to as the transformation of an organ's identity toward the identity of another organ. This term was coined in 1894 and soon thereafter; these organ transformations were attributed to the variation in certain genes. These genes became known as "homeotic" genes and examples were highlighted in several model organisms including *Drosophila melanogaster* and *Arabidopsis thaliana*. In *Arabidopsis*, the first of these homeotic genes was cloned in 1990, and it was named *AGAMOUS* (*AG*), as plants without a functional copy of this gene do not develop reproductive organs.

The phenotypes of several other floral homeotic genes led to the proposal of the ABC model of flower development. This model states that the combinatorial action of three classes of homeotic genes, which encode transcription factors, direct the specification of the four types of floral organs, *i.e.* sepals, petals, stamens, and carpels. Although *AG* was cloned over two decades ago; its role in specifying reproductive organ identity has remained rather vague until recently. A number of genes have now been identified as being regulated by *AG*, however, genome-wide *in vivo* binding data of the *AG* protein and a developmentally resolved view of gene expression changes upon perturbation of *AG* function is still missing.

Therefore, the first and main focus of this Ph.D. thesis was to characterise the role of *AG* on a genome-wide scale. Using several genomics technologies and genetical experiments, I have shown that *AG* activates other transcription factors to promote developmental processes such as reproductive organ development and floral organ boundary maintenance. I also present evidence that *AG* ultimately regulates key factors that maintain and promote epigenetic modifications, a process that is crucial for sexual reproduction in plants. Furthermore, using an inducible perturbation system to knockdown *AG* function, I have shown that *AG* controls distinct processes in early and late morphogenesis.

Another focus of my thesis was to characterise the role of a downstream target of *AG*, *KNUCKLES*. This gene is involved in floral meristem determinacy; therefore, it may perform a distinct subset of *AG* functions. Alternatively, once activated, *KNU* might act in parallel with *AG* to terminate the floral meristem. This work has yielded some

valuable transgenic lines that will allow the molecular and phenotypic characterisation of *KNU*.

The study of the roles of *AG* and *KNU* in early flower development relies mainly on the floral induction system, which allows the collection of large amounts of synchronously developing flower buds at early stages of development. Collection of this type of material is next to impossible without use of the floral induction system due to the sequential initiation of flower buds in a developing inflorescence. *APETALA 1* (*API*) is important for the initiation of flower development; therefore, the floral induction system is based on the activation of the *APETALA 1* protein, which is overexpressed in this system. I improved the floral induction system by replacing the overexpression construct with the endogenous *API* promoter. In summary, the use of the floral induction system generated in this study is more appropriate for future genome-wide studies.

Abbreviations

(a)miRNA	(artificial) microRNA
6xHA	six times haemagglutinin tag
A	adenine
AGI code	Arabidopsis Genome Initiative code
<i>Agrobacterium</i>	<i>Agrobacterium Tumefaciens</i>
AlcA	an AlcR responsive promoter from <i>Aspergillus nidulans</i>
AlcR	EtOH responsive transcription factor from <i>A. nidulans</i>
<i>Arabidopsis</i>	<i>Arabidopsis thaliana</i>
ARF-	
Aux/IAA	auxin response factor
arg/R	arginine
BASTA	a commercial herbicide containing ammonium-glufosinate
BLAST	basic local alignment search tool
bp	base pair
C	cytosine
ca	carpel
CaMV 35S	Cauliflower Mosaic virus 35S promoter
cDNA	complementary DNA
ChIP	chromatin immunoprecipitation
ChIP-Seq	ChIP combined with ultrahigh-throughput sequencing
cm	centimeter
	Crossing-point, a value calculated by the Lightcycle 480 software (Roche)
Cp	
CPIC	Complete Protease Inhibitor Cocktail
d	days
dCAPs	derived cleaved amplified polymorphism
ddH2O	double distilled H2O
dH2O	deionized water
DIG	digoxigenin
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
<i>E. coli</i>	<i>Escherichia coli</i>
EAR-domain	ERF-associated amphiphilic repression domain
EDTA	ethylenediaminetetraacetic acid
EMSA	electrophoretic mobility shift assay
EtBr	ethidium bromide
EtOH	ethanol
FIS	floral induction system
G	guanine
GFP	green fluorescent protein
GO	gene ontology
GR	glucocorticoid
GRN	Gene Regulatory Network
GUS	β -glucuronidase
hr	hour
kb	kilobase
kDA	kilodalton
LB	Luria Bertrani
<i>L-er</i>	<i>Landsberg erecta</i>
	a chimeric transcriptional activator based on the lac repressor mutant
LhG4	lacHis,
LiCl	lithium chloride
lys/K	lysine
M	molar
	intensity ratio (M), average intensity (A) for a gene/probe in a
MA-plot	microarray study
mg, mL, mM	milligram, millilitre, millimolar
mg, mL, mM	microgram, microlitre, microgram

MgCl ₂	magnesium chloride
MgSO ₄	magnesium sulphate
min	minutes
mRNA	messenger RNA
MS	Murashige and Skoog medium
Na ₂ CO ₃	sodium carbonate
Na ₂ H ₂ PO ₄	sodium phosphate, dibasic
NaCl	sodium chloride
NaH ₂ PO ₄	sodium phosphate, dibasic
NaOH	sodium hydroxide
NASC	European Arabidopsis Stock Centre
NCBI	National Center for Biotechnology Information
NEB	New England Biolabs
ng, nm, nM	nanogram, nanometer, nanomolar
NOS	nopaline sythetase
°C	celcius
OP	optimized lac operator
ORF	open reading frame
PBS	Phosphate buffered saline
PCR	polymerase chain reaction
pe	petal
primer	oligonucleotide primer
qRT-PCR	quantitative real-time polymerase chain reaction
RNA	ribonucleic acid
RNAi	RNA interference
RNase	ribonuclease
rpm	revolutions per minute
RT-PCR	reverse transcriptase PCR
s.e.m.	standard error of the mean
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrilimide gel electrophoresis
se	sepal
Ser/S	serine
SRDX	an EAR repression domain
st	stamen
T	thymine
TAIR	The <i>Arabidopsis</i> Information Resource
TCd	Trinity College Dublin
T-DNA	transferred DNA of the tumour-inducing plasmid
TE	tris-EDTA buffer
thr	threonine
T _m	annealing temperature of primers in PCR reactions
Tris-HCl	tris buffer, pH calibrated with HCl
UCD	University College Dublin
UTR	untranslated region
V	volts
w/v	weight/volume
WT	wild-type

CHAPTER 1 – DISSECTING THE GENE REGULATORY NETWORK UNDERLYING FLOWER DEVELOPMENT: AN OVERVIEW.....	4
1.1 INTRODUCTION.....	4
1.1.1 <i>The initiation of flower development</i>	5
1.2 THE ABC MODEL OF FLORAL ORGAN IDENTITY SPECIFICATION	7
1.3 THE FLORAL HOMEOTIC GENE <i>AGAMOUS</i>	10
1.3.1 <i>Transcriptional control of stamen development by AG</i>	12
1.3.2 <i>Transcriptional control of gynoecium development by AG</i>	14
1.3.3 <i>Transcriptional control of floral meristem determinacy by AG</i>	16
1.4 GLOBAL IDENTIFICATION OF GENES DOWNSTREAM OF THE HOMEOTIC FACTORS	18
1.5 DISENTANGLING GENE REGULATORY NETWORKS	23
1.5.1 <i>The GRN of flower development</i>	24
1.5.2 <i>Strategies for gene network perturbation</i>	24
1.5.2.1 <i>Artificial microRNAs</i>	25
1.5.2.2 <i>RNA interference</i>	26
1.5.2.3 <i>Dominant negative repression</i>	27
1.5.2.4 <i>Null mutant alleles</i>	28
1.5.3 <i>Chromatin immunoprecipitation</i>	28
1.6 PROJECT OUTLINE	30
1.6.1 <i>Perturbation analysis</i>	32
1.6.2 <i>Identification of direct AG targets by ChIP-Seq</i>	33
1.6.3 <i>Identification of molecular processes regulated by AG</i>	34
CHAPTER 2 – MATERIALS AND METHODS.....	35
2.1 STRAINS AND GROWTH CONDITIONS	35
2.1.1 <i>Plant Lines</i>	35
2.1.2 <i>Bacterial Strains</i>	35
2.1.3 <i>Plant Growth Media</i>	36
2.1.4 <i>Bacterial Growth Media</i>	36
2.2 CLONING AND GENOTYPING	36
2.2.1 <i>Plasmid DNA preparation</i>	36
2.2.2 <i>Genomic DNA extraction using the CTAB method</i>	37
2.2.3 <i>cDNA synthesis</i>	37
2.2.4 <i>Preparation and transformation of competent Escherichia coli</i>	37
2.2.5 <i>Preparation and transformation of competent Agrobacterium tumefaciens</i>	38
2.2.6 <i>Quick Arabidopsis genomic DNA preparation</i>	39
2.3 PLANT TRANSFORMATION AND SELECTION OF TRANSFORMANTS	39
2.3.1 <i>Plant transformation and seed sterilization</i>	39
2.3.2 <i>Selection of transgenic plants</i>	39
2.4 MICROSCOPY TECHNIQUES	40
2.4.1 <i>Light microscopy</i>	40
2.4.2 <i>Epifluorescence microscopy</i>	40
2.4.3 <i>Confocal imaging</i>	40
2.5 TREATMENTS OF LINES ENCODING INDUCIBLE PROMOTERS	40
2.5.1 <i>Ethanol treatment of plants</i>	40
2.5.2 <i>Treatment of plants with a dexamethasone-containing solution</i>	41
2.6 GENERATION OF AMIRNAS.....	41
2.7 QUANTITATIVE REAL-TIME PCR.....	42
2.8 THERMAL ASYMMETRIC INTERLACED PCR	42
2.9 WESTERN BLOTTING	43
2.9.1 <i>Preparation of nuclear extracts for SDS-PAGE electrophoresis</i>	43
2.9.2 <i>Transfer of proteins from an SDS-PAGE gel to a membrane</i>	43
2.10 CHROMATIN IMMUNOPRECIPITATION	44
2.10.1 <i>Fixation and nuclei extraction</i>	44
2.10.2 <i>Sonication</i>	45
2.10.3 <i>Immunoprecipitation</i>	45
2.10.4 <i>Purification of ChIP DNA</i>	46
2.10.5 <i>Quantification of genomic DNA</i>	47

2.10.6 Illumina ChIP-Seq library preparation.....	47
2.11 <i>IN SITU</i> HYBRIDISATION	48
2.12 DATA ANALYSIS AND BIOINFORMATICS	48
2.12.1 Re-annotation of the Agilent 44K array to the latest genome release.....	49
2.12.2 Low-level microarray data processing and basic data handling	50
2.12.3 Higher level microarray analysis	51
2.12.4 Short read alignment, quality control and general data handling for ChIP-Seq	52
2.12.5 Peak calling.....	52
2.12.6 Motif Detection.....	53
2.13 TRANSGENIC LINE GENERATION.....	53
2.13.1 Generation and expression of an amiRNA that targets the AG mRNA	53
2.13.2 Generation of an RNAi construct to disrupt AG function.....	54
2.13.3 Generation of dominant negative versions of AG.....	54
2.13.4 Generation of lines carrying the pAG::AG-6xHA rescuing construct	55
2.13.5 Generation of lines carrying the pAG::AG-mGFP5 rescuing construct.....	56
2.13.6 Crossing pAG::AG-mGFP5 ag-1 lines into the FIS.....	57
2.13.7 Introduction of the pAG::AG-mGFP5 fusion into an alternative FIS.....	58
2.13.8 Isolation of pAlcA::AG4-amiRNA lines in the pAPI-FIS.....	58
2.13.9 Generation and expression of an amiRNA that targets the KNU mRNA	59
2.13.10 Over-expression of the KNU coding sequence	59
2.13.11 Generation of an epitope-tagged version of KNU resistant to the KNU-amiRNA.....	60
2.13.12 Generation of a kanamycin-resistant 35S::API-GR floral induction system	61
2.13.13 A floral induction system based on the genomic locus of API	61
2.13.14 A floral induction system based on the over-expression of CAL	62
2.13.15 A floral induction system based on the genomic locus of CAL.....	62
2.13.16 Identification of the T-DNA insertion sites in plants harbouring the pAPI::API-GR transgene.....	63
2.13.17 Generation of plants expressing the AG4-amiRNA in specific developmental domains	63
2.13.18 Generation of plants expressing the KNU-amiRNA in specific developmental domains.....	64
CHAPTER 3 – CHARACTERIZATION OF THE TRANSCRIPTIONAL NETWORK CONTROLLED BY AGAMOUS.....	65
3.1. INTRODUCTION.....	65
3.2 RESULTS	65
3.2.1 Perturbation of AG function.....	65
3.2.1.1 Identifying an amiRNA that efficiently targets AG.....	66
3.2.1.2 Perturbation of AG activity using RNA interference	68
3.2.1.3 Perturbation of AG function by fusions with dominant repressor domains	71
3.2.2 Generation and characterization of a transgenic line that allows an inducible perturbation of AG activity.....	74
3.2.2.1 Generation of a 35S::AlcR/pAlcA::AG4-amiRNA line	74
3.2.2.2 Effects of AG4-amiRNA expression in wild-type plants	75
3.2.2.3 Molecular characterization of the AG4-amiRNA expression in a floral induction system	82
3.2.3 Generation of AG antibodies and of lines expressing epitope-tagged AG.....	83
3.2.3.1 Translational fusion of a haemagglutinin (HA) epitope tag to AG.....	84
3.2.3.3 Generation of a floral induction system expressing AG-mGFP5	87
3.2.3.4 Generation of AG peptide antibodies.....	89
3.2.4 Identification of genes acting downstream of AG through expression profiling of ag mutant flowers	90
3.2.4.1 Experimental design and implementation	91
3.2.4.3 Interplay between AG and B function genes.....	101
3.2.5 Identification of genes acting downstream of AG through inducible gene perturbation.....	106
3.2.5.1 Experimental design and implementation	106
3.2.5.2 Characterization of AG activity using the inducible AG4-amiRNA perturbation system	111
3.2.6 Establishment of a ChIP protocol.....	115
3.2.6.1 Optimisation of ChIP Protocol 1	117
3.2.6.2 Optimisation of ChIP protocol 2.....	123
3.2.7 Global identification of genomic regions bound by AGAMOUS.....	124
3.2.7.1 Generation of an Illumina ChIP-Seq library from ChIP enriched material	124
3.2.7.2 Quality assessment of the ChIP-Seq data.....	126
3.2.7.3 Detection of AG binding sites in the <i>Arabidopsis</i> genome.....	129

3.2.8 Discussion.....	136
3.2.8.1 Transgenic line generation, experimental setup and implementation	136
The efficiency of RNA interference-mediated gene silencing	140
The over-expression of a dominant negative version of AG	141
Complementing the <i>ag-1</i> phenotype with an epitope-tagged version of AG	141
The activated AlcR transcription factor can influence the expression of endogenous <i>Arabidopsis</i> genes	143
Global expression profiling using the <i>ag-1</i> mutant resulted in relatively small fold-change differences of downstream genes.....	146
3.2.8.2 The role of <i>AGAMOUS</i> in flower development.....	146
AG promotes the expression of other transcription factors involved in reproductive organ development	147
AG directly regulates genes involved in floral organ boundary formation	150
AG is involved in the epigenetic regulation of flower development	152
AG controls distinct developmental processes at early and late stages of flower development	154
3.2.9 Future Work	156
CHAPTER 4 – INVESTIGATING THE ROLE OF KNUCKLES IN FLOWER DEVELOPMENT... 158	
4.1 INTRODUCTION.....	158
4.2 RESULTS	159
4.2.1 Identification of an amiRNA that efficiently perturbs KNUCKLES activity	159
4.2.2 Identification of a transgenic line that allows an inducible perturbation of KNU activity.....	164
4.2.3 Generation of KNU over-expressing plants	166
4.2.4 Translational fusion of an mGFP5 reporter gene to KNU.....	167
4.3 DISCUSSION.....	168
4.3.1 Identification of a functional KNU-amiRNA	168
4.3.2 Functionality of mKNU-GFP fusion protein.....	170
4.3.3 Inducible or constitutive over-expression of KNU did not result in novel phenotypes	171
4.4 FUTURE WORK	171
CHAPTER 5 - GENERATION OF ALTERNATIVE FLORAL INDUCTION SYSTEMS 173	
5.1 INTRODUCTION.....	173
5.2 RESULTS	175
5.2.1 Generation of a floral induction system with an alternative selection marker.....	175
5.2.2 Generation of an improved floral induction system driving <i>API-GR</i> under the control of the endogenous <i>API</i> promoter.....	175
5.3 GENERATION OF FLORAL INDUCTION SYSTEMS BASED ON INDUCIBLE ACTIVATION OF <i>CAL</i>	183
5.4 DISCUSSION	185
5.4.1 Floral induction systems based on the activation of the <i>API-GR</i> protein	185
5.4.2 Floral induction systems based on the activation of a <i>CAL-GR</i> fusion protein.....	187
5.5 FUTURE WORK	188
CHAPTER 6 – DOMAIN-SPECIFIC PERTURBATION OF AG AND KNU FUNCTION THROUGH AMIRNAS 190	
6.1 INTRODUCTION.....	190
6.2 RESULTS	192
6.2.1 Expression of the <i>AG4-amiRNA</i> in the <i>CLV1</i> expression domain	192
6.2.2 Expression of the <i>AG4-amiRNA</i> in the <i>CLV3</i> expression domain.....	194
6.2.3 Expression of the <i>AG4-amiRNA</i> in the <i>WUS</i> expression domain	197
6.2.4 Expression of the <i>KNU-amiRNA</i> in the <i>CLV1</i> , <i>CLV3</i> and <i>WUS</i> expression domains	199
6.3 DISCUSSION.....	201
6.3.1 General considerations on the expression of amiRNAs in specific domains	201
6.3.2 Expression of the <i>AG4-amiRNA</i> in the <i>CLV1</i> , <i>CLV3</i> and <i>WUS</i> mRNA domains.....	203
6.3.3 Expression of the <i>KNU-amiRNA</i> in the <i>CLV1</i> , <i>CLV3</i> and <i>WUS</i> mRNA domains	206
6.4 FUTURE WORK	208
CHAPTER 7 – GENERAL DISCUSSION AND FUTURE PERSPECTIVES..... 209	
CHAPTER 8 - REFERENCES 212	

Chapter 1 – Dissecting the Gene Regulatory Network Underlying Flower Development: An Overview

1.1 Introduction

The generation of specific cell types from undifferentiated stem cells is largely dependent upon differential gene expression. The developmental program of an organism is encoded in its genome in the form of *cis*-regulatory elements that determine when and where genes are expressed (Wellmer and Riechmann, 2005). Transcription factors interact with these sequences in order to activate or repress transcription. Transcription factors are themselves subject to control by other regulators and together they form a complex regulatory network (Materna and Oliveri, 2008). Gene regulatory networks are difficult to dissect, however, they are fundamental to developmental processes such as embryogenesis and organogenesis. Flower development has served as a model system over the past 20 years to investigate the molecular basis of organogenesis in plants (Lohmann and Weigel, 2002).

One of the milestones in the field of flower development was the conception of the ABC model of floral organ identity specification (Bowman et al., 1991b; Coen and Meyerowitz, 1991). This model describes how the combinatorial action of three classes of homeotic genes, which encode transcription factors, direct the specification of the four types of floral organs, *i.e.* sepals, petals, stamens, and carpels. While ample information is available regarding the steps that precede the activation of the floral homeotic genes, very little is known about which genes their cognate proteins regulate (Ausin et al., 2005; Lohmann and Weigel, 2002; Parcy, 2005). In spite of concentrated efforts, relatively few genes have been identified that are directly regulated by the floral homeotic transcription factors (Ito et al., 2004; Sablowski, 2009; Sablowski and Meyerowitz, 1998). Forward genetic screens have identified many genes that are involved in flower development, however, the functional redundancy that appears to be present among genes involved in early flower development requires alternative approaches such as the genome-wide analysis of gene expression by microarray analysis (Wellmer et al., 2006). This approach allows the generation of high-resolution stage-specific datasets to reveal sets of genes whose expression needs to be activated or repressed in order for a specific developmental program

to take place. Moreover, these datasets will identify genes that are involved in the same regulatory networks. Expression and phylogenetic analysis can then be used in order to infer functional redundancy. Finally, null alleles for functionally redundant genes can be generated in the same plant in order to reveal their functions.

In the paragraphs that follow, I will provide an overview of the molecular basis of flower development and outline the current state of the field. I will discuss the original developmental models that were proposed and their subsequent extensions. Because the focus of this Ph.D. thesis is the functional characterization of the floral homeotic gene *AGAMOUS* (*AG*), I will review topics relating to its role in floral organ identity specification in detail. In addition, I will discuss methods that can be used for the identification of components downstream of transcription factors in *Arabidopsis* and outline a strategy to identify the stage-specific targets of *AG* on a genome-wide scale. I will also describe transgenic lines generated in this study that perturb the activity of a putative direct target of *AG*, *KNUCKLES* (*KNU*) (Chapter 4). To complement these projects, an approach that aims to perturb *AG* and *KNU* function in specific domains of the developing flower will be presented (Chapter 6). This approach will allow me to identify the regions of the developing flower where *AG* and *KNU* activities are necessary to direct the gene regulatory networks under their control. Finally, an improved system that allows the collection of large amounts of early-stage flower buds will be described (Chapter 5).

1.1.1 The initiation of flower development

The life cycle of the cruciferous plant *Arabidopsis thaliana* can be divided into two main stages of development, the vegetative and the reproductive phase. The reproductive phase itself can be further subdivided into three distinct phases that occur in a temporal sequence. First, the plant switches from a vegetative to a reproductive development (Figure 1.1A, B). This transition is characterized by the reprogramming of the shoot apical meristem (SAM) into an inflorescence meristem (IM). The key phenotypic difference between these meristems is that SAMs generate rosette leaves and axillary meristems (Figure 1.1A), whereas IMs generate a stem and cauline leaves before establishing floral

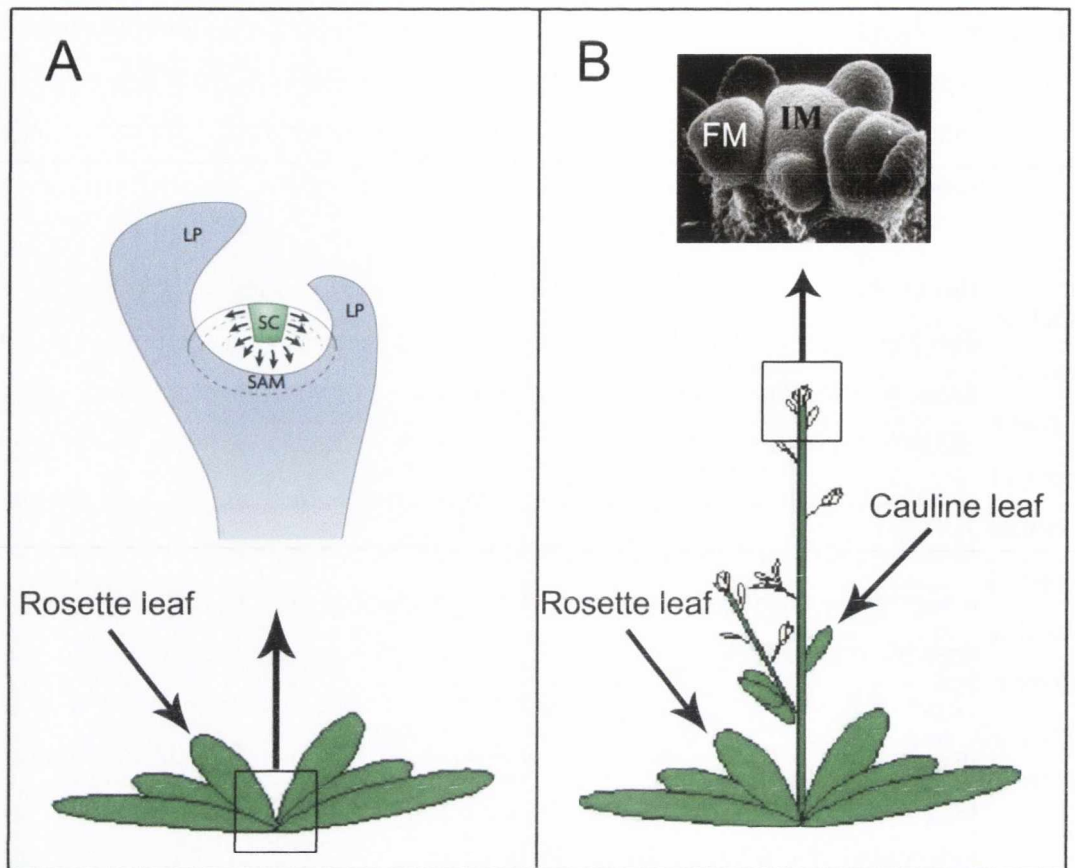


Figure 1.1 Vegetative and reproductive *Arabidopsis* development. (A) Schematic of a shoot apical meristem (SAM) producing leaf primordia (LP). A pool of stem cells (SC) is maintained within the shoot meristem. (B) A plant that has transitioned to flowering producing a stem, cauline leaves and flowers. Environmental signals have induced the specification of an inflorescence meristem (IM) from the vegetative shoot meristem. The inflorescence meristem produces floral meristems (FM) at its lateral edges and eventually mature flowers form. Figure modified from <http://biology.kenyon.edu> and (De Veylder et al., 2007).

meristems (FM) on their flanks to generate flowers (Figure 1.1B). Genes known as flowering time genes control this process through several distinct pathways (for review see (Ausin et al., 2005)). The flowering time genes, which are regulated by cues such as light and temperature, ultimately activate a group of genes known as the floral meristem identity genes (for review see (Parcy, 2005)). Plants that carry deleterious homozygous mutations in the floral meristem identity genes produce aberrantly patterned flowers, and floral organ identity is affected or abolished (Kempin et al., 1995; Schultz and Haughn, 1991; Weigel et al., 1992). These mutant phenotypes hinted that the transcription factors encoded by the floral meristem identity genes activate the floral homeotic genes in specific regions of the emerging floral bud (for a review see (Liu et al., 2009)). Once the floral homeotic genes are

activated, the corresponding transcription factors direct the development of the immature flower by activating, directly or indirectly, thousands of other genes which ultimately specify the cell types and tissues of the mature flower (Jack, 2004). The manner in which these homeotic transcription factors direct floral organ formation can be schematically explained by the ABC model of floral organ identity specification.

1.2 The ABC model of floral organ identity specification

The *Arabidopsis* flower is composed of four organ types, which are arranged in concentric rings, or whorls (Figure 1.2A). Four green leaf-like organs called sepals occupy the outermost whorl (the first whorl), the second whorl contains 4 white petals, the third whorl is comprised of six stamens that eventually produce pollen, while the fourth whorl includes two fused carpels, which encompass the ovules (Smyth et al., 1990). Genetic screens have led to the identification of a large number of mutants with abnormal flowers. Arguably, the most striking of these mutants are the floral homeotic mutants where individual floral organs develop normally but in an inappropriate whorl at the expense of another organ type (Figure 1.2B - G). Each mutant phenotype can be classified by the homeotic transformations observed. "A class" mutants include *apetala1* (*ap1*) (Figure 1.2B) and *apetala2* (*ap2*) (Figure 1.2C). *ap2* plants generate flowers where sepals and petals are transformed into carpels and stamens, respectively (Bowman et al., 1989) (Figure 1.2C), although second whorl organs are often absent. *ap1* flowers (Figure 1.2B) often lack petals and in place of sepals, bract-like organs form (Irish and Sussex, 1990). "B class" mutants include *apetala3* (*ap3*) and *pistillata* (*pi*) (Bowman et al., 1989). The flowers of either mutant have sepals in the second whorl instead of petals, and carpels or filamentous organs in the third whorl instead of stamens (Figure 1.2D, E). The "C class" mutant *ag* has a dual phenotype. The flowers of *ag* plants have no stamens or carpels but extra whorls of sepals and petals are present in an indeterminate manner (see Figure 1.2F above). The organs that develop within an *agamous* flower can be represented as: (sepal, petal, petal)_n (Bowman et al., 1989). Given these mutant phenotypes, it was proposed that the formation of the four floral organ types is developmentally controlled by the overlapping activities of the genes affected in these mutants and that each gene is active in two adjacent whorls (Bowman et al., 1991b; Coen and Meyerowitz, 1991) (Figure 1.2H). It was also suggested that A class

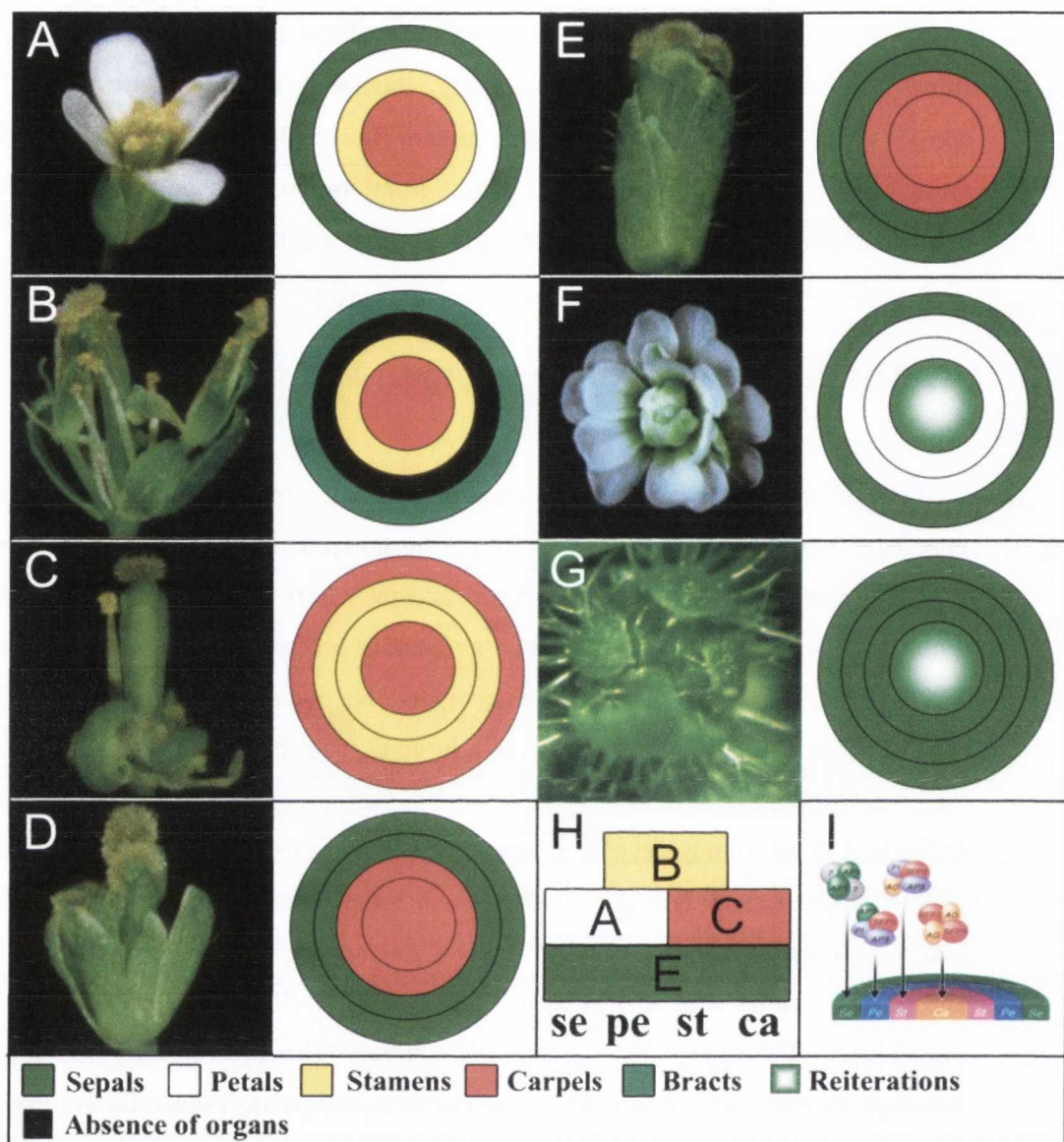


Figure 1.2 Phenotypes of floral homeotic mutants and graphic representation of the organ transformations that occur in each whorl. (A) A wild-type flower containing 4 sepals, 4 petals, 6 stamens and 2 fused carpels. (B) An *ap1-1* flower. (C) An *ap2-2* flower. (D) A *pi-1* flower. (E) An *ap3-3* flower. (F) An *ag-1* flower. (G) A quadruple *sep1-4* mutant flower. (H) A graphical representation of the ABC model. The *SEP* genes, which function as E class genes, are active in all whorls of the developing flower. (I) The floral homeotic transcription factors form higher-order protein complexes in order to activate or repress their target genes. Se: sepals; pe: petals; st: stamens; ca: carpels. Figure modified from (Krizek and Fletcher, 2005; Lohmann and Weigel, 2002; Wellmer et al., 2004).

and C class gene activities are mutually exclusive since the corresponding mutants show phenotypes that are essentially mirror images of each other.

According to the ABC model, *AP1* and *AP2* are required in the first whorl to direct sepal fate correctly. *AP1*, *AP2*, *AP3* and *PI* are required for normal petal development in the second whorl. *AP3*, *PI* and *AG* are required for proper stamen development, whereas *AG* is the only homeotic gene that directs carpel development. All of the homeotic genes, except *AP2*, encode MADS domain-containing transcription factors and are expressed in the domains in which they are active. Although *AP2* was initially thought to be expressed uniformly throughout the developing flower (Jofuku et al., 1994), it was recently shown that *AP2* mRNA is largely confined to the two outer whorls during flower development (Wollmann et al., 2010), as expected for an A class gene.

Plants in which all three classes of the homeotic genes are mutated contain leaf-like organs (Bowman et al., 1991b), however, over-expressing the homeotic genes in leaves failed to convert them into floral organs (Krizek and Meyerowitz, 1996; Mizukami and Ma, 1992). This suggested that there are additional factors required to specify floral organ fate. Subsequently, it was discovered that the floral homeotic transcription factors require the MADS domain-containing SEPALLATA (*SEP*) proteins as transcriptional cofactors (Pelaz et al 2000). There are four *SEP* genes (*SEP1-4*) in *Arabidopsis* and plants mutant for all of them have flowers that are entirely comprised of leaf-like structures (Figure 1.2G) (Ditta et al., 2004). It has been shown that a simultaneous over-expression of the homeotic genes with the *SEP* genes is sufficient to direct floral organ identity (Honma and Goto, 2001; Pelaz et al., 2001).

Another important feature of the current model of flower development is that these homeotic factors act in higher-order complexes with the *SEP* proteins in order to influence transcription (Figure 1.2I). The role of the *SEP* proteins in these complexes may be two-fold: the *SEP* proteins may provide activation domains to those MADS domain proteins that cannot activate transcription alone and/or the higher order complexes may increase DNA binding affinity in comparison to lower order complexes (Honma and Goto, 2001). It has also been shown that *SEP3* binding sites are enriched for predicted binding sites of other transcription factor families, apart from the MADS domain family (Kaufmann et al., 2009). This raises the possibility that the floral homeotic transcription factors act in complexes with other families of transcriptional regulators in order to control development. However, no experimental evidence is available for this conjecture at present. It has been shown that *AG* interacts with *AP3*, *PI* and a *SEP* protein in order to direct stamen development, and an *AG* protein dimer appears to interact with a *SEP* protein dimer to direct carpel fate

(Castillejo et al., 2005; Honma and Goto, 2001), although a final verification of this *in planta* is lacking. Moreover, it has been shown that simultaneous over-expression of *AP3*, *PI*, *AP1* and *SEP3* is sufficient to transform leaves into petals (Honma and Goto, 2001), indicating that these complexes influence the transcription of their target genes to generate reproductive organs.

The identification of protein complexes that are required for organ specification alludes to a general mechanism of how these transcription factors differentially regulate specific genes. Perhaps these MADS-box protein complexes interact with other proteins in order to direct the expression of specific gene sets. This conjecture is supported by the observation that *AP1* and *SEP3* interact with the transcriptional co-repressor proteins *LEUNIG* and *SEUSS* in order to repress target genes such as *AGAMOUS* (Gregis et al., 2006; Sridhar et al., 2006). Below, the progress that has been made of how *AG* influences floral organ identity specification and meristem determinacy is described.

1.3 The floral homeotic gene *AGAMOUS*

The C class gene *AGAMOUS* was cloned in the early 1990s (Yanofsky et al., 1990). It encodes a protein that is similar to the transcription factors *MCM1*, *DEFICIENS* and Serum Response Factor (SRF), which are present in yeast, *Antirrhinum majus* and humans, respectively. These proteins are the founders of a large family of proteins present in eukaryotes that play fundamental roles in development and signal transduction, and are known as the MADS domain transcription factor family (for review see (Messenguy and Dubois, 2003)). The MADS domain is involved in protein dimerization and DNA binding (Schwarz-Sommer et al., 1990).

AG is a ~28 kD type II MADS domain protein (Alvarez-Buylla et al., 2000). Type II MADS domain proteins contain three other protein domains, termed the I, K and C domains. The K domain has been shown to be important for protein-protein interactions and it is linked to the C domain, which has been suggested to act as a trans-activation domain, by the weakly conserved I domain (Riechmann and Meyerowitz, 1997). *AG* is thought to participate in large protein complexes to control transcription as mentioned above. It appears to interact with *PI*, *AP3* and a *SEP* protein to specify stamens, whereas the

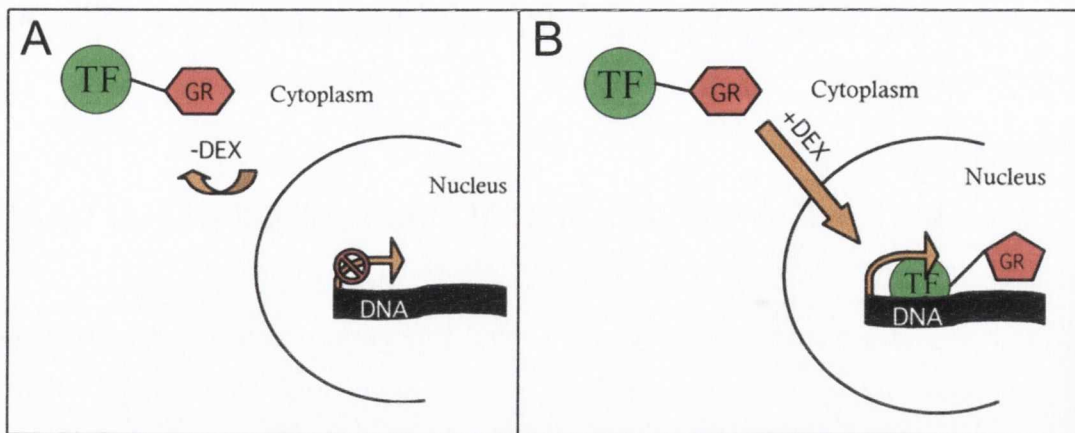


Figure 1.3 A dexamethasone-inducible system for the specific activation of plant transcription factors.

The system is based on a fusion between a transcription factor (TF) and the hormone-binding domain of the rat glucocorticoid receptor (GR). **(A)** The GR portion of the fusion protein excludes the transcription factor of interest from the nucleus in the absence of dexamethasone (DEX), rendering it inactive. **(B)** Once dexamethasone is added, the fusion protein is released from its cytoplasmic retention and is imported into the nucleus. Dexamethasone binds directly to the GR protein fragment to release it from its cytoplasmic retention (Mangelsdorf et al., 1995).

development of carpels is likely to be mediated by a complex including an AG protein dimer and a SEP protein dimer (Ditta et al., 2004; Honma and Goto, 2001). The flowers of plants lacking a functional copy of AG are indeterminate and have no stamens and carpels (Figure 1.2F). This phenotype has been shown to correlate with an expansion of A class function into the third and fourth whorls in the absence of AG protein (Bowman et al., 1991b). In concert with the B class genes, the A class genes produce petals in the third whorl and nested flowers in the fourth whorl of *ag-1* flowers that have a sepal-petal-petal pattern of reiteration (Bowman et al. 1989).

The phenotype of *ag-1* flowers indicates that *AG* is a master regulator of reproductive organ development. It is possible for AG to perform this function in several ways: AG could activate other transcription factors and signalling pathways very early in development, thereby passing on its functions to other genes that ultimately specify the appropriate floral organs, or it could directly activate genes that are involved in basic cellular processes such as cell division and metabolism. Alternatively, AG could perform both of these functions simultaneously or its activity may change depending on the stage of flower development. Some insights have been given into the molecular mechanisms underlying AG function by examining its expression pattern and identifying some of its downstream targets, however, it

remains largely unclear how *AG* specifies organ identity over the course of flower development.

1.3.1 Transcriptional control of stamen development by *AG*

AG expression commences at stage 3 of flower development (developmental stages according to Smyth et al., 1990) in the centre of the floral meristem. *AG* mRNA is found in stamen primordia from stage 5 until stage 14 when an RNA hybridization signal is seen in specific tissues of the stamen, such as the endothecium of the anther wall and nectaries. However, *AG* expression is absent from the microspore cell lineage from the time meiosis occurs (Bowman et al., 1991a; Drews et al., 1991). The expression pattern of *AG* mRNA in stamens suggests that *AG* functions not only in early development, but also in stamen maturation. Moreover, the expression pattern of *AG* during late stamen development is much more restricted than its expression pattern in early development (Drews et al., 1991). This suggests that *AG* may have distinct roles, and target genes, during late and early flower formation.

Stamen development is initiated by the activation of the B and C class genes, and by cofactors such as *UNUSUAL FLORAL ORGANS* (*UFO*) or the *SEP* genes, in the third whorl. Although hundreds of stamen-specific genes appear to act downstream of the B- and C-class transcription factors, evidence is available for the direct regulation of only a handful of genes by these homeotic factors (Wellmer et al., 2004). One of these genes, *SPOROCTELESS* (*SPL*; also known as *NOZZLE*), was identified in a screen of *Ds* transposon insertion lines by its complete male and female sterility (Sundaresan et al., 1995). Subsequently it was revealed that *SPL* is a transcription factor required for microsporogenesis and megasporogenesis (Schiefthaler et al., 1999; Yang et al., 1999). Using a transgenic line in which *AG* can be specifically activated (Figure 1.3 explains the inducible system that was used), Ito et al. (2004) showed that *AG* is sufficient to induce microsporogenesis. An expression profiling study with this line was then performed and it was demonstrated that *SPL* expression is responsive to *AG* activation. Next, using an electro-mobility shift assay it was shown that *AG* can bind the putative binding motif for MADS domain transcription factors, known as a CARG-box, that is present in the 3' region of the *SPL* gene. Moreover, a transcriptional fusion between β -glucuronidase (*GUS*) and a

putative promoter fragment of the *SPL* gene containing a mutated version of this CArG-box sequence disrupted the GUS activity that was directed by the non-mutated version of this transcriptional fusion. Finally, a steroid inducible version of *SPL*, when activated, could induce microsporogenesis (Ito et al., 2004). These results together indicated that *SPL* is directly activated by AG and acts as a mediator for AG in microsporogenesis. This was the first indication that AG controls organogenesis as part of a transcriptional regulatory hierarchy as opposed to directly regulating genes involved in basic cellular processes.

AG also appears to influence hormone signalling in order to direct stamen development. Jasmonate has been found to be important for the maturation and release of pollen while gibberellic acid (GA) has been shown to be important for microsporogenesis and stamen elongation (Cheng et al., 2004; Stintzi and Browse, 2000). Yu et al. (2004) demonstrated that the over-expression of the protein Repressor of GA (RGA) results in the down-regulation of the floral homeotic genes, including *AG* (Yu et al., 2004). RGA is a member of the DELLA protein family, which contains proteins known to be developmental repressors involved in the signalling pathway of the phytohormone gibberellic acid (GA) (Silverstone et al., 2001). Therefore, GA signalling appears to be a positive regulator of *AG* expression. Recently, it was suggested that AG itself directly and positively regulates a gene that is involved in GA biosynthesis, *GA4* (Gomez-Mena et al., 2005). It is known that GA upregulates genes involved in jasmonic acid synthesis (Cheng et al., 2009) and one of these genes, *DADI*, has been suggested to be a direct target of AG (Cheng et al., 2009; Gomez-Mena et al., 2005; Ito et al., 2007). Ultimately, it appears that jasmonic acid (JA) upregulates transcription factors that control late stamen development (Cheng et al., 2009; Mandaokar et al., 2006). It is known that other hormones, apart from JA and GA, are also important for the regulation of stamen development. However, it has not been shown that AG controls these processes directly (Cecchetti et al., 2008; Cheng et al., 2006; Kaufmann et al., 2009). The genes that have been identified support the notion of *AG* acting as master transcriptional regulator in early and late stamen development. The regulation of *SPL* demonstrates that AG passes its function to another transcription factor in a region of the stamen where *AG* is not expressed, thus indirectly controlling sporogenesis. In addition, the complex hierarchy observed for the regulation of hormone signalling by AG demonstrates how AG controls a myriad of other genes that ultimately activate genes involved in tissue specification. These studies offer valuable insights into the regulation of stamen development by AG; however, they are too few to build a model of AG activity

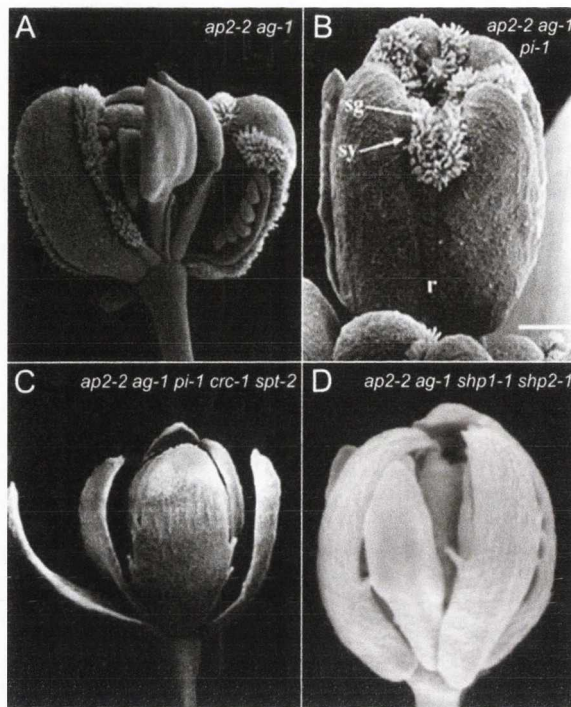


Figure 1.4 The *CRC*, *SPT* and *SHP* genes direct carpel fate in parallel with *AG*. (A) An *ag-1 ap2-2* double-mutant flower. Carpelloid tissue is present in the outer whorls of these flowers suggesting that genes other than *AG* may also specify carpel fate. (B) An *ag-1 pi-1 ap2-2* triple-mutant flower. The *pi-1* allele was crossed with *ag-1 ap2-2* in order to remove lateral patches of stamen tissue that arises in the first whorl of *ag-1 ap2-2* mutants. (C) An *ag-1 pi-1 ap2-2 crc-1 spt-2* quintuple mutant. No carpelloid tissue is present in these flowers. The organs present are more ovate and leaf-like, however, they retain some sepallod properties. (D) An *ag-1 ap2-2 shp1-1 shp2-1* flower. These flowers are very similar to the quintuple mutant shown in panel C. Figure adapted from (Alvarez and Smyth, 1999; Pinyopich et al., 2003).

during third whorl organ development.

1.3.2 Transcriptional control of gynoecium development by *AG*

As previously mentioned, *AG* expression is quite uniform in the centre of the floral meristem until around stage 6. At stage 9, *AG* expression is detected in ovule primordia and in fully differentiated carpel cells such as stigmatic papillae and the ovular endothelium at later stages (Bowman et al., 1991a). In the flowers of wild-type plants, *AG* expression is restricted to the central two whorls by the product of the A class gene *AP2*. In a strong *ap2* mutant flower, *AG* is expressed in all four floral organs and the outer whorl is composed of

carpels as opposed to sepals due to the action of AG in specifying the tissue in the absence of the AP2 protein (Figure 1.2C) (see Section 1.2) (Bowman et al., 1989; Drews et al., 1991). In flowers of *ap2 ag* double mutants, however, the first whorl retains some features unique to the fourth whorl such as stigmatic tissue and medial first whorl carpelloid leaves (Figure 1.4A) (Bowman et al., 1991b). This observation suggests that AG function can be at least partially replaced by other genes.

In 1999, two groups identified genes that appear to control carpel development in parallel with AG (Alvarez and Smyth, 1999; Bowman and Smyth, 1999). These genes, *CRABS CLAW* (*CRC*) and *SPATULA* (*SPT*), code for zinc-finger and basic helix-loop-helix (bHLH) transcription factors, respectively (Bowman and Smyth, 1999; Heisler et al., 2001). The carpels of *crc* mutant plants are shorter and wider, they are often unfused and the number of ovaries is reduced when compared to the wild type. The carpels of *spt* plants lack a transmitting tract, which is required for pollen tube growth. In addition, carpels are often unfused and the number of ovules present is reduced compared to wild-type plants. Apart from analysing the phenotypes and expression patterns of these genes, genetic interactions between *spt* and *crc* and some of the floral homeotic genes was investigated. When an *AP2* mutation was introduced into an *ag pi* double mutant background, the resulting plants produced flowers that contained carpelloid organs in every whorl except lateral first whorl organs. The addition of *spt-2* and *crc-1* alleles, to generate a quintuple mutant, abolished nearly all of these carpelloid features. The organs found within the flowers of these plants appear to be predominantly leaf-like (Figure 1.4C) (Alvarez and Smyth, 1999).

From analyzing mRNA expression patterns by *in situ* hybridization it was shown that both *SPT* and *CRC* are expressed independently of AG, however, it seems that AG is required to maintain or support their expression (Bowman and Smyth, 1999; Heisler et al., 2001). Recently, it was shown by electrophoretic mobility shift assay (EMSA) and chromatin immunoprecipitation (ChIP), that AG can bind the upstream promoter region of *CRC* (Gomez-Mena et al., 2005). However, no such data is available for *SPT*. In addition, it was shown that *AP2* negatively regulates both genes (Bowman and Smyth, 1999; Heisler et al., 2001). These analyses show that AG may regulate carpel development in parallel with *CRC* and *SPT*, and that the activation of these genes is independent of AG, although AG appears to at least boost their expression levels.

Another set of genes that is important for carpel development is comprised of the closely related *SHATTERPROOF1* (*SHP1*) and *SHP2* genes. When mutations in the *SHP*

genes are introduced into an *ag ap2* background (see above), flowers no longer have any carpel identity (Figure 1.4D). From the phenotypes, *SHP1/2* appear to be active in ectopic carpelloid tissue of *ap2 ag* flowers indicating that their expression can be independent of *AG* (Pinyopich et al., 2003; Savidge et al., 1995). It has been proposed that the proteins encoded by these two genes form higher order complexes with *SEP3* and *AG* in order to promote carpel development, which would support the idea that different combinations of these MADS domain proteins may be able to control different aspects of development in the fourth whorl (Favaro et al., 2003; Pinyopich et al., 2003).

The interactions between *AG*, *SPT*, *CRC* and the *SHP* genes are somewhat unclear at present. The expression levels of *SPT* and *CRC* are reduced in *ap2 ag* mutants in comparison to the wild type, but no data is available for expression levels of the *SHP* genes in *ap2 ag* mutants (Bowman and Smyth, 1999; Heisler et al., 2001). *AG* has been shown to bind both the *CRC* promoter and the *SHP2* promoter, however, the activity of both persist in the *ag ap2* background (Gomez-Mena et al., 2005; Savidge et al., 1995). Furthermore, it has been shown that *CRC* and *SPT* enhance the expression of *AG* (Bowman and Smyth, 1999).

A number of other genes have been identified as important for gynoecium development. These include the *STYLISH* genes (*STY1* and *STY2*), *ETTIN* (*ETT*), the *NGATHA* genes (*NGA1-4*) and the *HECATE* genes (*HEC1-4*). The *STY* and *NGA* genes are important for correct vascular and style development of the gynoecium (Alvarez et al., 2009; Kuusk et al., 2002; Trigueros et al., 2009), whereas *ETT* and the *HEC* genes are important for the development of the female reproductive tract (Gremski et al., 2007; Heisler et al., 2001; Sessions et al., 1997). Direct binding of *AG* to DNA regions near some of these genes has been suggested (Kaufmann et al., 2009), however, direct regulation has not been shown. Strong *ag* mutants are epistatic over these genes, however, their expression in *ap2 ag* double mutants has not yet been investigated (for a review of carpel development see (Ferrandiz et al., 1999)).

1.3.3 Transcriptional control of floral meristem determinacy by *AG*

The indeterminate phenotype of *ag-1* mutant flowers implies that the activity of *AG* is essential to generate a determinate flower. In order to terminate the floral meristem, *AG* could influence meristem maintenance genes, cell cycle genes or hormonal pathways. The

precise mechanism of *AG* activity in this regard is unknown, however, several working models have been proposed.

The floral indeterminacy displayed in *ag* flowers is correlated with the persistent expression of the homeodomain-containing transcription factor *WUSCHEL* (*WUS*). The main role of *WUS* is to maintain stem cell populations within the meristems of developing plants (Laux et al., 1996) and it is clear that the down-regulation of *WUS* is ultimately necessary for the termination of the floral meristem. In wild-type plants, *WUS* expression begins to be down-regulated in the floral meristem at stage 6. However, in *ag-1* plants, *WUS* expression persists long after stage 6. This persistent expression leads to an overproliferation of meristematic tissue in the centre of the flower resulting in the formation of extra floral organs. *AG* is involved in terminating *WUS* expression, though available evidence suggests that this repression is not direct (Lenhard et al., 2001; Lohmann et al., 2001).

The C2H2 zinc-finger protein *KNUCKLES* (*KNU*) has also been implicated in floral meristem termination (Sun et al., 2009). The carpels and siliques of *knu* plants contain ectopic reproductive organs (see Chapter 4, Figure 4.1). The appearance of ectopic stamens and carpels within the fourth whorl results in bulging of the carpels or siliques, giving a knuckle-like appearance to the developing gynoecium (Payne et al., 2004). It has been suggested that *KNU* is a direct target of *AG* (Sun et al., 2009). Moreover, it has been shown that over-expression of *KNU* is sufficient to repress *WUS* expression in the centre of the floral meristem. Although it appears that *KNU* can repress *WUS*, it is possible that other proteins are involved in this process. Furthermore, it is unclear whether this repression is direct. An alternative explanation of the *knu* phenotype is that the ectopic organs within the siliques of *knu* plants are a result of respecification of placental tissue into another floral meristem (Payne et al., 2004). Therefore, *KNU* may function as a repressor of floral meristems in the placenta.

Similar to *knu* plants, stamens were occasionally observed within the fourth whorl carpels of *crc-1* mutants (Alvarez and Smyth, 1999). The indeterminate phenotype of *crc* flowers was enhanced when mutations for *ULTRAPETALA* (*ULT*), *REBELOTE* (*RBL*) or *SQUINT* (*SQN*) were introduced into the *crc* background. While the role for *CRC* in floral meristem determinacy remains unclear, the functions of *ULT*, *RBL* and *SQN* appears to be to positively regulate *AG* and *SUPERMAN* (*SUP*), which is another gene involved in meristem determinacy (Bowman et al., 1992; Prunet et al., 2008). Flowers of *sup* plants

contain extra whorls of stamens instead of a determinate gynoecium. This indeterminacy is eliminated when mutations in *AP3* or *PI* are introduced (Bowman et al., 1992). Therefore, B-function may also be involved in meristem determinacy (Bowman et al., 1992).

1.4 Global identification of genes downstream of the homeotic factors

Classic genetic approaches have identified many of the key regulators of flower development (Alonso and Ecker, 2006), however, the regulatory hierarchy downstream of these factors is poorly understood. Genomic technologies offer the possibility of studying the expression of genes downstream of these master regulators on a global scale. Genes that are differentially expressed in response to the perturbation of a main regulator can be inferred to be part of the same gene regulatory network. Experimental approaches aimed at identifying the regulatory elements to which transcription factors bind, such as ChIP, can be combined with expression studies in order to establish the layout of a gene regulatory network (Materna and Oliveri, 2008). Recently many genome-wide studies have focused on discovering genes that are responsive to the floral homeotic genes.

One of the first genome-wide studies performed on the floral homeotic genes aimed at determining spatial gene expression in *Arabidopsis* flowers (Wellmer et al., 2004). In this study, the transcript profiles of inflorescences of five floral homeotic mutants (*ap1*, *ap2*, *ap3*, *pi* and *ag*) were compared with that of wild-type plants. This approach allowed the analysis of genes that are predominately expressed in one floral organ type. Using a flower-specific cDNA microarray and a whole genome oligonucleotide array, the expression profiles of mutant and wild-type flowers were determined and the resulting datasets were analyzed using a subtractive approach. For example, *ag* is the only mutant that lacks carpels, whereas the other homeotic mutants, such as *ap2* and *ap3*, have extra carpelloid tissue. Therefore, one would expect genes that are preferentially expressed in carpels to be down-regulated in *ag* and up-regulated in *ap2*, *pi* and *ap3* in comparison with the wild type (Figure 1.5A). Using this method, several hundred genes were identified as being differentially expressed in a particular organ type. Most of these transcripts were detected in the reproductive organs, which is in agreement with the complex architecture of the

reproductive organs in comparison with the organs of the perianth. These datasets offered a valuable resource for identifying candidate genes for reverse genetic studies and presented the first insight into the components of the floral homeotic gene network. Although this study offered important information regarding the floral homeotic gene network, very few genes involved in early flower development were identified since the RNA used for these experiments was isolated from whole inflorescences. This led to an overrepresentation of RNA from older flower buds and therefore transcripts that are expressed late in flower development were primarily detected.

Therefore, the next step in identifying the genes regulated by the homeotic factors was to identify genes that are responsive to their activity during early flower development. A central problem for this analysis was the fact that young flower buds are minute. Moreover, each flower is initiated successively, such that a single flower in any given inflorescence is at a distinct developmental stage (Smyth et al., 1990). This sequential initiation of flowers prevents the analysis of stage-specific targets as it would be extremely difficult and time consuming to collect enough early-stage flower buds at distinct developmental stages in order to perform a genome-wide study. Two separate groups sought to circumvent these issues in similar ways (Gomez-Mena et al., 2005; Wellmer et al., 2006). Both groups used a plant where *API* and its paralog *CAULIFLOWER (CAL)* were mutated. Flowering is significantly delayed in *apl cal* plants since floral meristems are not specified correctly. In the absence of these two proteins, plants undergo an over-proliferation of inflorescence meristem-like tissue, leading to a cauliflower-like appearance (Figure 1.5B) (Ferrandiz et al., 2000).

The first of these studies introduced a null mutation for *AG* into the *apl cal* background (Gomez-Mena et al., 2005). In addition, an *AG-GR* (see Section 1.3.1) translational fusion was overexpressed in the triple-mutant background. When the inflorescences of these plants were treated with dexamethasone, carpel and stamen development was induced in a synchronous manner. This allowed the collection of large amounts of tissue that are at identical stages of development. Once reproductive organ development was induced, tissue was collected after 1, 3 and 7 days and the purified RNA was hybridized to a whole genome oligonucleotide array. From the 22,810 genes represented on the array, 149 genes were considered up-regulated and 68 were considered down-regulated during the course of the experiment. Of the up-regulated genes, 26% were annotated as transcription factors. This represents a significant enrichment when compared

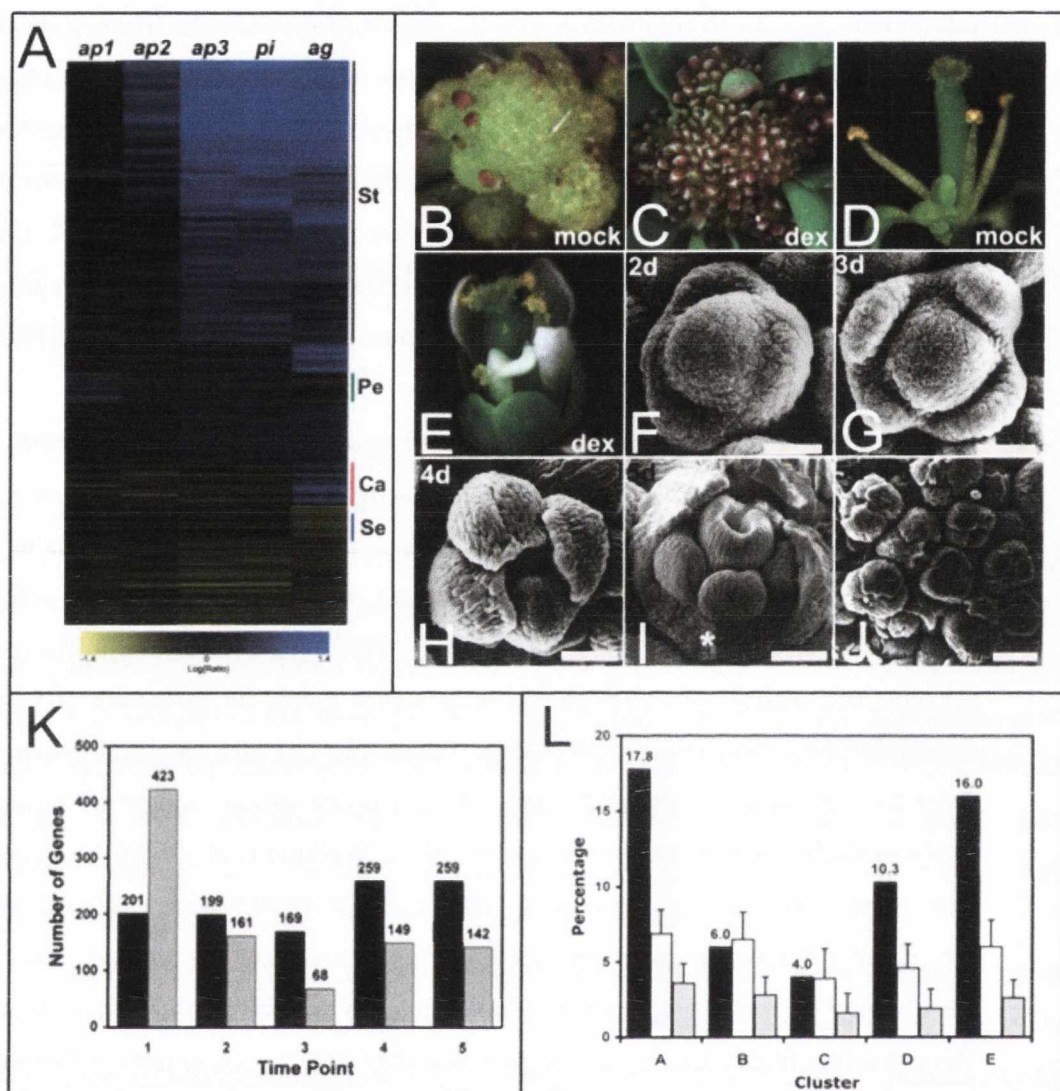


Figure 1.5 Global analysis of gene expression during early and late flower development by microarray analysis. (A) Spatial gene expression during late flower development. Comparison of the transcriptomes of wild-type, *ag-1*, *ap3-3*, *pi-1*, *ap2-2* and *ap1-1* mutant inflorescences. Genes whose expression was detected as being differentially expressed in at least one of the comparisons are represented. Elements highlighted in yellow indicate up-regulation in the wild type when compared to a mutant. Elements in blue indicate down-regulation. (B – J) The 35S:API-GR *ap1 cal* floral induction system. (B) A mock-treated inflorescence-like meristem. (C) An inflorescence-like meristem 6 days after dexamethasone treatment. A large number of developing flower buds are visible. (D) A flower lacking sepals and petals produced from a mock-treated inflorescence-like meristem. (E) A flower containing a full complement of floral organs produced from a dexamethasone-treated inflorescence-like meristem. (F–J) Scanning electron micrographs of flower buds produced from 35S::API-GR plants at different days after dexamethasone treatment (as indicated). These flower buds develop similarly to wild-type floral buds. (K) The number of genes up-regulated (black bars) or down-regulated (grey bars) 1-5 days after activation of flower development in a 35S::API-GR *ap1 cal*

to the genome-wide proportion of transcriptions factors of ~6%, indicating that a complex transcriptional hierarchy exists in early flower development. Moreover, this dataset was compared with that of Wellmer et al. (2004) (see above) and an overlap of only 20 genes was observed. This indicated that the transcriptional programs of late and early flower development are distinct. Finally, ChIP assays demonstrated that AG could bind its own promoter and that of *AP3*, in addition to several other genes. Therefore, it was concluded that AG auto-regulates itself in the third and fourth whorl and regulates *AP3* in the third whorl alone. This was already suspected given that 35S::AG-GR *ap1 cal* plants only require a single dexamethasone-treatment (Ito et al., 2004) in order for normal flower development to proceed, whereas, plants with an additional mutation in *AG* require continuous dexamethasone treatments in order to restore reproductive organ identity.

The second study that used the *ap1 cal* background reintroduced AP1 using the GR fusion technique (Wellmer et al., 2006). After dexamethasone treatment of the inflorescence-like meristems of 35S::AP1-GR *ap1 cal* plants (this line is referred to as the ‘floral induction system’ or FIS, hereafter), a synchronous proliferation of flower buds was observed (Figure 1.5B – J). Of the 26,700 genes represented on the oligonucleotide array used in this study, 1,653 were considered differentially expressed when compared to untreated plants over the time points taken (tissue was harvested before dexamethasone treatment and 1 – 5 days after dexamethasone treatment). The first day after treatment saw the majority of differentially expressed genes down-regulated, while at later time points, the majority of differentially expressed genes were up-regulated (Figure 1.5K). Of the 1,653 genes that were differentially expressed, 222 coded for transcription factors. This represented a 13.4% enrichment, in comparison to the genome-wide representation of transcription factors, again indicating an extensive transcriptional hierarchy in early flower development. Another finding of this study suggests a high degree of functional redundancy among genes controlling early flower development. Of the 222 transcription factors that

Figure 1.5 continued. background. (L) The proportion of closely related genes that were differentially expressed over several developmental stages (black bars). White and grey bars are the proportion of closely related sequences in equally sized sets of sequences that were randomly chosen from the dataset of differentially expressed genes during early flower development and the *Arabidopsis* genome, respectively. Clusters: genes predominantly expressed in inflorescence meristems (cluster A), in floral buds of stages 2-3 (cluster B), stages 3-4 (cluster C), stages 4-7 (cluster D), and stages 5-7 (cluster E). Floral stages are according to Smyth et al. (1990). Figures adapted from (Wellmer et al., 2006; Wellmer et al., 2004).

were differentially expressed, 91 fell into 37 distinct groups based on sequence similarity and co-expression data. Further analysis of all differentially expressed genes revealed the co-expression of genes with similar sequences in stages 4 – 7 of flower development and immediately after AP1 induction (Figure 1.5L). Therefore, if a large degree of functional redundancy did exist, disrupting the function of one gene would probably not result in a mutant phenotype as other genes may compensate for its loss.

Recently, several studies have been published that use chromatin immunoprecipitation combined with ultrahigh-throughput sequencing (ChIP-Seq) to identify the binding sites of transcription factors involved in flower development. The first genome-wide binding study attempted to identify the direct targets of the E-class gene *SEP3* using wild type and *ag-1* whole inflorescences (Kaufmann et al., 2009). *SEP3* appears to form higher-order complexes with the B and C class homeotic factors (Honma and Goto, 2001). Therefore, identification of *SEP3* binding sites should yield some information on the binding sites of the B and C class of transcription factors as well. Using a peptide antibody specific to *SEP3*, over 4,000 and 2,800 *SEP3*-binding sites were observed in wild type and *ag-1* inflorescences, respectively. The subsequent analyses revealed that the binding sites were enriched for sequences located near the coding sequences of transcription factors, indicating a transcriptional hierarchy, which is in agreement with the transcriptomics data described above. In addition, sequences close to the binding sites were enriched for consensus sequences belonging to transcription factor families other than the MADS-box transcription factors. This raised the possibility that the MADS-box proteins interact directly with non-MADS-box proteins to bind DNA and affect transcription. Unfortunately, the fact that *SEP3* interacts with the B- and C-function proteins precludes any interpretation of how B and C class proteins individually direct flower development. Moreover, the use of whole inflorescences may have resulted in identification of binding sites that are only relevant at later stages of flower development.

The next study to utilise ChIP-Seq focused on the transcription factor AP1 (Kaufmann et al., 2010b). Using the FIS (*i.e.* the *35S::AP1-GR ap1 cal* line described above) and an antibody specific to AP1, over 1,900 genomic regions were found to be significantly enriched. The inflorescence meristem-like tissue used for ChIP was collected 2 hours after dexamethasone treatment. In order to complement the binding studies, microarray analyses were performed on mRNA extracted from the same type of tissue collected 2, 4, 8 and 12 hours after treatment with a dexamethasone-containing solution. Ensuing analysis of the

data revealed that AP1 directly represses genes involved in the repression of flowering and therefore activates flowering by preventing the expression of negative regulators of this process. An important finding in this paper was that the number of genes bound by AP1 is much higher than the number of genes differentially expressed upon AP1 activation. For example, whereas ~2,000 genes were identified as putative AP1 targets (*i.e.* a binding site was detected within 3 kb upstream of the 5' end or 1 kb downstream of the 3' end of a gene), only 249 of these genes also showed a robust (> 1.8 -fold) change in expression. This indicates that the binding of a transcription factor to a promoter sequence of a given gene does not necessarily translate into expression changes. This also underlines the importance of generating complementary expression data when interpreting transcription factor binding data.

Although there have been massive advances in our understanding of flower development over the last decade due to genomic technologies, all the experiments described here share a common flaw. No tissue specific information has been derived from these studies, offering, at best, spatial resolution at the organ level (Jiao and Meyerowitz, 2010; Wellmer et al., 2004). Several strategies can be implemented to deliver tissue specific expression data including laser capture micro-dissection and fluorescence-activated cell sorting (Birnbaum et al., 2003; Kerk et al., 2003), however, they are beyond the scope of the current project.

1.5 Disentangling gene regulatory networks

In order to untangle a gene regulatory network (GRN), several criteria must be satisfied (Materna and Oliveri, 2008). First, the general layout of the GRN must be determined. Second, the main regulatory genes must be characterized with spatial and temporal resolution. Third, perturbation experiments must be performed in order to reveal the cause-effect linkages between regulatory genes. More elaborate perturbation approaches may be needed in order to verify the effect of one regulator on its downstream target if it is initially unclear. Ultimately these experiments would reveal a reasonably accurate description of the network architecture. Fourth, the interaction of transcription factors with *cis*-regulatory elements needs to be established in order to distinguish between direct and indirect regulation (Materna and Oliveri, 2008). Many components of the flowering

pathway have been characterized in detail; however, the genes that are activated by the homeotic factors are largely unknown. Next, the methods that can be used in *Arabidopsis* to identify genes regulated by the floral homeotic factors will be described.

1.5.1 The GRN of flower development

In the case of flower development, the first two criteria discussed in relation to characterizing a gene network (see above) have been satisfied. The inductive signal to transition from a vegetative state to a reproductive state is the starting point of this GRN. This process has been very well described in *Arabidopsis* (Ausin et al., 2005). In addition, the main regulatory genes, the floral homeotic genes, have been extensively characterized yielding high-resolution spatial and temporal expression patterns (Krizek and Fletcher, 2005).

Although very little information is available as to which genes the floral homeotic factors regulate, recent studies have shed some light on some of the genes these transcription factors affect (Gomez-Mena et al., 2005; Jiao and Meyerowitz, 2010; Kaufmann et al., 2009; Kaufmann et al., 2010b; Wellmer et al., 2006; Wellmer et al., 2004). ChIP is an extremely robust method for identifying binding sites *in vivo*, however, the information generated by this assay should be combined with expression data in order to identify direct target genes. To this end, perturbation of gene function is required. In *Arabidopsis*, several options are available for this purpose (Alonso and Ecker, 2006). Below, the strategies and methods most promising for the characterization of the role of *AG* in flower development will be described.

1.5.2 Strategies for gene network perturbation

As previously mentioned, cause-effect linkages in the GRN of interest can be identified by perturbation analysis. In *Arabidopsis* there are several options available that allow perturbation of transcription or translation and most of these options can be placed under an inducible system in order to perturb gene function at distinct developmental stages. These options include artificial microRNAs (amiRNAs), RNA interference (RNAi), dominant negative versions of a transcription factor of interest and complete null alleles.

1.5.2.1 Artificial microRNAs

In recent years, microRNAs (miRNAs) have been shown to be important for the regulation of many different biological processes in plants, including flower development (Chen, 2004; Laufs et al., 2004; Palatnik et al., 2003). Conserved miRNAs are often present in multiple copies in the genome, whereas, non-conserved miRNAs are usually present in a single copy (Fahlgren et al., 2007). MiRNAs are transcribed by RNA polymerase II into RNA hairpin structures that are bound and processed by several enzymes, such as DICER-LIKE1 (DCL1), and binding proteins, such as HYPONASTIC LEAVES1 (HYL1) into a miRNA duplex (Mallory et al., 2008). This duplex is then methylated on the ribose of the 3' nucleotide by the methyltransferase, HEN1, which prevents its degradation (Chen, 2005). Next, the RNA transporter HASTY exports either the miRNA duplex or single-stranded miRNA into the cytoplasm where the more stable of the miRNA duplex strands associates with its appropriate ARGONAUTE protein. This complex then targets mRNA transcripts for degradation, translational repression or, alternatively, epigenetic modification of the corresponding locus through the trans-activated small interfering RNA pathway (Mallory et al., 2008; Mi et al., 2008).

Schwab et al. (2006) recognized the potential that miRNAs have as a method for the characterization of gene function (Schwab et al., 2006). From extensive computational and experimental analysis several guidelines indicating how miRNAs function in plants were proposed (Schwab et al., 2005). In summary, it was noted that there were overrepresentations of certain nucleotides at certain positions, that mismatches at the 5' end of miRNAs were usually not present and that there were a limited number of mismatches allowed in the 3' portion of the miRNAs. From these guidelines, a system was generated in order to perturb gene function with artificially generated miRNAs (amiRNAs). Several important points were made about the advantages and disadvantages of the system. First, it appeared that there were minimal off target effects observed from over-expressing an amiRNA. This was shown by thorough and comprehensive microarray analysis. Second, it is possible to generate amiRNA-resistant versions of a gene of interest by creating synonymous mutations in the region bound by a given amiRNA, leading to less efficient target recognition. This is important so that rescue constructs can be generated. Third, transitivity (where an siRNA can seed the generation of new siRNAs) was not observed. Fourth, tissue-specific and inducible promoters can be used in order to drive amiRNA

expression, however, the amiRNAs may not function cell autonomously. Importantly, it was also observed that the effects from an inducible promoter were transient, therefore gene expression resumed as normal once the inductive conditions were removed. Fifth, different transcripts appear to differ in susceptibility to amiRNA directed regulation. Sixth, a single amiRNA can be designed in order to knockdown gene function of groups of genes, although this appears to be more difficult than disturbing the function of a single gene. Seventh, and finally, the endogenous backbones used to generate amiRNAs differed in their ability to down-regulate their target mRNAs.

Given the demonstrated efficacy of the amiRNAs by Schwab et al. and the subsequent effective use of amiRNA in perturbing gene function (Alvarez et al., 2009; Bomblies et al., 2007; Niu et al., 2006; Stahle et al., 2009), this system has a great potential to aid the discovery of gene functions. Moreover, the generation of a web-based amiRNA prediction software and a reasonably high-throughput system for designing and cloning amiRNAs, respectively, grant the plant community with an invaluable resource to characterize gene function.

1.5.2.2 RNA interference

RNAi is based on the conversion of inverted repeats in the genome into long double-stranded RNA (dsRNA) complexes that are ultimately converted into small interfering RNAs (siRNAs). DICER-LIKE3 processes this complex into smaller 24-nucleotide long siRNAs. Finally, these siRNAs associate with either AGO4 or AGO6 in order to effect gene expression. The processing and functionality of the siRNAs also require isoforms of RNA polymerase IV (Ramachandran and Chen, 2008).

The integration of inverted repeats complementary to a gene of interest into the genome has been shown to be efficient at perturbing gene expression. Moreover, the introduction of an intron between the inverted repeats appears to increase the efficiency of gene silencing (Wesley et al., 2001). Although RNAi has been shown to be a valuable system for perturbing gene expression, there are several disadvantages when compared to amiRNAs. First, the generation of a construct with inverted repeats is more time-consuming as care must be taken when choosing a region to clone in order to avoid regions that are highly conserved. This is particularly problematic in the case of *AG*, as *Arabidopsis*

contains many genes that are paralogous to *AG*. Second, there is a higher chance of creating unwanted off-target effects, as the RNAi constructs are much longer than amiRNAs. Moreover, the risk of transitivity is higher in this case as there are no mismatches present in the subsequent siRNAs. Third, generating an siRNA-resistant construct may be problematic as it would be unknown which siRNA(s) are targeting the mRNA of interest.

Nonetheless, RNAi still is a useful tool for gene perturbation. Well-established vectors have been generated and inducible systems have also been designed (Guo et al., 2003; Wielopolska et al., 2005). As Schwab et al. (2006) found it difficult to perturb the function of certain genes using amiRNAs; it is conceivable that RNAi might succeed where amiRNAs fail.

1.5.2.3 Dominant negative repression

Several domains within transcription factors that are involved in transcriptional repression have been identified and characterized (Hiratsu et al., 2002; Ikeda et al., 2009; Ohta et al., 2001; Payne et al., 2004; Sakai et al., 1995). It has been demonstrated that these domains can be fused to other transcription factors in order to convert them into dominant transcriptional repressors. Other domains, such as VP16, can convert transcription factors into transcriptional activators (Hiratsu et al., 2003; Mitsuda et al., 2006; Triezenberg et al., 1988). Chimeric repressors have been, in certain cases, shown to faithfully reproduce the null mutant phenotypes of the transcription factors to which they are fused, however, the over-expression of these chimeras is required in order to out-compete the endogenous transcription factors. This over-expression may lead to the binding of these chimeras to inappropriate loci, resulting in their undesired repression. In addition, inducible transient expression has not been demonstrated. Although RNAi and amiRNAs appear to be more specific and have certainly been characterized more thoroughly, they may fail to perturb gene function as desired (Schwab et al., 2006). Therefore, chimeric repressors present an alternative and potentially valuable method for knocking down gene function. However, given the potential problems outlined above, caution should be exercised when using this method to analyse gene expression.

1.5.2.4 Null mutant alleles

Comparing the transcriptome of a loss of function mutant to a wild-type counterpart has been extensively used to investigate the role(s) of a gene of interest on global gene expression (Guo et al., 2009; Morohashi and Grotewold, 2009; Wellmer et al., 2004). Although these experiments can be informative, they do not offer the same resolution as inducible perturbation since, in the case of the homeotic mutants, entire organs are absent. Therefore, the differential expression of a given gene may be due to the loss of function of *AG* or the absence of that organ type (see Section 1.4). Although inducible perturbations would reveal more about the transcriptional control of organogenesis throughout flower development, comparing the transcript profiles of the homeotic mutants to wild-type flowers would identify the components downstream of the homeotic factors.

1.5.3 Chromatin immunoprecipitation

Chromatin immunoprecipitation (ChIP) is a technique for assaying protein-DNA binding *in vivo* (Park, 2009). The premise of ChIP is to identify genomic regions that are bound by a DNA-binding protein of interest. However, it can also be used to study epigenetic modifications of chromatin such as histone methylation or acetylation. In a ChIP protocol, tissue is harvested and then DNA-binding proteins are chemically cross-linked to the DNA (Figure 1.6A, B). The tissue is then processed until only the chromatin remains. The chromatin is then fragmented in order to generate shorter fragments (Figure 1.6A, C), which can then be precipitated by an antibody specific to the protein of interest. The cross-linking is then reversed and the DNA is purified and analyzed (Figure 1.6A, D – F). The DNA sequences over-represented in the subsequent analysis should correspond to the binding sites of the protein of interest (Figure 1.6D) (Robertson et al., 2007; Wang et al., 2002).

There are several different methods available to analyse the DNA output from a ChIP assay including ultra high-throughput sequencing (ChIP-Seq), microarray hybridization (ChIP-chip) and PCR analysis. Quantitative real-time PCR (qPCR) analysis is a reliable method to analyse data of this type as well. The specificity and simplicity of qPCR are major advantages, though its low throughput and relatively high cost are limiting factors. In

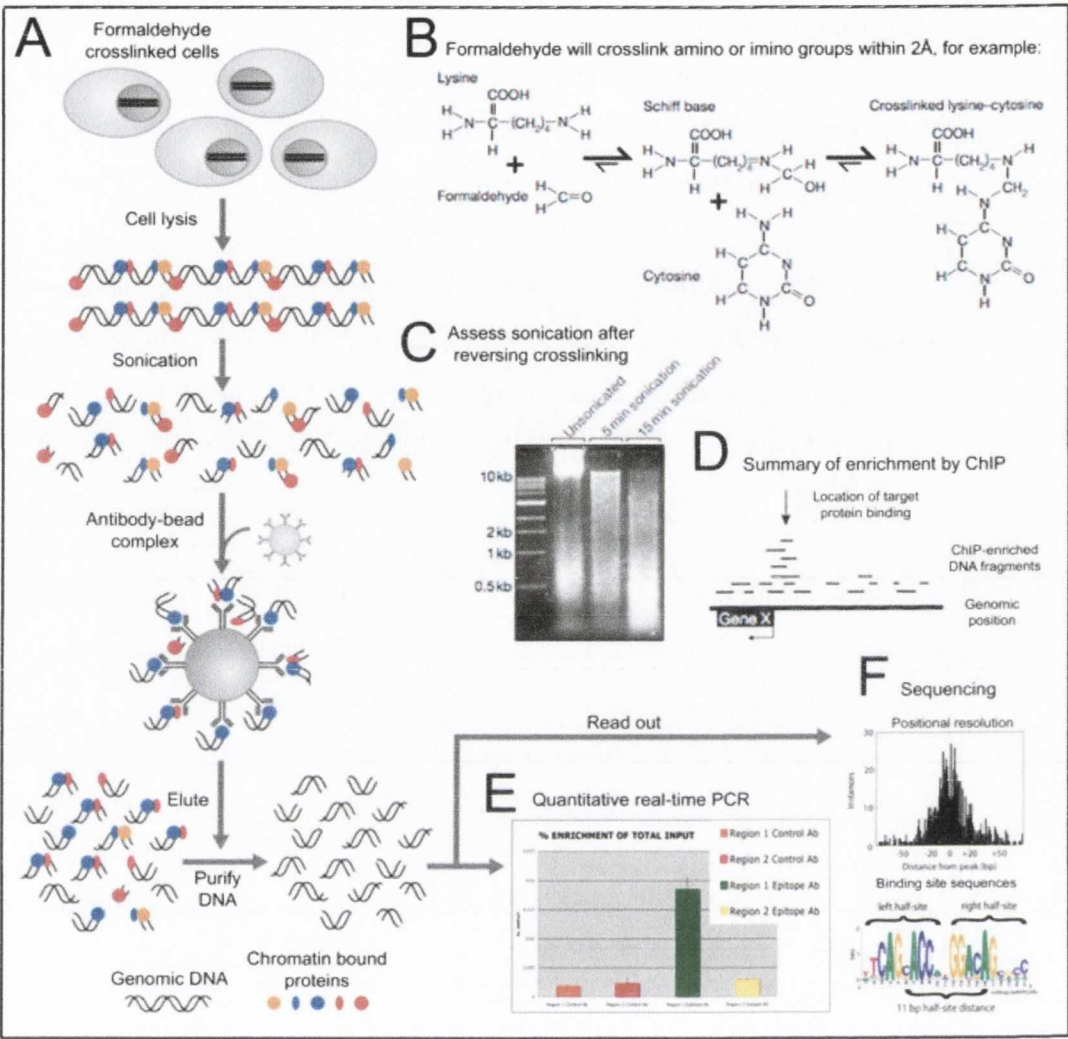


Figure 1.6 Chromatin immunoprecipitation (ChIP). (A) After formaldehyde cross-linking, cell lysis and chromatin fragmentation, genomic regions bound by a transcription factor are immunoprecipitated using a specific antibody against the protein of interest. (B) The chemical basis of chemically cross-linking chromatin with formaldehyde. (C) The effect of sonication on the fragment size of chromatin. (D) The expected distribution of genomic fragments after the ChIP procedure. Enrichments should be observed at specific points of the genome in comparison to the background. (E) Quantitative real-time PCR for selected target genes can be used to verify that the ChIP procedure was successful before a global target gene analysis is conducted (F). The lower segment of panel F shows identification of consensus binding sites. Figure modified from (Johnson et al., 2007; Massie and Mills, 2008).

order to perform a genome-wide study of transcription factor binding sites, sequencing or microarray technologies must be used (Gregory et al., 2008). Ultra high-throughput sequencing offers significant advantages over microarray platforms. First, a higher base pair resolution can be obtained from ChIP-Seq. Although the density of probes on whole-

genome tiling arrays is increasing, they are still relatively expensive and technical difficulties due to the hybridization process can still impede accuracy. Second, the signal to noise ratio is often much higher when ChIP-Seq is used. Again, the hybridization process can increase background signals due to inaccurate binding of a probe to non-complementary DNA.

In order to perform ChIP successfully, an antibody that efficiently and specifically recognizes the protein of interest is required (Lee et al., 2006). To this end, there are two different options: an antibody that recognizes the endogenous protein can be generated or the protein of interest can be fused to an epitope-tag. Generating antibodies is a time-consuming and expensive exercise, and their suitability for the use in ChIP assays is not guaranteed. In contrast, epitope tagging is cost-effective, however, the generation of a tagged protein which functions in the same manner as the equivalent endogenous protein can be laborious. The advantage of epitope tagging is that well-characterized commercially available antibodies can be used. In any case, the antibody chosen for immunoprecipitation must be extensively tested as many commercially available antibodies have been reported to be unsuitable for use in a ChIP assay (Park, 2009). Given the success of employing ChIP in other organisms and the gradually lowering costs of ultra high-throughput sequencing, ChIP-Seq offers the opportunity to deconstruct the regulatory hierarchy in flower development.

1.6 Project outline

In the following project outline, experiments are described that aim at obtaining detailed insights into the processes underlying flower formation in *Arabidopsis thaliana*. Although the homeotic factors have been extensively investigated, the genes which they regulate, the *cis*-regulatory elements to which they bind and the regulatory interactions between them remain, for the most part, unknown. Therefore, new experimental approaches are necessary to obtain a detailed insight into the developmental mechanisms that regulate flower development. As a master regulator of reproductive organ development, *AG* is of

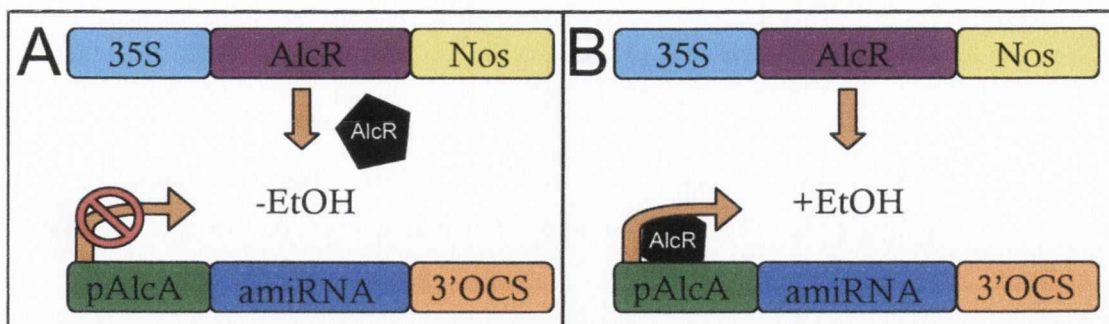


Figure 1.7 The *pAlcA/AlcR* ethanol-inducible system. (A) The *AlcR* transcription factor is constitutively expressed from a Cauliflower Mosaic Virus (CaMV) 35S promoter; however, it is unable to bind the *AlcA* promoter in the absence of ethanol (EtOH). (B) Once ethanol is present, the *AlcR* transcription factor is able to bind the *AlcA* sequence and drive the transcription of the *amiRNA* of interest. 3'OCS; octopine synthase gene terminator sequence, Nos; nopaline synthase gene terminator sequence.

central importance. As outlined in Section 1.3, to date, relatively few genes have been identified to be directly regulated by AG have been identified. Therefore, a strategy that will allow the identification of genes regulated by AG has been developed.

In order to characterize the GRN that AG controls, it is necessary to identify both the genes that are responsive to AG expression and those whose promoters are bound by AG. Inducible perturbation systems mentioned above can be used to identify genes responsive to AG, whereas ChIP-Seq can be used to discover the genomic regions bound by AG. Using the FIS, the direct and indirect targets of AG can be characterized at specific stages of development. The FIS is the more robust of the two systems generated to induce flower development (see Section 1.4). Indeed, the reintroduction of AP1 is more appropriate than that of AG since it compensates for the loss of the endogenous copy in the FIS. Moreover, the response from dexamethasone-treated AP1-GR plants is more quantitative, as all plants treated respond in the same manner, whereas as few as 30% of the dexamethasone-treated AG-GR plants respond to the inductive treatment. Finally, all organs are present in the dexamethasone-treated AP1-GR lines, whereas only stamens and carpels develop in AG-GR treated plants. Therefore, the AP1-GR FIS will be used to perform the experiments described below.

1.6.1 Perturbation analysis

As previously discussed, there are several options available to perturb specific components of a GRN. Here, all three inducible options and use of the null mutant approach will be tested. In order to make the appropriate constructs inducible, they will be placed under the *AlcA* promoter that is bound by the ethanol-responsive transcription factor, AlcR from *Aspergillus nidulans* (Figure 1.7) (Roslan et al., 2001). This system of regulation has been shown to be highly responsive to ethanol treatment, allowing transcription levels similar to a constitutive 35S promoter to be reached (Schwab et al., 2006). Initially, a construct (e.g. an *AG*-specific amiRNA) under the control of the 35S promoter that phenocopies *ag* flowers in a wild-type background will be identified. Once a suitable construct has been found, it will be placed under the control of the ethanol inducible system and tested in wild-type plants for an ethanol-induced phenocopy of the *ag-1* null mutant phenotype. When a suitably responsive line is identified, it will be crossed with the FIS. Ultimately, this system will allow the study of stage-specific targets of AG by first activating AP1-GR with a dexamethasone-containing solution, and then disrupting *AG* gene function at different stages of flower development using the ethanol-controlled expression system. Tissue will be harvested after ethanol-treatment and RNA will be isolated. The RNA will then be hybridized to a whole-genome oligonucleotide microarray platform.

Given that the amiRNAs, RNAi constructs and dominant negative constructs do not guarantee an *ag* phenocopy, the null mutant approach will also be used. Although it is not an ideal approach, it will provide the expression data required to complement the ChIP data. The *ag-1* mutation has been crossed into the AP1-GR FIS. In this case, flower development would be synchronously initiated, in both plants with an endogenous, functional copy of *AG* and in plants homozygous for the *ag-1* mutation. Tissue would then be collected at different points of flower development. Although this strategy may not be as informative as the inducible perturbation experiments, it would be an important complement to the ChIP data, should the other techniques fail to recapitulate the *ag-1* phenotype.

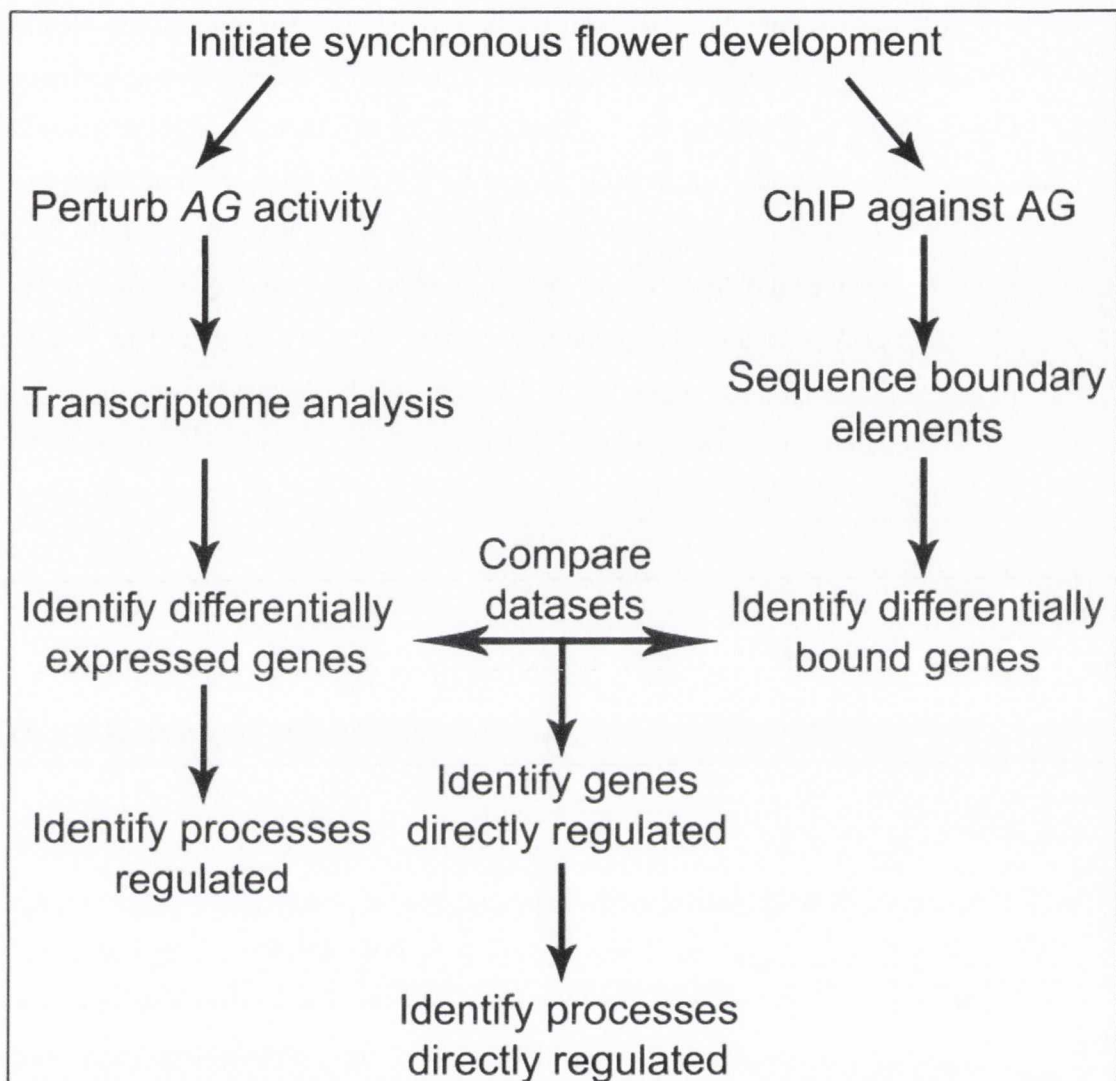


Figure 1.8 Experimental design to identify processes regulated by AG. Using the FIS to synchronize flower development, one line of experimentation will aim at determining the effects that a perturbation of AG activity has on the flowering GRN at distinct stages of flower development. In addition, via ChIP-Seq, genomic regions bound by AG will be identified. The datasets from these experiments can then be combined allowing the discovery of genes and processes controlled by AG either directly or indirectly.

1.6.2 Identification of direct AG targets by ChIP-Seq

Given the advantages high-throughput sequencing offers in comparison to tiling arrays, ChIP-Seq will be the method of choice for this study. Since no member of the Wellmer lab has ever performed a ChIP assay, it will first be optimized and validated by

qPCR using inflorescence tissue from the FIS. Several genes have been identified as direct targets of AG and can be analyzed for enrichment of AG at their promoters in order to validate and optimize the ChIP procedure. To ensure that an antibody of suitable quality can be used, a dual strategy of generating peptide antibodies complementary to the AG protein and generating tagged lines of AG that will be tested for their ability to rescue the *ag-1* phenotype will be followed. Peptides will be designed that are likely to be on the outer surface of AG, according to protein structure predictions. These designs will be synthesized and rabbit polyclonal antibodies will be generated. *Modified green fluorescent protein 5 (mGFP5)* and a *haemagglutinin (HA)* tag will be translationally fused to the entire *AG* locus. Commercially available antibodies are available for both epitope tags and several ChIP assays have been performed using these epitope tags (Cole et al., 2009; Znaidi et al., 2009).

1.6.3 Identification of molecular processes regulated by AG

The generation of lists of genes that are, directly or indirectly, responsive to AG activity will allow the identification of molecular processes that AG regulates. Moreover, doing this in a stage-specific manner using the FIS will allow a temporal characterisation of AG activity and further the understanding of how organ identity is specified. Identifying sequences to which AG binds will then allow the refinement of these data to separate processes that are directly regulated by C class activity from those indirectly regulated through a transcriptional hierarchy (Figure 1.8). Not only will these data allow the generation of a model of AG function, they will also identify candidate genes that can be analysed using reverse genetics tools in order to further dissect the GRN of flower development.

Chapter 2 – Materials and Methods

2.1 Strains and growth conditions

2.1.1 Plant Lines

The previously published plant lines used in this study are:

Landsberg *erecta* (L-*er*) wild-type

L-*er ag-1* (Bowman et al., 1989)

L-*er ap1-1 cal-1* (Bowman et al., 1993)

L-*er 35S::API-GR ap1-1 cal-1* (Wellmer et al., 2006)

L-*er pCLV1::LhG4* (Schoof et al., 2000)

L-*er pCLV3::LhG4* (Lenhard and Laux, 2003)

L-*er pWUS::LhG4* (Gross-Hardt et al., 2002)

L-*er 35S::API-GR ap1-1 cal-1 ag-1* (generated in this study)

2.1.2 Bacterial Strains

Escherichia coli strain **XL1-Blue** – Genotype *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F' *proAB lacIZΔM15 Tn10* (Tet)] (Stratagene)

Escherichia coli strain **Tp611** – Genotype *recBC hsdRM cya610 pcn-12* (Glaser et al., 1993)

Escherichia coli strain **STBL2** – Genotype F- *endA1 glnV44 thi-1 recA1 gyrA96 relA1 Δ(lac-proAB) mcrA Δ(mcrBC-hsdRMS-mrr) λ* (Invitrogen)

Agrobacterium tumefaciens strain **C58 pGV2260** (McBride and Summerfelt, 1990)

2.1.3 Plant Growth Media

Plants were grown on a medium consisting of compost, perlite and vermiculite (NAD, Dublin, Ireland) in a ratio of 5:2:3 (V:V:V).

0.5xMS agar plates (Murashige and Skoog medium) – 2.2 g/L MS salts, pH 5.7, 6 g/L agar

2.1.4 Bacterial Growth Media

LB (Luria-Bertani medium) – 10 g/L bacto-tryptone, 5 g/L bacto-yeast extract, 10 g/L NaCl (15 g/L agar was added to make LB plates).

2.2 Cloning and Genotyping

2.2.1 Plasmid DNA preparation

Plasmids were extracted from XL1-Blue and STBL2 *E. coli* cultures using the GeneJet™ plasmid miniprep kit (Fermentas). Plasmid preparation from Tp611 was done using a boiling lysis method. Briefly, a culture of Tp611 was pelleted by centrifugation and the pellet was resuspended in 60 µL TES buffer (see Appendix 2 for all buffer compositions). The bacterial suspension was then incubated on ice for 10 min with 20 µL of a solution containing 40 mg/mL lysozyme. 550 µL of MSTET buffer was then added and the mixture was incubated at room temperature for 10 min, and subsequently boiled for 60 s before being centrifuged to pellet the cell debris. The mixture was then incubated at room temperature for 10 min with a 10 µL solution of RNase (10 mg/mL) before binding buffers and columns from the GeneJet™ plasmid miniprep kit (Fermentas) were used to purify the plasmid DNA, following the instructions included in the GeneJet kit.

2.2.2 Genomic DNA extraction using the CTAB method

This procedure was applied to generate contaminant-free genomic DNA preparations from plant tissue, which were used as template in PCR reactions for cloning purposes. Briefly, plant tissue was ground in CTAB buffer (Appendix 6.2) and purified by phenol-chloroform extraction with phase-lock tubes (5-prime) twice. The DNA was then precipitated using isopropanol and resuspended in molecular biology grade water once the pellet had dried.

2.2.3 cDNA synthesis

Total RNA was isolated from plant tissue using the Spectrum Plant Total RNA kit (Sigma-Aldrich) according to the manufacturer's instructions. RNA was reverse transcribed using oligo(dT)₁₈ primers (Fermentas) and the RevertAid H Minus reverse transcriptase (Fermentas) according to the manufacturer's instructions. cDNA obtained from this reaction was used for PCR or quantitative real-time PCR.

2.2.4 Preparation and transformation of competent *Escherichia coli*

Two separate protocols were used to generate competent *E. coli*. Competent XL1-Blue and STBL2 were generated using protocol 1. Competent Tp611 was generated using protocol 2.

Protocol 1: Competent cells were generated as described in Inoue et al. (1990).

Protocol 2: The Tp611 strain was streaked on a LB agar plate and incubated at 37°C overnight. The following day, a single colony was grown at 37°C in 3 mL LB. This pre-culture was used to inoculate 40 mL of LB and incubated at 37°C until an OD₆₀₀ of 0.5 was reached. The culture was then chilled on ice for 10 min, and spun at 4,000 rpm for 6 min at 4°C. The pellet was then resuspended in 20 mL of an ice-cold solution containing 100 mM CaCl₂ and 25% glycerol, and incubated on ice for 20 min. The cell suspension was then

spun at 4,000 rpm for 6 min at 4°C and resuspended in 1 mL of an ice-cold solution containing 100 mM CaCl₂ 25% glycerol. Aliquots of 200 µL were frozen in liquid nitrogen and stored at -80°C.

All competent cells were transformed by incubating for 10 min cells that had been thawed at 4°C with the plasmid of interest. Afterwards the cells were incubated at 42°C for 1 min. The cells were then incubated at 37°C for 1 h with 1 mL of LB. After incubation an aliquot of the transformed cells were plated on an LB agar plate containing the appropriate antibiotic(s). If a low number of transformants were expected the cells were pelleted, the supernatant was removed and the pellet was resuspended in 200 µL LB before plating. Plates were incubated overnight at 37°C.

2.2.5 Preparation and transformation of competent *Agrobacterium tumefaciens*

In order to generate competent cells, a 5 mL culture of *A. tumefaciens* was grown overnight at 28°C. Then 250 mL of LB supplemented with 0.2 g/L MgSO₄ was inoculated with 0.5 mL of the 5 mL culture. This culture was grown at 28°C until an OD₆₀₀ between 1 and 2.5 was reached. The cells were then pelleted by centrifugation and washed with ice-cold CaCl₂. Subsequently, the suspension was pelleted and resuspended in 4 mL ice-cold CaCl₂. 100 µL aliquots were made and frozen in liquid nitrogen. In order to transform the competent cells, an aliquot of cells was thawed and incubated with the plasmid of interest for 5 min in liquid nitrogen. The cell suspension was then thawed at room temperature and 1 mL of LB was added. Afterwards, the cell suspensions were incubated at 28°C for 3-4 h. After incubation the cells were pelleted though centrifugation and the pellet was resuspended in 500 µL LB. An aliquot was then plated on an LB agar plate containing rifampicin (50 µg/mL), carbenicillin (100 µg/mL) and spectinomycin (100 µg/mL). Plates were incubated at 28°C for approximately 3 days.

2.2.6 Quick Arabidopsis genomic DNA preparation

The method used results in crude genomic DNA preparations, which were used for genotyping assays. The procedure was performed as described in Edwards et al. (1991).

2.3 Plant transformation and selection of transformants

2.3.1 Plant transformation and seed sterilization

Plants were transformed by the floral dip method as described in Zhang et al. (2006). If selection of transformants was to be done on MS agar plates, seeds from transformed plants were sterilized using hypochlorous acid generated by mixing 100 mL bleach (Domestos) with 3 mL concentrated hydrochloric acid. Sterilization was performed for 3-4 h in a sealed container under a chemical fume hood.

2.3.2 Selection of transgenic plants

In order to select for transformants that carry a T-DNA with a kanamycin resistance gene, sterilized seeds were plated on 0.5xMS agar containing 50 µg/mL kanamycin. The plates were placed in the dark at 4°C for 3 days and allowed to germinate under long day conditions. Plants that formed true leaves were transplanted onto soil and genotyped for the presence of the correct transgene.

In order to select for transformants carrying an ammonium-glufosinate (BASTA) resistance gene, seeds were germinated on soil and seedlings were sprayed three times in three-day intervals with a 0.15% (v/v) BASTA (Bayer) solution. The seedlings that formed true leaves were transplanted onto fresh soil and genotyped for the presence of the transgene.

2.4 Microscopy Techniques

2.4.1 Light microscopy

Dissection of flowers was done using an Olympus SZX7* zoom stereomicroscope. Photographs were taken using an Olympus SP-500UZ digital camera mounted directly to a microscope eyepiece using a camera-microscope adapter.

2.4.2 Epifluorescence microscopy

Visualization of GFP was performed using an Olympus BX61 fluorescence microscope with excitation wavelengths of 460/480 nm and emission wavelengths of 495-540 nm.

2.4.3 Confocal imaging

In order to visualize the AG-GFP fusion protein, inflorescences from homozygous *ag-1* plants carrying the *pAG::AG-mGFP5* transgene were dissected so that older flowers were removed. The stem of the dissected inflorescences was then placed in a 0.5xMS medium with 2% agarose, and the inflorescence itself was covered with de-ionized water. A Zeiss LSM 510 confocal microscope equipped with a 40x C-Apochomat Lens was used to visualize the GFP-tagged AG protein. GFP fluorescence was excited with a 488 nm argon ion laser.

2.5 Treatments of Lines Encoding Inducible Promoters

2.5.1 Ethanol treatment of plants

Plants were placed in a sealable container with two 50 mL centrifuge tubes (Corning) containing 10 mL 100% ethanol each. Each tube was placed at opposite ends of the container. Mock-treated plants were placed in a similar container, but ethanol was replaced with ddH₂O.

2.5.2 Treatment of plants with a dexamethasone-containing solution

Once plants had bolted > 1 cm, the inflorescence-like meristems of plants were treated with a solution containing 10 μ M dexamethasone (Sigma) and 0.015% Silwet-77 using a plastic Pasteur pipette.

2.6 Generation of amiRNAs

The amiRNAs were designed using the Web microRNA designer (<http://wmd3.weigelworld.org>). Briefly, the Arabidopsis Genome Initiative (AGI) number of the gene of interest was submitted to the database and a list of microRNAs were presented which have the potential of targeting the mRNA of the gene of interest. These amiRNAs are ranked by different criteria such as amiRNA sequence composition; mismatch positions, hybridization energy when paired to intended targets, number of other genes that have 5 or less mismatches to the amiRNA, and hybridization energies of other genes with 5 or less mismatches to the amiRNA. The ranking system divides the potentially functional amiRNAs into four different color-coded classes. The amiRNAs predicted to perturb gene function most efficiently are colored in green, followed by intermediate ones in yellow and orange, and the remaining ones in red.

Once an amiRNA sequence is chosen, an Oligo Designer application (<http://wmd3.weigelworld.org>) generated oligonucleotides that allowed site-directed mutagenesis to be performed on the pRS300 plasmid, which contains a copy of the endogenous *Arabidopsis* miRNA miR319a (At4g23713) (Palatnik et al., 2003). Overlapping PCR was used to generate the amiRNA precursor and the final PCR product was cut with *Bam*HI and *Eco*RI, and ligated to a vector (pBJ36) containing either the CaMV 35S promoter or the *AlcA* ethanol-inducible promoter. After verifying the sequence of the amiRNA, a fragment containing the promoter and the amiRNA was sub-cloned into either of the binary vectors pML-BART or pART-27 using *Not*I sites. This vector was then used to transform plants as described in Section 2.3.1.

2.7 Quantitative real-time PCR

Primers were designed to have a T_m of $60 \pm 1^\circ\text{C}$. Each primer pair did not differ in annealing temperature of more than 1°C . Additionally, primer pairs to amplify cDNA were designed to span exon-exon boundaries whenever possible and were as close to the 3' end of the coding sequence as possible. The annealing temperatures and primer dimerization potentials of these oligonucleotide primers were checked using the thermodynamic oligonucleotide software, Oligo Analyzer, available at the Integrated DNA Technologies website (<http://eu.idtdna.com>). The Lightcycler 480 (Roche) with SYBR green master 1 (Roche) was used to quantify relative enrichments of DNA. One reaction mix contained 5 μL of 2x SYBR green master 1, 1 μL (c)DNA, 1 μL of 10 μM primers and 3 μL of molecular biology grade water. An equivalent time of 60 s per 1 kb of DNA was given to generate the amplicons. Annealing temperatures between 58°C and 60°C were used depending on the primer pairs. The enrichment observed was calculated relative to the input sample. Therefore, for each locus analyzed, the enrichment was calculated as follows:

$$\begin{aligned}\text{For ChIP: } & (2^{\text{Cp}^{(\text{IN})} - \text{Cp}^{(\text{IP})}}) \times \text{VIN} \\ \text{For cDNA: } & (2^{\text{Cp}^{(\text{REF})} - \text{Cp}^{(\text{GENE})}})\end{aligned}$$

where Cp is the crossing point calculated by the Lightcycler 480, Cp(IN) is the Cp value for the input, Cp(IP) is the Cp value for the IP and V(IN) is percentage of the total volume of the IP used to generate the input sample. Cp(REF) is the Cp value for the chosen reference gene and Cp(GENE) is the Cp value for the chosen target gene. The reference gene used for qRT-PCR was DM-97/DM-98 (At4g34270), which was chosen from (Czechowski et al., 2005).

2.8 Thermal Asymmetric Interlaced PCR

TAIL PCRs were performed as described in (Liu et al., 1995). PCR products were analysed on a 2% agarose gel and the desired DNA bands were extracted, purified and sequenced using the primer DM-73. BLAST analyses (www.arabidopsis.org) were used to

identify the genomic position of the sequenced DNA. Primers were then designed in order to distinguish between the wild-type locus and presence of the T-DNA at the same locus.

2.9 Western blotting

2.9.1 Preparation of nuclear extracts for SDS-PAGE electrophoresis

The extraction of nuclei was performed as described in Section 2.10.1 with the exception that formaldehyde and glycine were not added and that after washing the nuclei with the M3 buffer (Appendix 2), they were resuspended in 50 μ L 2x SDS loading buffer (see Appendix 2 for the composition of all buffers used for SDS-PAGE electrophoresis and Western blotting) and incubated at 95°C for 5 min. 15 μ L of this denatured protein extract was loaded onto a 10% SDS-PAGE gel. One of the main differences between these protocols is that the tissue is treated with formaldehyde after cell disruption in Protocol 1, whereas intact whole tissue samples are treated with formaldehyde in Protocol 2 (*i.e.* before cell disruption).

2.9.2 Transfer of proteins from an SDS-PAGE gel to a membrane

Proteins were transferred at 50 V for 2 hours (4°C) from the SDS-PAGE gel to an Immobilon-P membrane (Sigma-Aldrich) (Appendix 2). The membrane was then incubated in a blocking solution for 15 min at room temperature. A 2,000-fold dilution of the Ab290 antibody (Abcam) in blocking solution was then incubated overnight at 4°C with the membrane. Non-hybridized antibody was removed by rinsing the membrane with 1xPBS supplemented with 0.05% (v/v) Tween-20 (noted PBS-T) for 5 min. This wash was repeated twice more. The secondary antibody (α -rabbit conjugated to horse radish peroxidase (Sigma)) was diluted 50,000-fold in the blocking solution and was added to the membrane. The latter was then incubated with the secondary antibody solution for 4 h at room temperature with mild shaking. The excess of secondary antibody was rinsed by washing the membrane 3 times 5 min with PBS-T. The presence of GFP was detected using

the Amersham ECL Plus Western Blotting Detection Reagent (GE Healthcare) in the LAS-3000 Intelligent Dark Box Image Reader (Fujifilm).

2.10 Chromatin immunoprecipitation

Two different chromatin immunoprecipitation protocols were tested. Protocol 1 was adapted from protocols by Dr. Adrian Bracken (Cancer Transcriptomics Laboratory, Smurfit Institute of Genetics, TCD) and by Ito et al. (1997). Protocol 2 was performed as described in (Kaufmann et al., 2010a). The main difference between the two protocols is the time at which the fixation is performed. In protocol 1, fixation was performed after the tissue had been ground into a fine powder. In protocol 2, fixation was performed on whole plant tissue, prior to any treatments. There are other differences between the two protocols, including buffer compositions, however, the point at which the tissue was treated with formaldehyde appeared to result in the greatest differences between the results obtained.

2.10.1 Fixation and nuclei extraction

In protocol 1, approximately 500 μ L tissue was ground in liquid nitrogen. The resulting tissue powder was resuspended in 900 μ L M1 buffer (the composition of all buffers used in this section is detailed in Appendix 2) and the tissue was ground further. 24.3 μ L of a 37% (w/v) formaldehyde solution (Sigma-Aldrich; final concentration was 1%) was then added to the tissue extract to crosslink proteins to the DNA. This mixture was incubated for 10 min at 4°C with rotation. 50 μ L of a 2.5 M glycine solution was added to stop the cross-linking reaction, followed by incubation for 5 min at 4°C with rotation. In order to remove the debris, the cross-linked tissue was filtered twice through a layer of miracloth that had been pre-wet with M1 buffer. The filtered extract was then centrifuged for 5 min at 7,700xg at 4°C to pellet the nuclei. Next, the nucleus pellet was resuspended in 0.9 mL M2 buffer, followed by centrifugation for 1 min at 7,700xg and 4°C. This step was repeated twice more. The next wash step was performed by resuspending the nucleus pellet in 0.9 mL M3 buffer, followed by centrifugation for 5 min at 380xg, 4°C. The nucleus pellet was finally resuspended in 0.2 mL lysis buffer and incubated for 10 min on ice in order to

lyse the nuclei. After incubation, the nucleus extract was diluted with 0.8 mL ChIP dilution buffer.

In protocol 2, approximately 1 mL of tissue was fixed in 2 mL MC buffer with 1% formaldehyde on ice under vacuum (800 mbar). The vacuum was released after 15 min, the tissue was mixed and the vacuum was applied 3 more times for another 15 min each (total fixation time was 60 min). The formaldehyde fixation was stopped by adding 50 μ L of a 2.5 M glycine solution. The tube was inverted several times, placed back on ice and vacuum (800 mbar) was applied for 5 min. The excess solution was removed and the tissue was washed in 1 mL MC buffer. The excess buffer was removed and replaced with 1 mL MC buffer. This wash step was repeated once more, after which the excess MC buffer was removed carefully. The tissue was then snap frozen in liquid nitrogen.

In order to extract the nuclei, approximately 200 μ L M1 buffer was added to the tissue, which was then ground thoroughly. More M1 buffer was then added until a final volume of 1.25 mL was reached. In order to remove the debris, the cross-linked tissue was filtered twice through a layer of miracloth that had been pre-wet with M1 buffer. At this stage, nuclei were extracted using M2 and M3 buffers, as detailed in protocol 1.

2.10.2 Sonication

Chromatin was sonicated using a Diagenode Bioruptor at 4°C. The nucleus extract was split into 0.3 mL aliquots. “High” intensity sonication cycles were applied, which consisted of 30 sec sonication, followed by 30 sec without sonication. In protocol 1, 10 cycles were necessary, while in protocol 2, 12 cycles were required in order to shear the chromatin into fragments averaging between 350 and 750 bp. The sonication efficiency was checked for each assay by purifying a 10 μ L aliquot of chromatin using a phenol-chloroform extraction (see below) and analyzing it on a 1.5% agarose gel.

2.10.3 Immunoprecipitation

The protein A sepharose beads (GE Healthcare) were washed with ChIP dilution buffer (Protocol 1) or IP buffer (Protocol 2) and blocked with a 0.5 mg/mL lipid-free bovine

serum albumine solution. Next, the chromatin was pre-cleared by incubating it with 60 μ L of washed protein A sepharose beads (15 min in Protocol 1 and 90 min in Protocol 2). A 100- μ L aliquot of this pre-cleared chromatin was taken as the input control sample. The remaining of the pre-cleared chromatin was pipetted out into a clean microtube and incubated with 4 μ L of GFP-specific antibody (Ab290; Abcam) at 4°C for 1 h, with rotation. After incubation with the antibody, 60 μ L of washed protein A sepharose beads were added, in order to purify the antibody/chromatin complexes. This mixture was incubated with rotation for 2 h at 4°C. Next, the protein A sepharose beads were washed using different sets of buffers, depending on the protocol used.

In Protocol 1, the beads were washed 3 times with 1 mL Mixed Micelle Wash Buffer, twice with 1 mL buffer 500, twice with 1 mL LiCl buffer and once with 1 mL 1xTE buffer. All wash steps were performed at 4°C by adding the wash buffer to the beads, followed by centrifugation for 1 min at 2,400xg and 4°C. After centrifugation, the supernatant was removed and replaced with fresh wash buffer. The antibody/chromatin complexes were eluted from the beads by incubating the latter with elution buffer 1 at 65°C for 2 hours. The mixture was then spun 1 min at 2,400xg to pellet the beads. The supernatant was collected and incubated overnight at 65°C to reverse the protein-DNA cross-links.

In Protocol 2, the protein A sepharose beads were washed 5 times with 1 mL IP buffer at 4°C. For each of the washes, the beads were incubated with the IP buffer on ice with mild shaking for 8 min, followed by centrifugation for 1 min at 2,400 g, 4°C. After the last wash, the antibody/chromatin complexes were eluted by adding 100 μ L elution buffer 2 to the beads. This mixture was incubated at 37°C with vigorous shaking for 1 min. After centrifugation for 1 min at maximum speed, the supernatant was removed and 50 μ L of a 1 M Tris pH 9.0 solution was added. The elution was repeated twice more and the eluates were combined.

2.10.4 Purification of ChIP DNA

In Protocol 1, 2 μ L DNase-free RNase (Roche) was added to the protein/DNA mixture eluted from protein A sepharose beads, as well as to the input fraction. These samples were incubated at 37°C for 1 h. Next, 2 μ L proteinase K (Roche) was added, and

the mixture was incubated at 55°C for 2 h. This was followed by a phenol-chloroform extraction of the DNA. To this end, 1 volume of phenol:chloroform:isoamyl alcohol (25:24:1) (Fluka) was added to the protein/DNA mixture. The latter was then centrifuged in a phase-lock tube (5-prime) and the aqueous phase was transferred to a clean tube for DNA precipitation.

In protocol 2, the eluate (*i.e.* antibody/chromatin complexes) and input fractions were treated with 4 µL proteinase K (Roche) and incubated overnight at 37°C. This was followed by addition of 2 µL DNase-free RNase (Roche) and incubation for 30 min at 37°C. Finally, another 4 µL proteinase K was added, followed by incubation at 65°C for at least 6 h, in order to reverse the cross-links. The DNA was then extracted using phenol-chloroform extraction, as described for Protocol 1.

Finally, independently of the protocol used, the DNA was precipitated by adding 2.5 volumes of 100% ethanol, 0.1 volume of 0.1 M sodium acetate pH5.4 and 2 µL glycogen (from a stock solution at 20 mg/mL, Sigma), followed by incubation at -20°C for 30 min. After centrifugation at maximum speed for 15 min at 4°C, the DNA pellet was washed with 80% ethanol, dried and resuspended in molecular biology grade water.

2.10.5 Quantification of genomic DNA

To obtain an accurate estimate of the amount of genomic DNA present in ChIP samples, the Quant-iT dsDNA HS assay (Invitrogen) was used according to the manufacturer's instructions.

2.10.6 Illumina ChIP-Seq library preparation

The preparation of Illumina ChIP-Seq libraries was performed as per manufacturer's instruction with some modifications. Briefly, ChIP DNA (5 – 10 ng) was treated with *T4* DNA polymerase, *E. coli* DNA Pol I large fragment (Klenow polymerase), and *T4* polynucleotide kinase (PNK) to convert the overhangs of the ChIP DNA into phosphorylated blunt ends. Next, an adenine base was added to the 3' end of the blunt phosphorylated DNA fragments using the polymerase activity of Klenow fragment (3' to 5'

exo minus). Then the ends of the DNA fragments were ligated to adapters (a dilution of 1:40 of adapter oligo mix was used instead of 1:10). Excess adapters were then separated by agarose gel electrophoresis. To this end, a 2% agarose gel (Low Range Ultra Agarose, Bio-Rad) was run for over six hours at 20 V. DNA fragments ranging from 300 to 500 bp were excised and purified using Qiagen's Gel extraction kit. Then adapter-specific primers were used to amplify the size-selected DNA. The products from these PCR reactions were again separated on an agarose gel. DNA fragments ranging from 300 to 500 bp were excised and purified from the gel using Qiagen's Gel extraction kit. These libraries were then analyzed on an Agilent Bioanalyser using the High Sensitivity DNA chips. The amount of DNA was measured as described in Section 2.10.5. Finally, these libraries were diluted to a final concentration of 50 pg/ μ L for qPCR analysis, and to a final concentration of 10 nM for Illumina sequencing.

2.11 *In situ* hybridisation

Non-radioactive *in situ* hybridisations were performed as previously described (Long and Barton, 1998). The primers DM-31/68 (Appendix 1) were used to amplify the *GR* fragment using the vector *pAPI::API-GR*-pBJ36 as a template. The resulting PCR product was ligated into pJET 1.2 (Fermentas) using blunt ends and sequenced in order to determine the orientation of the insertion. This vector was linearised using the restriction enzyme *Cla*I and anti-sense RNA was synthesized using T7 RNA polymerase (Roche). The probe was hybridised to 8 μ M sections of tissue from *pAPI::API-GR* plants.

2.12 Data analysis and Bioinformatics

Unless otherwise stated, all analyses were performed using the R statistical software Version 2.13.0 and the listed libraries implemented in the Bioconductor project (www.bioconductor.org) Version 2.8.

2.12.1 Re-annotation of the Agilent 44K array to the latest genome release

A custom-made Agilent 44K array was used throughout this study for microarray profiling experiments (array design performed by Dr. Jose-Luis Riechmann, Centre for Research in Agricultural Genomics, Barcelona). Because probe design was based on the TAIR version 8 genome release, all probes were re-annotated on the array to TAIR version 10. Figure 2.1 summarizes the approach. The standalone BLAST executable function *blastall* Version 2.2.24 (downloadable from <ftp://ftp.ncbi.nih.gov/blast/executables/>) (Altschul et al., 1990) was used to map all probes to the latest cDNA definitions (file TAIR10_cdna_20101214); if it had one or multiple hits with less than two or less mismatches or no more than one gap-extension it was assigned the corresponding transcripts. If a probe exhibited multiple hits at these criteria, it was assigned to all transcripts and labelled to have possible cross-hybridisation properties.

If no cDNA hit was found, the probe was blasted against the transposable elements (TE) sequences (file TAIR9_TE.fasta). If a probe showed only a single perfect hit with no mismatches or gap-extensions, it was assigned to the corresponding TE, if it had additional or several good hits only (same criteria as in the cDNA BLAST), it was also labelled to have cross-hybridization properties. If a probe had no hits, it was further blasted against genomic sequences (file TAIR10_bac_con_20101028). If a probe had no hits with less than four mismatches or three gap-extensions, it was labelled as “negative probe” to be used later as a guidance for filtering based on expression levels. The remaining probes were labelled with their original identifiers, generally identification codes (IDs) containing the genomic coordinates.

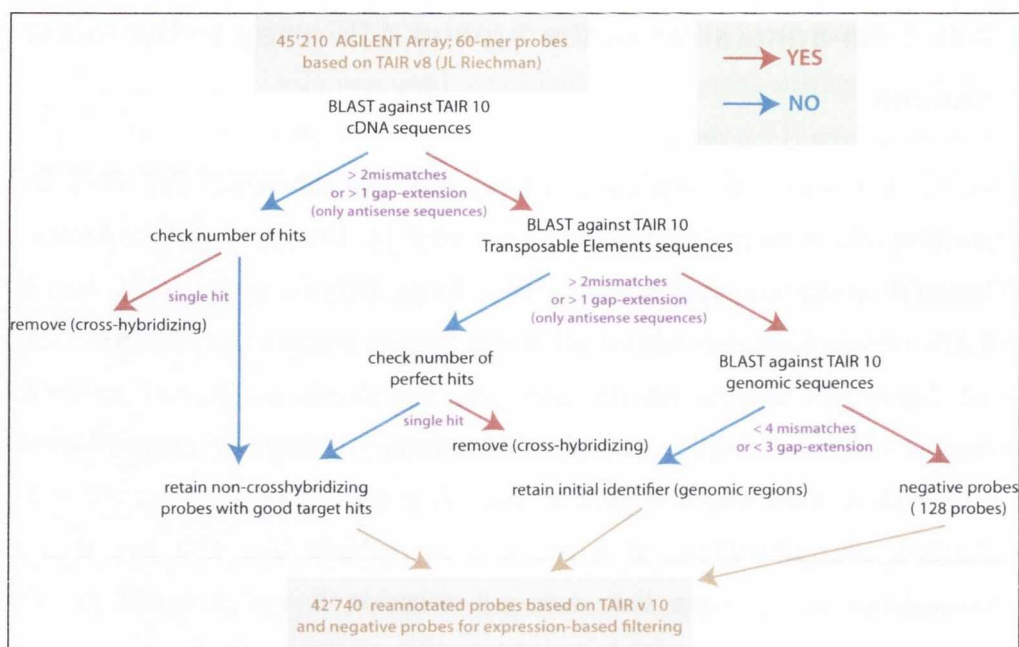


Figure 2.1 Reannotation of the Agilent 44K array to the latest genome release. Scheme summarizing the workflow used to re-annotate the 44K Agilent array - originally designed based on the TAIR version 8 genome release - to the TAIR 10 genome release using BLAST. The purple text outlines decision rules. The re-annotation was also used to remove potential cross-hybridizing probes. In case a probe could not be unambiguously assigned to a locus or to the pool of negative probes, its initial identifier was retained.

2.12.2 Low-level microarray data processing and basic data handling

Low-level data processing was performed using functionality provided by the Bioconductor package *limma* (Silver et al., 2009). Agilent median signals and background were read into R and background correction was performed using the *backgroundCorrect*-function (Ritchie et al., 2007) with maximum likelihood estimation of the background (Silver et al., 2009) and an offset of 50. Between-channel normalization was performed using loess-normalization, between-array normalization using quantile normalization (Smyth and Speed, 2003). Duplicated probes were averaged using the *avereps*-function provided in the *limma* package. Non-specific filtering was performed using the *genefilter*

package, based on a signal cutoff at the median of all negative probe signals, so that only those probes with higher signals in at least three arrays were retained.

Principal component analyses were performed using the *prcomp*-function in R. Testing for differential expression: Separate-channel analysis (Smyth, 2005) was performed using linear models (Smyth, 2004) as described in the *limma* users guide.

2.12.3 Higher level microarray analysis

Base-level annotations were retrieved from the Bioconductor homepage, (package *arabidopsis.db0* Version 2.5.0) and an annotation package for the custom-made 44K Agilent array was generated using *AnnotationDBI* package as described in the SQL Forge vignette. Gene ontology terms contained within the *Biological Process* class used for further analyses, and terms with less than 10 genes associated with them were removed. Parent-children relationship information as contained within the GO.db data package (Version 2.5.0) were used to remove redundancy between terms, and parent terms were removed if they had more than 95% percent gene-association overlaps with their daughter terms.

Pfam- and Gene family information was retrieved from the TAIR homepage (files *gene_families_sep_29_09_update.txt* and *TAIR10_all.domains*) and parsed in R to generate family-to-gene mappings. AGIs for the gene family records of B3 transcription factors were missing and included as described by Romanel et al. (2009). Only families with 10 or more associated genes were included in the analysis. Redundancies amongst protein families were removed by searching for protein families that had more than 95% overlaps in gene-associations with another family, and removing the smaller family from further analyses. The same procedure was used to remove redundant gene families.

Adjustments for multiple testing were performed using the *p.adjust*-function in R, using Benjamini & Hochberg adjustments (Benjamini and Hochberg, 1995). Gene set enrichment analysis was performed using the *mroast* function as implemented in the *limma* package (Wu et al., 2010) with 5000 rotations, and no adjustment of the *p*-values were performed (because the number of tests was limited by testing only terms that had been identified in the GO-term enrichment analysis before). Heatmaps were generated using the *gplots* package (Warnes et al.).

2.12.4 Short read alignment, quality control and general data handling for ChIP-Seq

Reference datasets were downloaded from the NCBI SRA database (www.ncbi.nlm.nih.gov/sra), and sequence information from the sra-files dumped into fastq-files using the *sratoolkit* (Version 2.0.1) software as described in the manual at www.ncbi.nlm.nih.gov/books/NBK47540. Quality control of the sequenced libraries was performed using the *FastQC* software (www.bioinformatics.bbsrc.ac.uk/projects/fastqc/).

Arabidopsis genomic sequences were downloaded from TAIR version 10 (www.arabidopsis.org - single chromosome fasta files TAIR10_ch1.fas- TAIR10_ch5.fas, including chloroplast and mitochondrial sequences). Short read alignments were performed using the *SOAP2* software (Li et al., 2009b) with default settings.

Data were read into R using the *ShortRead* package from the Bioconductor project, and functionality provided by various Bioconductor packages (including the *IRanges* and *GenomicRanges* packages) was used for various data handling operations, with some help from example scripts described in the online tutorial provided by Thomas Girke (<http://manuals.bioinformatics.ucr.edu/home/ht-seq>). Pattern matching was performed using the *Biostrings* package. Genomic sequences within R were retrieved from the R data package BSgenome.Athaliana.TAIR.TAIR9 (Version 1.3.18) that uses infrastructure from the *BSgenome* package to store genomic information (note that TAIR 9 and TAIR 10 genomic sequences are identical).

2.12.5 Peak calling

Peak calling was performed using the *CSAR* package (Muino et al., 2011b), and enriched regions were identified at a FDR cutoff of 0.001 (with > 500'000 permutations for each dataset). Only peak regions with stretches of >50 bp were retained. Furthermore, an independent peak calling method, as implemented in the *BayesPeak* (PP > 0.9) was used to confirm enriched regions (Cairns et al., 2011; Spyrou et al., 2009): *CSAR* peaks were retained only if they overlapped with a region detected by the *BayesPeak* method.

Peaks were associated with genes located in the vicinity using *CSAR* functionality, whereby a peak had to reside in a genomic region ranging from 3kb upstream of a gene's transcriptional start position to 1 kb downstream of its transcriptional end position. Gene range definitions were downloaded from TAIR (file TAIR10_GFF3_genes.gff). GO-term enrichments were performed as described above by using GO-mappings provided by the org.At.tair.db data package (version 2.5.0) downloaded from the Bioconductor homepage. Relative location of peaks to nearest gene features were determined using the *annotatePeakInBatch*-function from the *ChIPpeakAnno* package (Zhu et al., 2010). Visualization of genomic regions was performed using the *GenomeGraphs* package (Durinck et al., 2009).

2.12.6 Motif Detection

For the detection of motifs at putative binding sites, 200 bp core sequences centred around peak locations (as obtained from the *CSAR* analysis) from a dataset were written into a fasta-file and submitted to the browser-based *DREME* analysis suite (<http://meme.sdsc.edu/meme/>) (Bailey, 2011).

2.13 Transgenic line generation

2.13.1 Generation and expression of an amiRNA that targets the AG mRNA

A detailed description of how amiRNAs are designed, cloned and expressed in *Arabidopsis* is available in Section 2.6. The sequences of the primers used are available in Appendix 1. In brief, amiRNAs were amplified, cleaved with restriction enzymes and ligated to pBJ36 vectors that contained either the *CaMV 35S* (vDM-1) or the *AlcA* promoter (vDM-2) sequence. Once the sequences of these vectors were verified, the promoter-amiRNA cassette was excised from pBJ36 using the *NotI* restriction enzyme and ligated to pML-BART (vDM-3), which had been digested with *NotI* and treated with alkaline phosphatase. In order to generate ethanol inducible expression plant lines, a pML-BART

vector that contained the 35S::*AlcR* cassette was used for ligation (vDM-4). Plants were transformed as described in Section 2.3.1. Colony PCRs and genotyping assays were performed using primers DM-3, DM-7, DM-11, DM-15, DM-19, DM-23 and DM-27 for the different AG amiRNAs, in combination with DM-31.

2.13.2 Generation of an RNAi construct to disrupt AG function

The sense strand complementary to 300 bp of the *AG* coding region was amplified using the primers DM-43/DM-44 and inserted upstream of the intron within the pHANNIBAL (vDM-5) vector and downstream of the CaMV 35S promoter. The primers used incorporated *XhoI* and *EcoRI* sites into the amplified product. The anti-sense strand was amplified using primers DM-45/DM-46, which incorporated a *BamHI* and an *XbaI* site into the amplified product and subsequently was inserted downstream of the intron within pHANNIBAL (Wesley et al., 2001). Using the 2 *NotI* sites that flank the promoter and terminator sequences, this fragment was transferred into the binary vector pML-BART that had been digested with *NotI* and treated with alkaline phosphatase. Plants were transformed as described in Section 2.3.1. Colony PCRs and genotyping assays were performed using primers DM-43 to DM-46.

2.13.3 Generation of dominant negative versions of AG

Three different dominant negative fusions with AG were overexpressed in *Arabidopsis*. Two of these fusions were with the SRDX domain (SLDLDLELRGFA) from the plant transcription factor SUPERMAN. Two different linker regions were used to fuse the AG coding sequence with the SRDX domain (Linker 1: PGFEIDKLG and Linker 2: RGIRYPGFEIDKLG). The SRDX domain was translationally fused to the coding region of *AG* under the control of the CaMV 35S promoter. The primers DM-47/DM-48, which incorporated *XhoI* and *SmaI* sites into the amplified product, were used to amplify the *AG* coding region from wild-type (accession: L-er) flower cDNA. This PCR product was then ligated between the 35S promoter and the SRDX domain of a pBJ36 vector (vDM-6) using

XhoI and *SmaI* sites to create an in frame translational fusion containing linker 1 (see above).

The primers DM-63/DM-64, which incorporated *XhoI* and *PstI* sites into the PCR product, were used to amplify the *AG* coding region from wild-type (accession: L-*er*) flower cDNA. This PCR product was then ligated between the 35S promoter and the SRDX domain of a pBJ36 vector using *XhoI* and *PstI* sites to create an in frame translational fusion containing linker 2.

A third repressor domain, noted WUSB (HRTLPLFPMHGED) (Ikeda et al., 2009), was translationally fused to the coding sequence of *AG*. Primers DM-47/DM-48, which incorporated *XhoI* and *SmaI* sites into the PCR product, were used to amplify the *AG* coding region from wild-type (accession: L-*er*) flower cDNA. This PCR product was then ligated between the 35S promoter and the WUSB domain of a pBJ36 vector (vDM-7) using *XhoI* and *SmaI* sites to create an in frame translational fusion containing linker 1.

The restriction enzyme *NotI* was then used to excise each of these fragments from pBJ36 and ligate them to pML-BART (vDM-3) using complementary *NotI* sites. Plants were transformed as described in Section 2.3.1. Colony PCRs and genotyping were performed using primers DM-47/DM-99 (SRDX) or DM-47/DM-100 (WUSB).

2.13.4 Generation of lines carrying the *pAG::AG-6xHA* rescuing construct

Primers DM-37/DM-38 were designed that incorporated a *PstI* site 1.6 kb upstream of the translational start site of *AG*, and a *Clal* site at the very end of the final exon (excluding the stop codon). Due to recombination with the *E. coli* genome, there was difficulty in cloning this PCR fragment into a pBJ36 vector (vDM-8) encoding six copies of the haemagglutinin tag (denoted: 6xHA). In order to alleviate this problem, a different construct was made that encoded a genomic region starting 1.6 kb upstream of the initial ATG up to the end of the second intron, followed by the remaining cDNA sequence of *AG* (Yanofsky et al., 1990). The genomic region used has been shown to contain essential transcriptional regulatory elements and should therefore be sufficient to express *AG* in a pattern similar to that of the endogenous gene.

In order to generate this construct, 2 PCRs were carried out using primers DM-

38/DM-40 and genomic DNA as a template, and DM-37/DM-39 and flower cDNA as a template (accession: L-*er*). The 2 PCR products were then assembled together by overlapping PCR using primers DM-37/DM-38. The resulting PCR product was digested with *Cla*I and *Pst*I and cloned into a pBJ36 vector that encoded the 6xHA tag, which had been digested with the same enzymes. Despite the removal of most introns, recombination problems with the *E. coli* genome persisted after ligation with the pBJ36 6xHA vector. In order to bypass this problem, an *E. coli* strain (Tp611, (Glaser et al., 1993)) that reduces plasmid copy number and is deficient in homologous recombination (due to mutations in the *pcnB* and *RecBC* genes, respectively) was used for transformation after ligation of the PCR product with pBJ36 6xHA. pBJ36 encoding the *pAG::AG-6xHA* rescue construct could then be obtained and was controlled by sequencing. The DNA fragment encoding the rescue construct was then sub-cloned into the binary vector pML-BART (vDM-3) using *Not*I sites and ultimately transformed into plants as described in Section 2.3.1. Colony PCRs were performed using primers DM-39/DM-40.

To genotype *ag-1 pAG::AG-6xHA* plants for the rescue transgene and the *ag-1* allele, DNA was extracted from transformants as described in Section 2.2.6. The primers DM-39/DM-40 were used to verify the presence of the correct transgene. The presence of the *ag-1* allele was confirmed by using a derived cleaved amplified polymorphic sequence (dCAPS) (Neff et al., 1998). Briefly, a PCR was performed with primers DM-35/DM-36, which resulted in a product of 144 bp. The primer DM-35 carried mismatches that resulted in the incorporation of an *Afl*III site in the PCR product, if the *ag-1* mutation was present. After digestion of the PCR product with *Afl*III, the DNA was analyzed on a 3% (w/v) agarose gel. A fragment of 144 bp indicates a wild-type allele and a fragment of 114 bp indicates the presence of an *ag-1* allele.

2.13.5 Generation of lines carrying the *pAG::AG-mGFP5* rescuing construct

The entire *AG* genomic locus was amplified using primers DM-38/DM-41. The use of these primers incorporated a *Pst*I site 1.6 kb upstream of the translational start site and an *Eco*RI site at the end of the final exon of *AG* (excluding the stop codon). This PCR fragment was cloned into the vector pBJ36 (vDM-9) containing the *mGFP5* gene using the

E. coli strain Tp611 (Glaser et al., 1993). The fragment encoding the *pAG::AG-mGFP5* rescue construct was then sub-cloned into the binary vector pML-BART (vDM-3) using *NotI* sites and transformed into plants as described in Section 2.3.1. The presence of the correct transgene in plants resistant to BASTA was verified by extracting DNA as described in Section 2.2.6 and genotyping for the presence of mGFP5 using primers DM-31/DM-59.

The removal of the introns downstream of the second intron in the *pAG::AG-6xHA* rescue construct also removed the annealing site for primer DM-35, which is used together with primer DM-36 to genotype for the presence of the *ag-1* mutant allele. Therefore, no modification of the genotyping assay for *ag-1* needed to be made in transgenic plants encoding the *pAG::AG-6xHA* fusion. This was not the case when the *pAG::AG-mGFP5* transgene was introduced, as all introns were present in this rescue construct. Thus, to genotype the *ag-1* mutant in the presence of the *pAG::AG-mGFP5* rescue construct, a modified assay was developed. In this case, primers DM-35/DM-34 were used in order to amplify a fragment containing part of the 3'UTR, which is missing in the *pAG::AG-mGFP5* transgene. This PCR product was then digested with *NlaIII* and *AflIII*. A band of ~500 bp was present in all samples, whereas a band of ~220 bp indicated the presence of the *ag-1* mutation and a band of ~250 bp showed the presence of a wild-type allele.

2.13.6 Crossing *pAG::AG-mGFP5 ag-1* lines into the FIS

Pollen from a *35S::API-GR ap1-1 cal-1 ag-1/+* line was used to fertilize emasculated flowers from a *pAG::AG-mGFP5 ag-1* plant. F₁ plants were genotyped for the presence of the *ag-1* allele, as described above, and seeds were collected from plants homozygous for this allele. F₂ plants that exhibited an *ap1 cal* mutant phenotype were genotyped for the presence of the *35S::API-GR* transgene using primers DM-61/DM-62, as well as for the presence of the *pAG::AG-mGFP5* transgene using primers DM-31/DM-59. F₃ plants that were homozygous for all mutations and for the T-DNA encoding the *35S::API-GR* fusion were isolated (genotyped using primer pairs DM-61/DM-62 (absence of T-DNA) and DM-DM-60/61 (presence of T-DNA)). However, because no positional genotyping assay was available for the T-DNA encoding the *pAG::AG-mGFP5* fusion, a segregation analysis had to be carried out on the offspring of individual F₃ plants in order to isolate a line that was homozygous for the rescue transgene.

2.13.7 Introduction of the pAG::AG-mGFP5 fusion into an alternative FIS

As described in Chapter 5, an alternative FIS was generated, which is based on the introduction of a transgene encoding the AP1-GR fusion under the control of the endogenous *AP1* promoter (see Section 2.13.13) into an *ap1-1 cal-1* mutant background. This alternative FIS is denoted *pAP1-FIS* hereafter. Pollen from a 35S::AP1-GR pAG::AG-mGFP5 *ap1-1 cal-1 ag-1* line was used to fertilize flowers of a pAP1::AP1-GR *ap1-1 cal-1* plant (line #3). F₁ plants were genotyped for the presence of the pAP1::AP1-GR transgene using primers DM-70/DM-73. F₂ plants were genotyped in order to isolate plants homozygous for the *ag-1* mutant allele and for the pAP1::AP1-GR transgene (using primer pairs DM-69/DM-70 (genomic locus in the absence of T-DNA) and DM-70/DM-73 (presence of T-DNA at the expected genomic locus)). The absence of the 35S::AP1-GR transgene was also checked using the PCR reactions described above. Finally, the presence of the pAG::AG-mGFP5 transgene was checked using primers DM-59/DM-31, but because the position of this T-DNA in the genome was unknown, it was not possible to check whether plants were homozygous for this T-DNA. In order to isolate plants that were homozygous for the pAG::AG-mGFP5 rescue T-DNA, a segregation analysis was carried out on the offspring of individual plants of the F₂ population.

2.13.8 Isolation of pAlcA::AG4-amiRNA lines in the pAP1-FIS

Pollen from *pAP1-FIS* plants (line #5) was used to fertilise emasculated flowers of a plant encoding the pAlcA::AG4-amiRNA (line #52). F₁ plants were genotyped for the presence of the pAP1::AP1-GR transgene using primers DM-73/DM-74. F₂ plants were checked for the presence of an *ap1-1 cal-1* double mutant phenotype and treated with a dexamethasone-containing solution (see Section 2.5.2). Plants that responded to the treatment with synchronous flowering (i.e. plants that had at least one copy of the pAP::AP1-GR transgene) were then genotyped for the presence of the *AlcA::AG4-amiRNA* transgene using primers DM-15/DM-31. These plants were then genotyped for the homozygous presence of the pAP1::AP1-GR-containing T-DNA using primer pairs DM-74/75 (genomic

locus in the absence of T-DNA) and DM-73/DM-74 (presence of T-DNA at the expected genomic locus). Because the location of the T-DNA encoding the *AlcA::AG4-amiRNA* construct was unknown, a segregation analysis was carried out in the offspring of plants from the F₂ population to isolate plants that were homozygous for this T-DNA.

2.13.9 Generation and expression of an amiRNA that targets the *KNU* mRNA

A detailed description of how amiRNAs were designed, cloned and expressed in *Arabidopsis* is available in Section 2.6. The sequences of the primers used to generate the *KNU-amiRNA* are available in Appendix 1 (DM-55-58). In brief, the amiRNA was amplified as described in Section 2.6 and the resulting PCR fragment was cleaved with *Bam*HI and *Eco*RI. The resulting DNA fragment was ligated into pBJ36-based vectors that contained either the *CaMV 35S* (vDM-1) or the *pAlcA* (vDM-2) promoter sequence which were also cleaved with *Bam*HI and *Eco*RI. Once the sequences of these vectors were verified, the *35S::KNU-amiRNA* cassette was excised from pBJ36 using the *Not*I restriction enzyme and ligated to the binary vector pML-BART (vDM-3) that had also been cleaved using *Not*I and treated with alkaline phosphatase (Roche). In order to generate lines for an ethanol-inducible expression of the *KNU-amiRNA*, a pML-BART vector that contained the *35S::AlcR* cassette was used for ligation in combination with the *AlcA::KNU-amiRNA* cassette. Plants were transformed as described in Section 2.3.1. First generation transformants were identified by their herbicide resistance and by using PCR-based genotyping. Colony PCRs and genotyping assays were performed using primer DM-57 in combination with primer DM-31.

2.13.10 Over-expression of the *KNU* coding sequence

In order to generate a line that constitutively over-expressed *KNU*, the *KNU* coding sequence was amplified from cDNA generated from an mRNA population, which was extracted from *L-er* wild-type flowers, using primers DM-86/DM-87 that incorporated *Pst*II and *Acc*65I restriction sites, respectively. The resulting PCR product was cleaved and

ligated to a pBJ36 vector containing the 35S promoter (vDM-1). Once the presence of the insertion had been verified by colony PCR, this vector was cleaved using *NotI* and ligated to the binary vector pML-BART that had also been cleaved with *NotI* and treated with alkaline phosphatase (Roche). The resulting vector was used for plant transformation. The sequence of the 35S::*KNU* cassette was verified by sequencing. First generation transformants were identified by their herbicide resistance and were genotyped by PCR. Colony PCRs and genotyping assays were performed using primers DM-86 and DM-31.

In order to generate a line that conditionally overexpressed *KNU*, the *KNU* coding sequence was amplified from cDNA generated from mRNA, which was extracted from *L-er* wild-type flowers, using primers DM-88/DM-89 that incorporated *PstI* and *EcoRI* restriction sites, respectively. This PCR product was cleaved and ligated into a pBJ36-based vector containing six copies of the haemagglutinin (denoted: 6xHA) tag. Once the presence of the insertion had been verified by colony PCR, this vector was cleaved using the enzymes *PstI* and *SpeI*. The *SpeI* site was within the 3'OCS terminator sequence. A pBJ36 vector containing the *AlcA* promoter sequence was also cleaved using *PstI* and *SpeI* and the *KNU*-6xHA fragment was ligated into it. This ligation reaction reconstituted the 3'OCS sequence. Next, this vector was cleaved using *NotI* and ligated into the binary vector pML-BART (vDM-3) that had also been cleaved with *NotI* and treated with alkaline phosphatase (Roche). The resulting vector was used for plant transformation. The sequence of the *AlcA*::*KNU*-6xHA cassette was verified by sequencing. First generation transformants were identified by their herbicide resistance and were genotyped by PCR. Colony PCRs and genotyping assays were performed using primers DM-33 and DM-89.

2.13.11 Generation of an epitope-tagged version of *KNU* resistant to the *KNU*-amiRNA

The primer combinations DM-90/DM-93 and DM-91/DM-92 were used to generate two PCR fragments that were subsequently used in an overlapping PCR with primers DM-90 and DM-91. The primers DM-92/DM-93 introduced synonymous nucleotide substitutions into the region of the *KNU* coding sequence that the *KNU*-amiRNA targets. The primers DM-90/91 introduced the restriction sites *PstI* and *EcoRI* that were ~2.6 kb

upstream of the translational start site and at the end of the *KNU* coding sequence, respectively. This PCR product was then ligated to a pBJ36-derived vector containing the *mGFP5* coding sequence (vDM-9) followed by the 3'OCS terminator sequence. Once the presence of the *KNU*-encoding fragment was confirmed by colony PCR, this vector was cleaved using *NotI* and ligated to the binary vector pART27 (vDM-10) that had been cleaved with *NotI* and treated with alkaline phosphatase (Roche). Plants containing the *35S::KNU-amiRNA* transgene (line #2) were transformed with this construct. First generation transformants were identified based on kanamycin resistance and were genotyped using PCR. Colony PCR and genotyping assays were performed using the primers DM-59 and DM-31.

2.13.12 Generation of a kanamycin-resistant *35S::AP1-GR* floral induction system

A pBJ36 vector containing the *35S::AP1-GR* transgene, which had been previously generated (Wellmer et al., 2006), was digested with *NotI* and ligated into the plant transformation vector pART27 (vDM-10). This construct was then transformed into *ap1-1 cal-1* plants as described in Section 2.3.1. The primers DM-53/DM-54 were used for colony PCRs and genotyping assays (see Appendix 1). All transformants were tested by placing a solution containing 10 μ M dexamethasone, 0.015% (v/v) Silwet L-77 in water on the inflorescence-like meristems and observing the phenotype several days later. In order to identify lines homozygous for the transgene, T3 plants were tested by segregation analysis; by genotyping for the presence of the transgene; and by testing their response to dexamethasone treatment.

2.13.13 A floral induction system based on the genomic locus of *AP1*

The last exon of the genomic *AP1* locus was translationally fused to a portion of the hormone-binding domain of the rat glucocorticoid receptor (GR) using the primers DM-51/DM-52 (see Appendix 1), which incorporated an *XhoI* site 3.2 kb upstream of the

translational start site and an *Xba*I site at the end of the last exon of *API*. The PCR fragment was ligated into a pBJ36 vector containing the GR coding sequence (vDM-11). Finally, *Not*I sites were used to transfer this fusion to the plant transformation vectors pART27 (vDM-10) and pML-BART (vDM-3), respectively. Plants were transformed and transformants were selected for as described in Section 2.3. The primers DM-49/DM-50 were used to perform colony PCRs and the primers DM-31/DM-68 genotyping assays (see Appendix 1). The inflorescences of transformants were treated with 10 μ M dexamethasone and 0.015% (v/v) Silwet L-77 and the resulting phenotypes were monitored over several days.

2.13.14 A floral induction system based on the over-expression of *CAL*

The coding region of *CAL* was amplified from Col-0 cDNA using the primers DM-80/DM-81 (Appendix 1) that incorporated *Bam*HI sites at both the 5' and 3' ends of the coding region. After restriction digestion, this PCR fragment was ligated with pGREEN-0229-GR (vDM-12) that had also been cleaved with *Bam*HI. The presence of the insertion and correct orientation were confirmed by restriction digestion. This vector was made available to Dr. Jiao Jin (Riechmann group, Catalan Institution for Research and Advanced Studies, Barcelona) for plant transformation.

2.13.15 A floral induction system based on the genomic locus of *CAL*

The final exon of the genomic *CAL* locus was translationally fused to a portion of the hormone-binding domain of the rat glucocorticoid receptor using the primers DM-82/DM-83 (see Appendix 1), which incorporated an *Mlu*I site 1.7 kb upstream of the translational start site and an *Xho*I site at the end of the last exon of *CAL*. The PCR fragment was ligated into a pBJ36 vector containing the GR coding sequence (vDM-11). Finally, *Not*I sites were used to transfer this fusion to the plant transformation vector pML-BART (vDM-3). This vector was used by Dr. Jiao Jin (Riechmann group, Catalan Institution for Research and Advanced Studies, Barcelona) for plant transformation.

2.13.16 Identification of the T-DNA insertion sites in plants harbouring the *pAP1::AP1-GR* transgene

The genomic DNA of two independent transgenic lines, pAP1-3 and pAP1-5, was used in a Thermal Asymmetric Interlaced (TAIL)-PCR reaction (see Section 2.8). The resulting PCR products were analysed on agarose gels containing ethidium bromide and DNA bands were excised, purified and sequenced. The pAP1-5 transgene identified is positioned within the third exon of At4g24204. Based on the location of the T-DNA, the primers DM-74/DM-75 can be used to determine the absence of the T-DNA, while the primers DM-73/74 can be used to determine the presence of the T-DNA. The pAP1-3 transgene is situated in the intergenic region between At3g59050 and At3g59040. The primers DM-69/DM-70 can be used to determine the absence of the T-DNA, while the primers DM-70/DM-73 can be used to test for the presence of the T-DNA.

2.13.17 Generation of plants expressing the AG4-miRNA in specific developmental domains

The sequence for the *AG4-miRNA* was excised from the *pAlcA::AG4-miRNA*-pBJ36 vector (see Chapter 3) using the restriction enzymes *EcoRI* and *BamHI*. This DNA fragment was ligated to the *6xOP*-pBJ36 vector (vDM-13) that had also been cleaved with *EcoRI* and *BamHI*. The *E. coli* strain STBL2 (Invitrogen) was transformed with the ligation mixture (see Section 2.2.4). Colony PCRs were performed using the primers DM-15/31 to verify the presence of the *6xOP::AG4-miRNA* transgene. Once vectors containing the miRNAs were purified from the transformed *E. coli*, they were digested using the restriction enzyme *NotI*. The resulting DNA fragment was ligated into pART27 (vDM-10), which had also been cleaved with *NotI* and treated with alkaline phosphatase. This vector was transformed into *Agrobacterium tumefaciens*. The *Agrobacterium* strain containing the *6xOP::AG4-miRNA* transgene were then used to transform *pCLV1::LhG4* (Schoof et al., 2000), *pCLV3::LhG4* (Lenhard and Laux, 2003) and *pWUS::LhG4* (Gross-Hardt et al., 2002) driver lines using the floral dip method (see Section 2.3.1). First generation

transformants were selected on MS agar plates containing kanamycin as described in Section 2.3.2 and were genotyped using the primers DM-15/31.

2.13.18 Generation of plants expressing the KNU-amiRNA in specific developmental domains

The generation of plants containing the *6xOP::KNU-amiRNA* transgene was performed in the same manner as described in Section 2.13.17, except that the *pAlcA::KNU-amiRNA*-pBJ36 vector was used to excise the *KNU-amiRNA* sequence and the primers DM-57/DM-31 were used for colony PCRs and genotyping assays.

Chapter 3 – Characterization of the transcriptional network controlled by *AGAMOUS*

3.1. Introduction

As outlined in Chapter 1, *AGAMOUS* (*AG*) is essential for the development of reproductive organs in *Arabidopsis* (Bowman et al., 1989) (see Chapter 1 for a detailed discussion of the known functions of *AG*). The aim of this project was to identify the genes directly and indirectly regulated by *AG* on a genome-wide scale using microarray technology and chromatin immunoprecipitation combined with ultrahigh-throughput sequencing (see Section 1.6). To this aim, several transgenic lines were generated including an ethanol-inducible system to perturb *AG* activity (Sections 3.2.1 – 3.2.2) and an epitope-tagged version of *AG* that can restore reproductive development in an *ag-1* mutant (Section 3.2.3). These transgenic lines were then crossed into an improved version of the floral induction system (see Chapter 5) in order to investigate early stages of flower development. Additionally, the *ag-1* mutation was isolated in the floral induction system (Section 3.2.4). The generation of these lines is described below in addition to the expression profiling (Sections 3.2.4 and 3.2.5) and ChIP-Seq analyses (Sections 3.2.6 – 3.2.7) performed.

3.2 Results

3.2.1 Perturbation of *AG* function

As outlined above (Section 1.5), it is possible to identify genes that are, directly or indirectly, dependent on *AG* function by perturbing its activity followed by a global analysis of gene expression. In order to identify the most efficient method of perturbing *AG* activity, three different strategies were tested: amiRNAs and RNAi constructs complementary to the *AG* mRNA were designed and translational fusions of *AG* with transcriptional repressor motifs were generated. The results for these approaches are presented below.

3.2.1.1 Identifying an amiRNA that efficiently targets AG

Prior to designing amiRNAs against *AG*, an amiRNA targeting the *AP3* mRNA was designed and overexpressed in wild-type plants by Adrian Bonner, an undergraduate student in the Wellmer laboratory. The overexpression of this amiRNA recapitulated the *ap3-3* null allele phenotype (data not shown). Given this result and the reported high success rate of the amiRNA approach in *Arabidopsis* (Schwab *et al.* 2006 and Ossowski *et al.* 2008), a single amiRNA (AG1) against *AG* was initially designed using the Web MicroRNA Designer (WMD, <http://wmd3.weigelworld.org>). This amiRNA was expressed in wild-type plants under the control of the constitutive Cauliflower Mosaic Virus (CaMV) 35S promoter as illustrated in Figure 3.1A and described in Sections 2.6 and 2.13.1. None of the 24 independent transformants obtained had phenotypes similar to those of an *ag* mutant (Figure 3.1B and Table 3.1), suggesting that this amiRNA was non-functional. Therefore, two additional amiRNAs (designated AG2 and AG3) were generated and expressed in wild-type plants under the control of the CaMV 35S promoter. When designing these amiRNAs, it was noted that the prediction software had been updated, and no longer predicted the AG1 amiRNA as a suitable candidate amiRNA to perturb *AG* activity. Transgenic plants expressing the AG2 and AG3 amiRNAs did not display any deficiencies indicative of a significant perturbation of *AG* activity. While most AG2 transformants displayed some mild carpel and stamen defects (Figure 3.1C), all AG3 transformants appeared to be phenotypically normal (Table 3.1).

It was noted that the three AG-amiRNAs tested targeted the 3' untranslated region (3'UTR) of *AG* (Figure 3.1H). Given that the functional amiRNA against *AP3* mentioned above as well as most endogenous miRNAs identified in *Arabidopsis* target coding regions (Ehrenreich and Purugganan, 2008), four more amiRNAs were designed to tile the *AG* coding sequence (Figure 3.1H). Two of these amiRNAs, AG4 and AG5, were predicted by the WMD software as being optimal for a gene knockdown (*i.e.* they were marked as “green”; see Section 2.6 for a description of the WMD software output). The other two, AG6 and AG7, were judged as possibly sub-optimal (marked as “yellow” and “orange”, respectively). Although these amiRNAs should still be functional, they have a potential to produce off-target effects.

Expression of *AG5*, *AG6* and *AG7* under the control of the CaMV 35S promoter in wild-type plants resulted only in mild carpel and stamen defects (Figure 3.1E – G). In

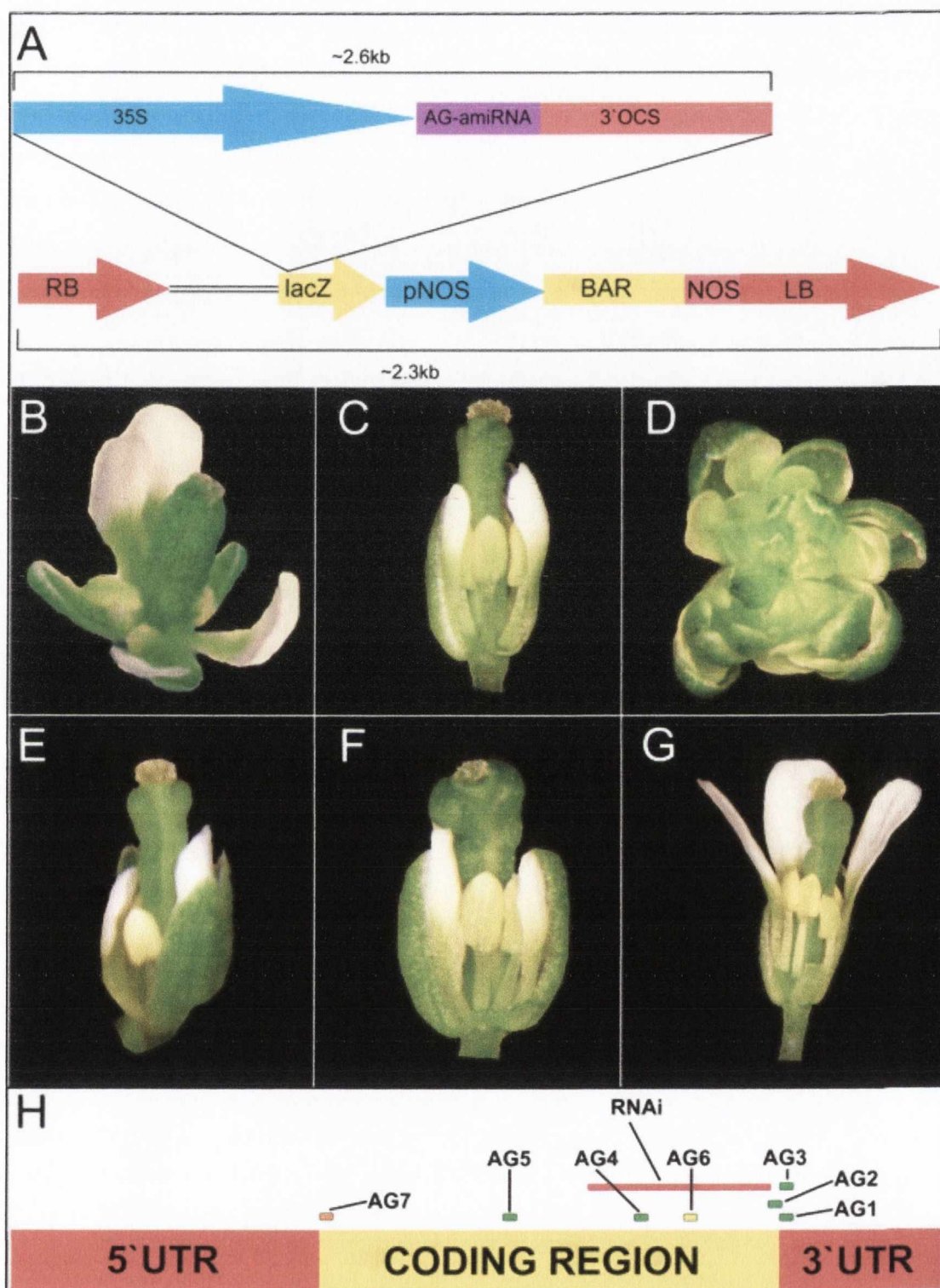


Figure 3.1 Overexpression of amiRNAs targeting the *AG* mRNA. (A) A schematic of the T-DNA insertion mediating expression of the amiRNAs under the control of the 35S *CaMV* promoter. The binary vector pML-BART was used in conjunction with *Agrobacterium tumefaciens* to transform wild-type plants. For full details, see section 3.2.1. (B to G) Floral phenotypes of plants constitutively over-expressing different

contrast, the remaining AG4 amiRNA resulted in *ag* mutant-like phenotypes when ectopically expressed in wild-type plants (Figure 3.1D). Of 21 primary transformants, one produced flowers that were almost indistinguishable from those of an *ag-1* mutant, while the flowers of several other lines showed partial stamen and carpel transformations as well as determinacy defects or misshaped carpels with immature stamens (see Table 3.1). The petals in the third whorl were slightly green and smaller than second whorl petals, however, no stamenoid characteristics were observed. Moreover, no carpelloid tissue was seen in any of the whorls present.

The *ag* mutant-like phenotypes observed in lines overexpressing the *AG4-amiRNA* strongly suggested that *AG* gene function was being efficiently disrupted. This conjecture was further tested by placing the *AG4-amiRNA* sequence under the control of the *35S::AlcR/pAlcA* ethanol-inducible promoter system (Figure 1.7) and by monitoring *AG* mRNA levels in response to an activation of amiRNA expression. The results of these experiments are presented in Section 3.2.2 below.

3.2.1.2 Perturbation of AG activity using RNA interference

In addition to the amiRNAs described above, an RNA interference (RNAi) gene perturbation approach was tested. To this end, a construct was designed which allowed the expression of a double-stranded RNA molecule that was homologous to a portion of the *AG* mRNA. In order to minimize the chance of creating unwanted off-target effects, the *AG* cDNA sequence was subjected to BLAST analyses. A 300 bp segment was identified that

Figure 3.1 continued. amiRNAs against *AG* in a wild-type background. **(B)** AG1; **(C)** AG2; **(D)** AG4; **(E)** AG5; **(F)** AG6; **(G)** AG7. **(H)** A schematic representation of the *AG* mRNA showing the predicted complementary binding sites of the different amiRNAs (as indicated) and the RNAi construct used in this study. The relative positions of the amiRNAs and the RNAi construct designed to target the *AG* mRNA are shown. The colours of the amiRNAs (AG1-7) correspond to the ranking system of the amiRNA design software (see Section 2.6 for further information). In brief, amiRNAs coloured in green are predicted to perturb target gene activity specifically and efficiently, while the other classes of amiRNAs (labelled in orange and yellow, respectively) may be sub-optimal. 3'OCS, octopine synthase terminator sequence; RB, right border of T-DNA; LB, left border of T-DNA. lacZ, a modified version of the β -galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA® resistance gene; NOS, nopaline synthase terminator sequence.

Table 3.1 Summary of amiRNAs designed to perturb *AG* activity. The amiRNA sequence, class (see Section 1.5.2.1 for a detailed description) and position of the target sequence of each amiRNA relative to the *AG* start codon are indicated. In addition, the total number of transformants obtained and the phenotypic effects of an constitutive over-expression of the different amiRNAs in wild-type plants are listed.

amiRNA	Distance from start codon (bp)	Class	Sequence of amiRNA	No. Transformants and phenotypic effects				
				Total	Strong	Intermediate	Weak	None
AG1	+771	Green	TTT TTA TTC ACT CCG GGC CGT	24	0	0	9	15
AG2	+736	Green	TTT ACA CTA ACT GGA GAG CGC	20	0	0	12	8
AG3	+771	Green	TTT TTA TTC ACT CCC GGC GAT	28	0	0	0	28
AG4	+516	Green	TTT ATC GTT ATG CAA ATC AGC	21	1	15	5	0
AG5	+300	Green	TTG TTG ATA ATA CTG TGC GTT	43	0	0	19	24
AG6	+597	Yellow	TTC GTA GTT AGA TCC TGC CGG	51	0	0	45	6
AG7	0	Orange	TTA GCT CCG ATT GGT AGG CCG	13	0	0	13	0

had little sequence similarity to other annotated *Arabidopsis* transcripts. The sense and anti-sense strands of this region were cloned under the control of the CaMV 35S promoter (see Section 2.13.2 and Figure 3.2A) and expressed in wild-type plants.

Fifty-two independent transformants were isolated and screened for the presence of an *ag*-like phenotype. Forty-eight of these primary transformants displayed a wild-type phenotype. Of the remaining four lines, one produced flowers with mildly deformed carpels and fully mature stamens (Figure 3.2D), while the flowers of two transformants exhibited stamens that were partially converted to petals as well as enlarged carpels (Figure 3.2B). One additional transformant showed a stronger floral phenotype that was characterized by a conversion of stamens to petals and a loss of floral determinacy (Figure 3.2C). However, fused carpels remained in the centre of these flowers, and stamens had not been fully transformed into petals as observed in *ag-1* mutant flowers. Although these mutant phenotypes suggested some degree of effect on *AG* function, the penetrance of the RNAi transgene was low (only 4 transformants out of 52 displayed any mutant phenotype) when compared to the AG4 amiRNA (see Table 3.1). In addition, overexpression of the AG4 amiRNA transformants resulted in phenotypes that more closely resembled those of *ag* mutants than in any of the RNAi lines obtained. Therefore, the amiRNA appeared to be more suitable for a specific disruption of *AG* activity than the RNAi approach. Consequently, no inducible RNAi system was generated.

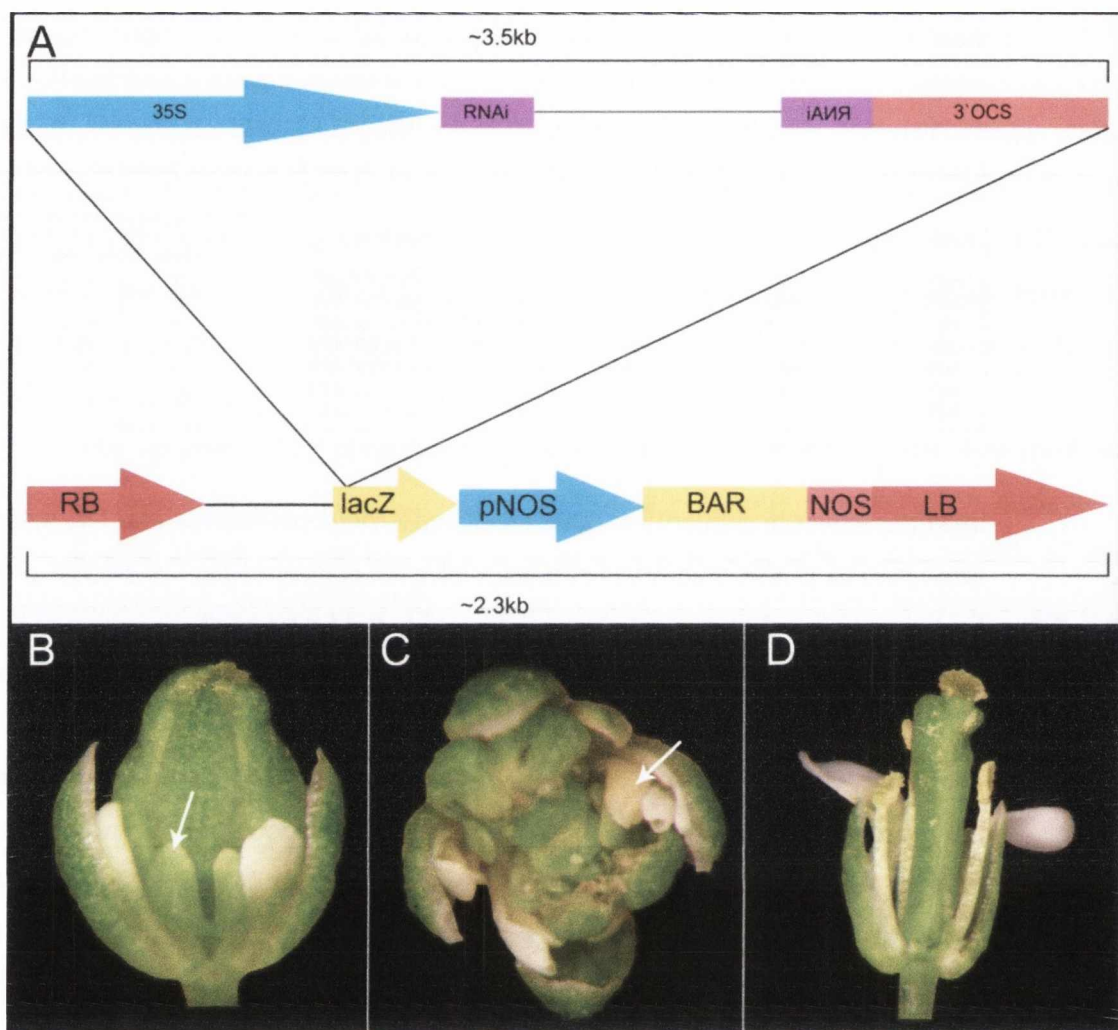


Figure 3.2 Perturbation of *AG* activity by RNAi. (A) A schematic of the T-DNA insertion mediating expression of an AG-RNAi construct under the control of the 35S *CaMV* promoter. The binary vector pML-BART was used in conjunction with *Agrobacterium tumefaciens* to transform wild-type plants. For full details see section 2.13.2. (B) to (C) Floral phenotypes of different AG-RNAi lines. (B) Flower of an AG-RNAi plant with a bulged gynoecium and small green stamens (arrow). (C) Flower of an AG-RNAi plant with petaloid stamens (arrow) and an indeterminate phenotype. (D) Flower of an AG-RNAi plant with mildly affected carpels. The fertility of this plant did not appear to be impaired. 3'OCS, octopine synthase terminator sequence; RB, right border of T-DNA; LB, left border of T-DNA. lacZ, a modified version of the β -galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA® resistance gene; NOS, nopaline synthase terminator sequence.

3.2.1.3 Perturbation of AG function by fusions with dominant repressor domains

As discussed in Section 1.5.2.3, chimeric transcriptional repressors can be generated by translationally fusing a repression domain to a transcription factor. Two different repressor motifs have been described in recent years that are suitable for a chimeric repressor approach. They are the WUSCHEL-box (WUSB) domain (Ikeda *et al.*, 2009) from the *Arabidopsis* transcription factor *WUSCHEL* and the SRDX domain, which is a modified version of EAR-like repression motif from the *Arabidopsis* transcription factor *SUPERMAN* (Hiratsu *et al.*, 2003). To test the suitability of these repressor domains for a disruption of AG function, the DNA sequences encoding both of these motifs were translationally fused to the C-terminus of AG and the resulting fusion proteins were expressed under the control of the constitutive 35S promoter in wild-type plants (see Section 3.2.3). Fifteen primary transformants were obtained that contained the 35S::AG-SRDX transgene. All of these plants were phenotypically abnormal but their flowers did not resemble those of *ag* mutant plants. Three lines produced deformed flowers with immature sepals or petals, and stamens were underdeveloped (Figure 3.3C). Carpels were present, however, they developed abnormally. A further two transformants formed undifferentiated tissue at the inflorescence meristem (Figure 3.3D), except for the appearance of malformed carpels. In a further four plants all floral organs were affected, however, petal and sepals were less affected than the previously discussed transformants (Figure 3.3B). The remaining six transformants displayed very mild phenotypic defects in all four whorls (not shown). These results strongly suggest that the AG-SRDX fusion protein was augmenting the expression of genes that are not usually regulated by AG. This finding was unexpected given that this system has been reported to work very effectively for AG and other transcription factors (see Section 3.2.8.1 for Discussion). Given the lack of a clear *ag-1* phenotype this fusion construct was not placed under the control of an inducible promoter.

Next, the WUSB domain was translationally fused to the AG coding sequence and overexpressed in plants as described in Section 2.13.3. Sixteen independent transformants were obtained that contained the 35S::AG-WUSB transgene. Eight of them formed flowers that showed some features of *ag* mutant flowers (Figure 3.3F). These flowers were indeterminate and contained petaloid stamens (Figure 3.3F, arrowhead) and malformed carpels. All flowers observed contained unfused carpels. The cauline leaves of these plants

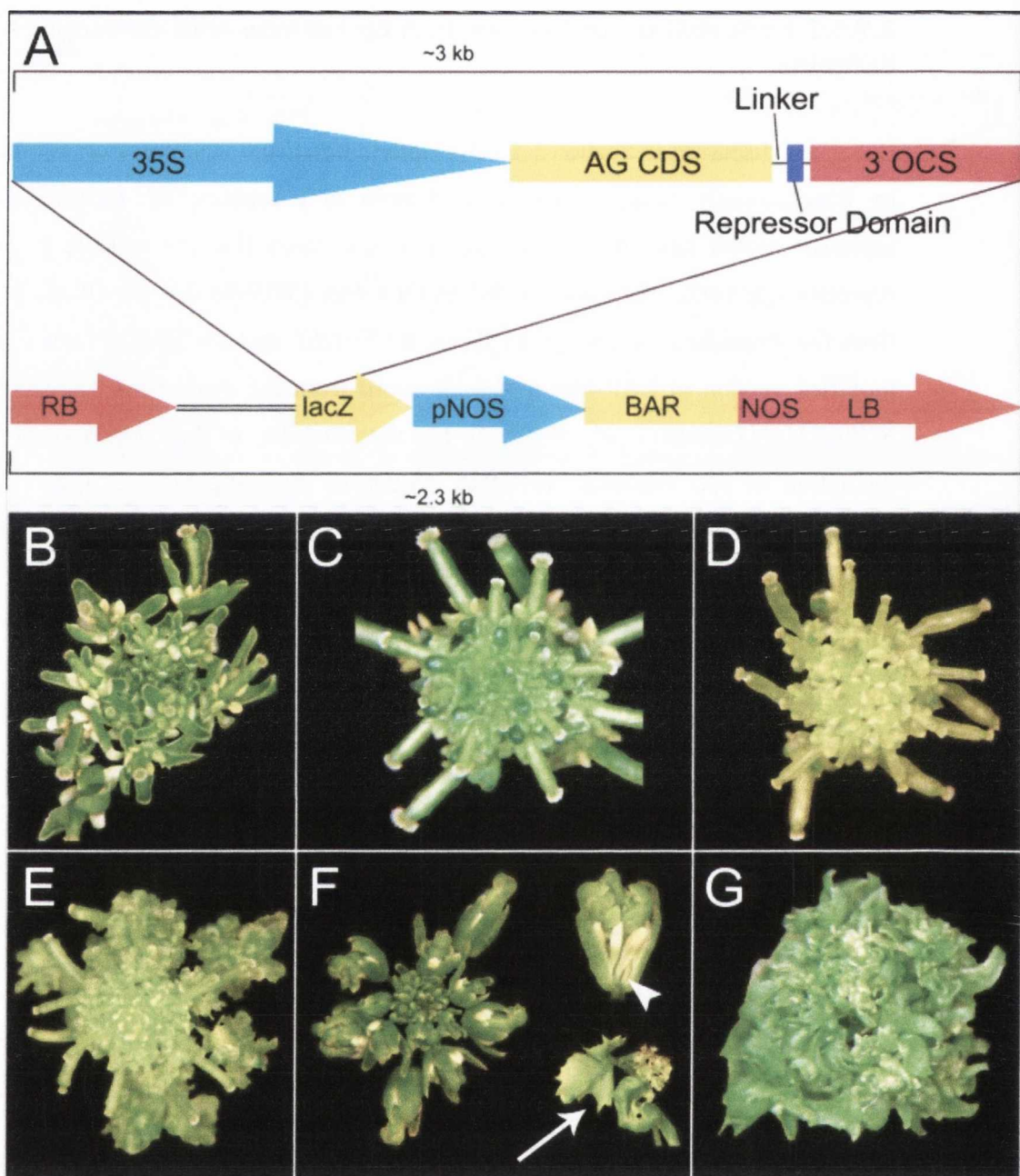


Figure 3.3 Effects of AG-SRDX and AG-WUSB fusion protein expression. (A) A schematic of the T-DNA insertion mediating expression of the fusion proteins under the control of the 35S *CaMV* promoter. The binary vector pML-BART was used in conjunction with *Agrobacterium tumefaciens* to transform wild-type plants. For full details see section 2.13.3. (B – D) Phenotypes of independent primary transformants carrying the 35S::AG-SRDX construct with linker #1. (B) Inflorescence of a plant containing flowers with abnormal sepals, petals and stamens. (C) Inflorescence of a plant containing flowers lacking sepals and petals. Stamens were under-developed and did not produce pollen. (D) Inflorescence of a plant containing flowers without identifiable sepals, petals and stamens. (E – G) Phenotypes of independent primary transformants carrying the 35S::AG-WUSB transgene with linker #1.

were often serrated (Figure 3.3, arrow) although the rosette leaves were unaffected. It is known that overexpression of *AG* in leaves causes developmental defects, however, this phenotype is not equivalent to what has been described in the literature (Mizukami and Ma, 1997). Seven other transformants resembled *35S::AG-SRDX* transformants (Figure 3.3E), which suggests that these proteins may affect plant development in a similar manner. Finally, one independent transformant developed leaf-like structures on the floral meristem instead of flowers (Figure 3.3G). This plant eventually produced *ag*-like flowers (not shown) similar to those presented in Figure 3.3F. To summarise, some plants containing the *35S::AG-WUSB* transgene presented *ag*-like phenotypes, which indicated that the chimeric repressor had repressed genes normally regulated by the endogenous AG protein. However, the presence of the phenotypes depicted in Figures 3.3E and 3.3G suggested that, like the *35S::AG-SRDX* transgene, this fusion protein was augmenting the expression of genes that are not usually regulated by AG.

The overexpression of both the AG-WUSB and -SRDX fusion proteins in wild-type plants produced transgenic plants with similar inflorescence-meristem defects (Figure 3.3C, D, E, G). This may be because the activity of these proteins had been affected in a similar way. I hypothesized that fusing the repressor domains to AG may have resulted in abnormal folding of the chimeric proteins or that the repressor domains may have interfered with the normal protein or DNA interactions endogenous to AG. It was reasoned that a longer peptide sequence linking the AG protein to the repressor domains may prevent any unwanted interactions or folding problems. Therefore, an alternative sequence of DNA was placed between the *AG* coding sequence and the SRDX domain sequence. This sequence was longer than the sequence previously used and resulted in an extra six amino acids being placed between AG and the SRDX domain (see Section 2.13.3 for details). This translational fusion was overexpressed in plants and nine transformants were obtained. No

Figure 3.3 continued. (E) Inflorescence of a plant containing flowers lacking sepals and petals. Furthermore, the morphology of stamens and carpels was abnormal. (F) An inflorescence containing flowers with determinacy defects and partially transformed stamens (arrowhead). A cauline leaf with serrated margins of the same plant is shown as well (arrow). (G) An inflorescence producing leaf-like structures. Floral organs eventually developed (not shown). 3'OCS, octopine synthase terminator sequence; RB, right border of T-DNA; LB, left border of T-DNA. lacZ, a modified version of the β -galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA® resistance gene; NOS, nopaline synthase terminator sequence

abnormal phenotypes were observed (not shown), however, expression levels of the transgene have yet to be examined.

3.2.2 Generation and characterization of a transgenic line that allows an inducible perturbation of AG activity

The *amiRNA* described in Section 3.2.1.1 efficiently disrupted *AG* function. In order to identify genes that are directly responsive to *AG* activity, this *amiRNA* was placed under the control of the *35S::AlcR/pAlcA* ethanol-responsive promoter system. This system has been described to allow high levels of expression of a gene of interest in an inducible and transient manner (Roslan et al., 2001). The generation of plants that express the *AG4-amiRNA* in an ethanol-dependent manner is described below.

3.2.2.1 Generation of a *35S::AlcR/pAlcA::AG4-amiRNA* line

The *AG4-amiRNA* was placed downstream of the *AlcA* promoter sequence and wild-type plants were transformed with a T-DNA that contained the *pAlcA::AG4-amiRNA*, *35S::AlcR* expression cassette as described in Section 2.13.1. In total, 63 independent transformants were obtained. These plants were treated with ethanol vapour (see Section 2.5.1 for details of the ethanol treatment) in their second generations and analysed for an ethanol-dependent *ag*-like phenotype. In addition, the expression levels of the *AG* mRNA after ethanol treatment were examined. In brief, T₂ plants that were segregating the *35S::AlcR/pAlcA::AG4-amiRNA* transgene (these plants are denoted as "pAlcA::AG4" plants hereafter) were treated with ethanol vapour by sealing plants in a container containing tubes with 100% ethanol. Wild type and the progeny of all 63 primary transformants were incubated in this sealed container for 48 hours and placed in normal growing conditions (see Section 2.5.1). Once 48 hours had elapsed, the inflorescences from three individual T₂ plants were collected, snap frozen in liquid nitrogen and stored at -80°C. In total, 192 inflorescences were collected (three inflorescences per T₂ line and 3 inflorescences from wild-type *L-er*). The remaining plants were phenotyped for the presence of an *ag*-like phenotype. In addition, untreated T₂ plants were grown to allow for

morphological comparisons between treated and untreated plants. All of the untreated plants were fertilized with pollen from the floral induction system based on the *API* locus (denoted as *pAPI-FIS* hereafter, see Chapter 5) in order to shorten the time between identification of a suitable *pAlcA::AG4* transgenic line and the generation of a FIS containing the *pAlcA::AG4* transgene. The problems encountered with the *35S-FIS*, AG-GFP rescue line (see Section 3.2.3.3) prompted the use of the *pAPI-FIS* instead of the *35S-FIS*.

The ethanol-treated plants were screened for an *ag*-like phenotype for 14 days after ethanol-treatment. Several days after ethanol-treatment, abnormal phenotypes were observed in 25 individual populations. The primary stems of these plants did not elongate and leaves began to senesce earlier than wild-type ethanol-treated plants (not shown). One independent transformant appeared to precociously express the *AG4-miRNA* as an *ag*-like phenotype was observed without ethanol treatment. A further 15 independent transformants displayed mild phenotypes such as a failure of stamens to elongate. Sixteen independent transformants presented carpels that were misshapen and stamens that did not elongate. Finally, six independent transformants produced flowers that displayed indeterminacy and mild stamen to petal homeotic transformations (not shown). The tissue harvested from sibling plants that had been ethanol-treated and stored at -80°C was then used for RNA extraction. Once cDNA had been synthesized, quantitative real-time PCR was performed on this cDNA in order to determine the extent of *AG* mRNA degradation. Line #52 showed the most dramatic reduction in *AG* mRNA levels (down to as little as 10% when compared to wild-type *AG* mRNA levels, not shown). Therefore, this line was isolated in a homozygous state and seeds were propagated. In addition, the cross between this plant line and the *pAPI-FIS* line was planted and homozygous plants for all appropriate transgenes and mutations were identified, and seeds were propagated (see Section 2.13.7). The characterisation of these lines is presented below.

3.2.2.2 Effects of *AG4-miRNA* expression in wild-type plants

A morphological characterization of the *AG4-miRNA* mediated *AG* perturbation was performed in a wild-type background. Initially, plants were treated for 6 hours with ethanol vapour every day for up to 5 days. Although three treatments with ethanol did not

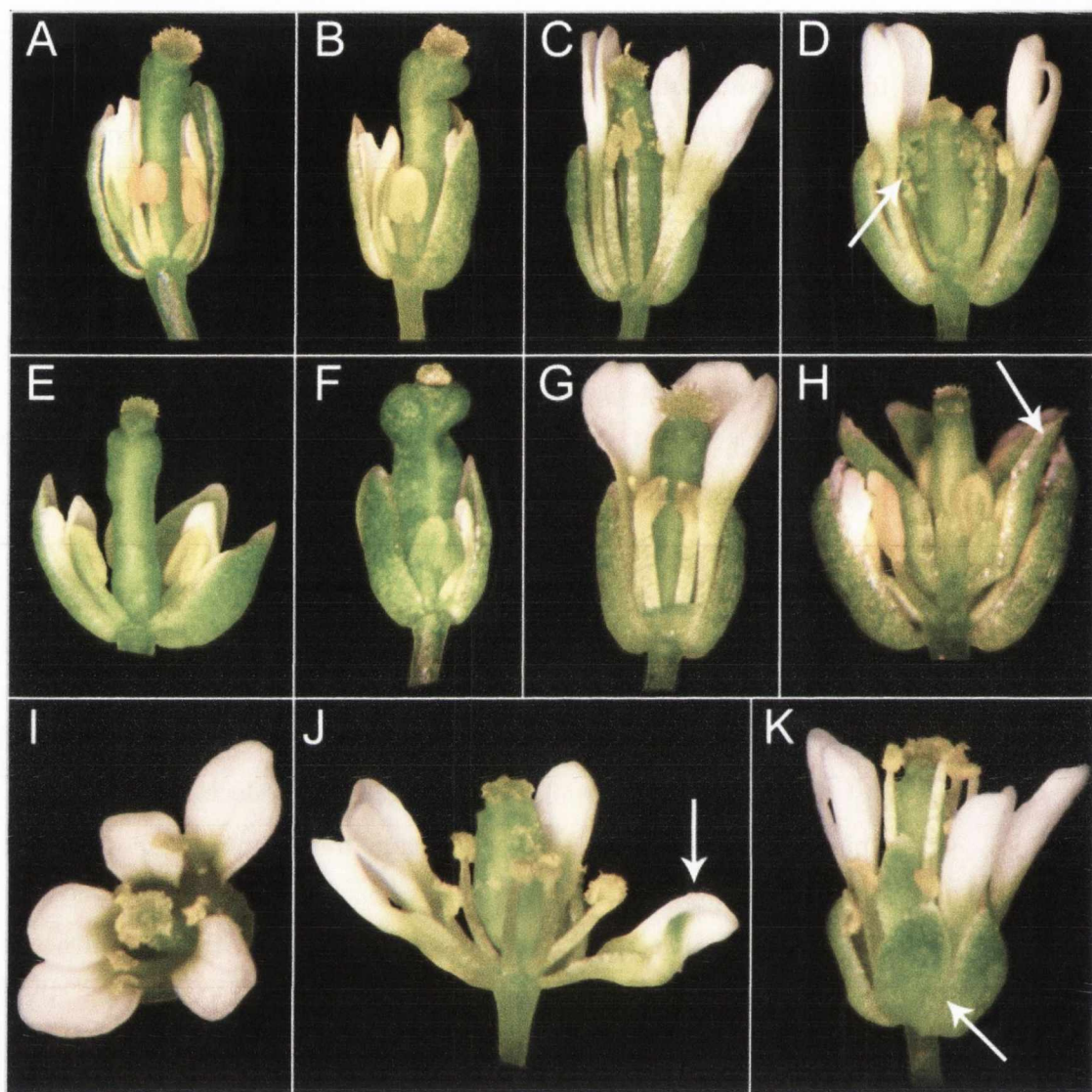


Figure 3.4 Specific knockdown of *AG* activity by inducible *AG4*-amiRNA expression. (A – D) Floral phenotypes observed after *35S::AlcR/pAlcA::AG4* plants were exposed to ethanol vapour for 6 hours. Flowers reached maturity 5 d (A), 9 d (B) 12 d (C), and 15 d (D) after ethanol treatment. An arrow in (D) indicates the presence of ovules and stigmatic tissue. This stigmatic tissue was present on unfused carpel-like tissue. (E – H) Phenotypes observed after *35S::AlcR/pAlcA::AG4* plants were treated twice with ethanol vapour for 6 hours. Treatments were separated by 42 hrs. Flowers reached maturity 6 d (E), 12 d (F), 14 d (G), and 16 d (H) after ethanol treatment. (H) The fourth whorl organs had lost all carpel identity and appeared to adopt a sepal-like morphology. In addition, stamens and carpels developed within this whorl, however, these flowers did not produce seeds. (I – K) Flowers of ethanol-treated plants that were gradually reverting back to a wild-type morphology from separate ethanol treatment experiments to those shown in (A – H). Flowers shown here were all observed in separate ethanol treatment experiments. Flowers reached maturity 20 d (I), 18 d (J), and 19 d (K) after ethanol treatment. Flowers reverting to a wild-type morphology after ethanol treatment often contained extra petals (I), petaloid-sepals (arrow) (J), or fused sepals (arrow) (K).

appear to affect the plants, treatment with ethanol four or more times did result in loss of apical dominance. This made it difficult to assess the phenotype. Therefore, in order to reduce the negative effects of ethanol treatment, plants were treated for 6 hours of 100% ethanol every second day for up to 10 days. Different floral phenotypes were observed depending on what stage of development a given flower was at when ethanol treatment began. The first flowers that arose from an ethanol-treated inflorescence on a given plant were flowers that were treated with ethanol late in development. The last flowers that arose were the flowers that were treated with ethanol early in development. The phenotypes observed after various lengths of ethanol treatment are described below.

First, pAlcA::AG4 plants were treated once with 6 hours of ethanol vapour as described in Section 2.5.1. The first abnormal flowers that arose in these plants contained stamens that had failed to elongate or dehisce pollen. In addition, the shapes of the gynoecia were mildly affected (Figure 3.4A). The next flowers that arose contained underdeveloped stamens and more severely misshapen gynoecia (Figure 3.4B). The flowers that developed subsequently to this displayed petaloid anthers and slightly bulged gynoecia (Figure 3.4C). Finally, flowers were produced that displayed an indeterminate phenotype (Figure 3.4D). The carpels in the inner whorl had begun to separate and an extra carpel had formed within. The tissue surrounding the fused carpel in the innermost whorl developed stigmatic tissue and contained ovules (Figure 3.4D, arrow). Stamen development appeared to have been completed normally.

Next, plants were treated with 6 hours of ethanol vapour twice, with an interval of 42 hours between treatments. Although the progression of floral phenotypes was similar to plants treated once with ethanol vapour, the extent of the mutant phenotypes was more severe. Flowers that were in late development when treated with ethanol produced carpels that were misshapen and stamens did not elongate (Figure 3.4E). The next flowers that arose contained greenish stamens that failed to elongate. In addition, the shapes of the gynoecia were severely affected (Figure 3.4F). Moreover, the petals of these plants were often narrower and shorter than wild-type petals. Although this could be due to ethanol sensitivity, petals at other stages of development appeared normal. The next flowers that arose contained slightly bulged gynoecia and stamens with moderate petal identities (Figure 3.4G). Finally, flowers that were at an early stage of development when plants were treated

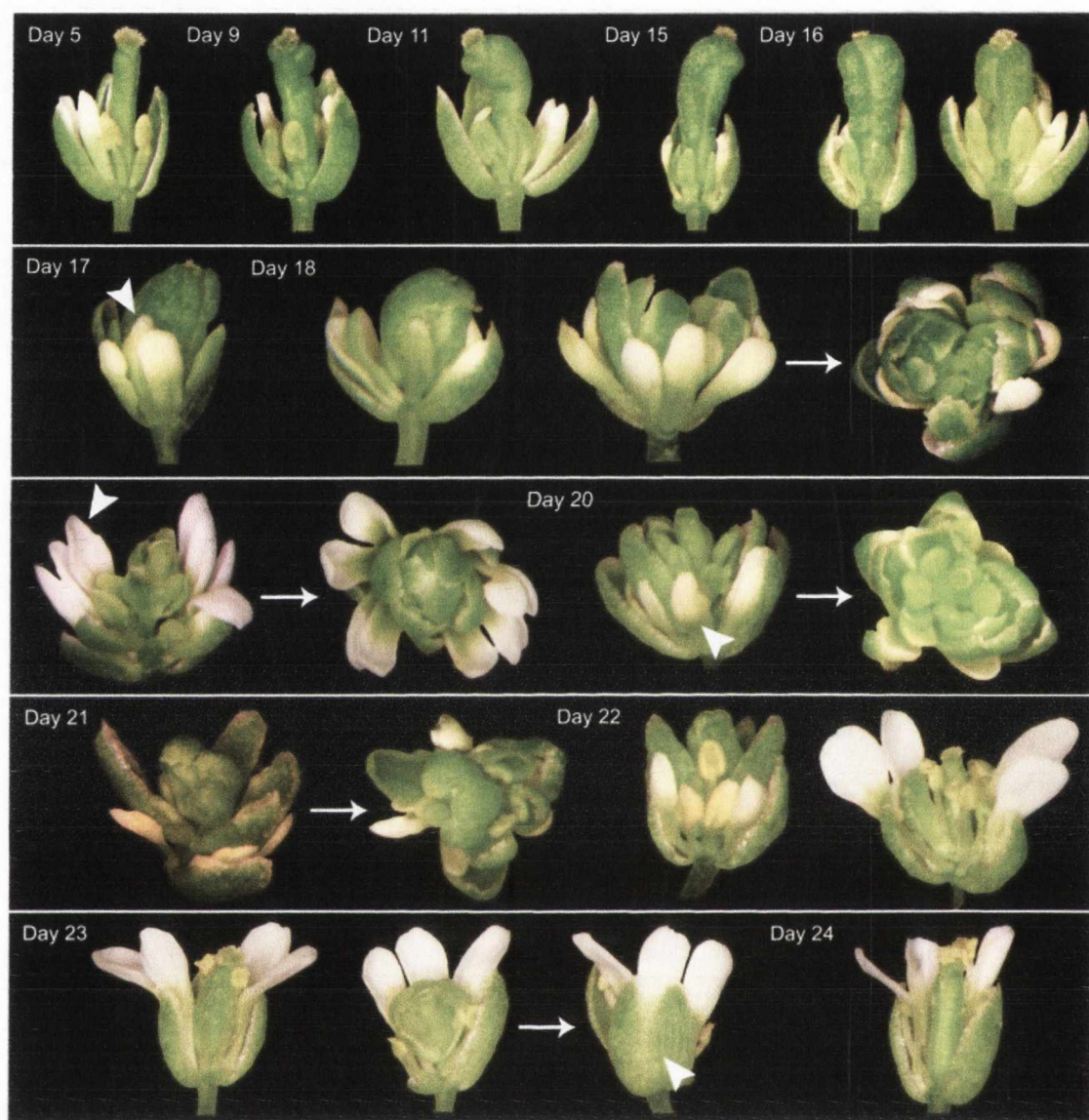


Figure 3.5 Effects of prolonged AG4-amiRNA expression. Plants that were homozygous for the *35S::AlcR/pAlcA::AG4* transgene were treated with ethanol vapour for 6 hours every second day for 10 days in total. Over a 24-day period after the first ethanol treatment, 53 flowers were produced by one representative plant. These flowers exhibited various defects including organ transformations. The most notable mutant phenotypes are presented here. The number of days indicated in the figure refers to the number of days between the first ethanol treatment and the time a photograph was taken. Arrows between individual flowers indicate the same flower photographed from different angles. Flowers observed between day 5 and day 8 (6 flowers in total) contained stamens that failed to elongate or dehisce pollen. Flowers observed between day 9 and day 10 (4 flowers) contained stamens that failed to elongate or dehisce pollen. In addition, the shape of the gynoecium was abnormal. Flowers observed between day 11 and day 14 (11 flowers) produced small, green stamenoid

with ethanol vapour displayed indeterminacy and fourth whorl identity was lost. In addition, the anthers of stamens were rounded although they produced pollen grains (Figure 3.4H). Plants treated with 6 hours of ethanol vapour 3 times, with 42 hours between treatments, contained stamens that had been partially transformed into petals, however, these organs remained a light-yellow colour and locules appeared (not shown). Third whorl stamen identity was lost in flowers of plants treated with 6 hours of ethanol vapour 4 times or more, with 42 hours between treatments. These organs had been transformed completely into petals (not shown).

Flowers from plants treated with 6 hours of ethanol vapour 5 times, with 42 hours between treatments, are shown in Figure 3.5. The most notable changes in floral organ type and shape have been documented. In total, 53 flowers were observed over 24 days that contained abnormal organs. These phenotypes ranged from organ growth defects to organ identity transformations. For simplicity, the changes in stamen and carpel morphology will

Figure 3.5 continued. organs in the third whorl and the shapes of the gynoecia were more distorted. Flowers observed on day 15 (4 flowers) contained very small stamenoid organs that were very broad at their base and the gynoecia remained bulged. Flowers observed on day 16 (6 flowers) contained stamenoid organs in the third whorls of flowers that had become wider and more petal-like in shape and the gynoecia became more bulged. A single flower was documented on day 17. This third whorl organs observed in this flower were white at the tip and green at the base (arrowhead). Locules were still observed (not shown). An extremely bulged gynoecium was observed that contained additional reproductive organs when opened (not shown). Flowers observed at days 18 and 19 (5 flowers) became more indeterminate over time and third whorl organs appeared to have been fully transformed into petals (arrowhead). In addition, carpels had been transformed into four sepal-like structures that did not have any discernable carpel identity. No stigmatic tissues or ovules were observed in these sepal-like organs; however, extra organs including stamens and carpels had begun to form within this whorl. Flowers observed on day 20 (2 flowers) contained third whorl organs that were yellowish and stamenoid (arrowhead). Fourth whorl organs still appeared sepal-like. Flowers observed on day 21 (4 flowers) contained third whorl organs that were broad at their base. The top of these organs produced a yellow petal-like structure that contained locules. No pollen was observed. Fourth whorl organs still appeared sepal-like. Day 22 flowers (6 flowers) produced dehiscent anthers that produced pollen. Furthermore, unfused fourth whorl carpel structures containing stigmatic tissue were observed. Day 23 flowers (5 flowers) observed phenotypically normal stamens and fused carpels with stigmatic tissue at the tip. These gynoecia still contained internal floral organs. The next flowers observed on day 23 contained unfused carpels in the fourth whorl. The sepals of flowers that were reverting to a wild-type phenotype were often fused (arrowhead). The final flower photographed on day 24 was phenotypically normal. Flowers that reached maturity subsequently were phenotypically normal as well.

be discussed separately.

The first flowers that arose after ethanol treatment suffered from stamen elongation defects. In addition, anthers failed to dehisce and release pollen grains (Figure 3.5, Day 5). These stamen elongation defects were observed up to 10 days after ethanol treatment when stamens began to change in morphology (Figure 3.5, Day 11). Stamens became green at day 11 and the base of these organs became broader with every maturing flower. At day 16, stamens still contained locules, however, the colour had begun to shift from green to white (Figure 3.5, Day 16). The first flowers observed at day 18 contained third whorl organs where the top halves had become completely white while the bottom halves were still green (Figure 3.5, Day 18). Additionally, these organs were often yellow in patches. The last flowers observed at day 18 produced full stamen to petal transformations (Figure 3.5, Day 18). Six petals were observed in the third whorl that did not contain any stamen identity. At day 20, third whorl organs had developed locules again although they remained white (Figure 3.5, Day 20). These organs became yellow by day 21 and the base had become narrower (Figure 3.5, Day 21). By day 22, fully formed stamens had appeared and although mild elongation defects were observed, the anthers produced pollen (Figure 3.5, Day 22).

Carpel defects were first observed in flowers 5 days after the first ethanol treatment (Figure 3.5, Day 5). The gynoecia of flowers that had matured between 9 and 16 days after ethanol treatment were abnormally shaped (Figure 3.5, Day 9 – 16) and eventually became bulged at day 17 (Figure 3.5, Day 17). These bulged gynoecia contained extra reproductive organs. Carpels became unfused on day 18 (Figure 3.5, Day 18) and all carpel identity was lost in the fourth whorl. Sepal-like organs replaced carpels in the fourth whorl; however, carpels and stamens had formed in repeated whorls (Figure 3.5, Day 18). Organs in the fourth whorl retained a sepal-like morphology until day 22 when stigmatic tissue was observed (Figure 3.5, Day 22). Fully fused, gynoecia were observed 23 days after ethanol treatment (Figure 3.5, Day 23). The carpels often separated again (Figure 3.5, Day 23) after production of a fully fused gynoecium before a wild-type gynoecium was present at day 24 (Figure 3.5, Day 24).

Next, plants were treated 5 times for 6 hours with ethanol vapour with 66 hours in between treatments. Although fourth whorl carpels were transformed to sepal-like organs in these plants, stamens did not completely transform into petals (not shown). Third whorl organs remained slightly yellow and locules usually developed (not shown). Otherwise, these flowers strongly resembled the flowers represented in Figure 3.5.

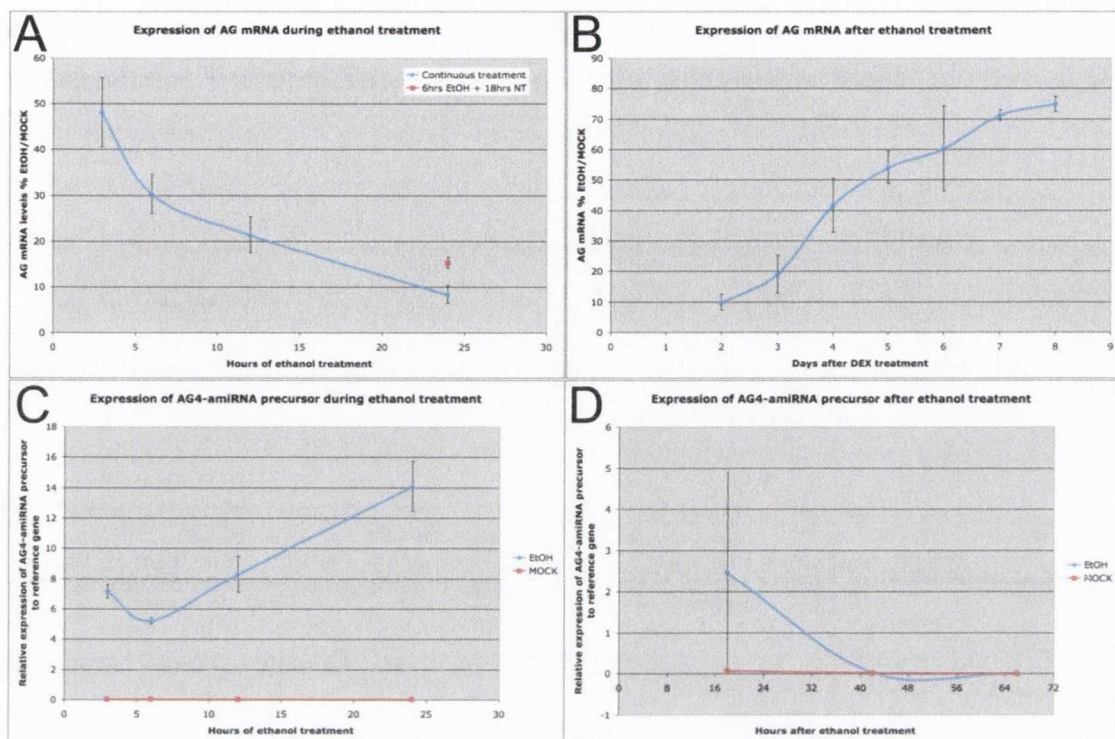


Figure 3.6 Analysis of *AG* mRNA and *AG4-amiRNA* precursor expression in *pAPI-FIS*, *35S::AlcR/pAlcA::AG4* plants. RNA levels were determined by qRT-PCR. **(A)** Expression of *AG* mRNA during ethanol treatment. Plants were first treated with dexamethasone to induce flower formation. After 4 d, they were exposed for 3, 6, 12 and 24 hours to ethanol vapour (blue line) or were mock-treated. Tissue was harvested after each treatment. Other plants were treated for 6 hours with ethanol vapour or were mock-treated. Tissue was collected 18 hours after the end of the treatments (red data point), *i.e.* 24 hours after the experiment commenced. **(B)** Expression of *AG* mRNA after ethanol treatment. Plants that had been treated with dexamethasone were exposed 1 d after the initial treatment for 6 hrs to ethanol vapour or were mock-treated. Tissue was harvested 18 hours later and then again every 24 hours for 6 days. **(C)** Expression of the *AG4-amiRNA* precursor during ethanol treatment. The same RNA as in **(A)** was used. **(D)** Expression of the *AG4-amiRNA* precursor after ethanol treatment. The same RNA samples as in **(B)** were used. Error bars are s.e.m. $N=3$. The primer pairs DM-76/DM-77, DM-78/DM-79 and DM-97/DM-98 were used to amplify the *AG*, *amiRNA-precursor* and reference gene (*At4g34270*) cDNAs, respectively.

Finally, flowers in which first or second whorl organs were observed (Figure 3.4I – K). This was independent of treatment length. Fused sepal, petaloid sepals and extra petals were the most frequent phenotypes observed. These flowers were always observed after the most severe organ transformations had been observed, indicating that these flowers were affected by amiRNA expression very early in development. Possible explanations relating to these phenotype are discussed in Section 3.2.8.2.

In summary, strong *ag*-like phenotypes were generated when pAlcA::AG4 plants were treated with ethanol. These phenotypes included complete transformations of stamens and carpels into petals and sepals, respectively. In addition, indeterminate flowers were generated by ethanol treatment. These phenotypes showed that *AG* activity could be inducibly perturbed in a temporary manner.

3.2.2.3 Molecular characterization of the *AG4-miRNA* expression in a floral induction system

The main goal of this project is to investigate the function of *AG* at early stages of flower development, which can only be done in the FIS. Therefore, the molecular characterization of AlcA-AG4 line was done in the FIS. The floral induction system based on the *API* locus was crossed with pAlcA::AG-52 plants. Plants harbouring all of the appropriate transgenes and mutations were isolated as described in Section 2.13.8. This transgenic line is referred to as “FIS-AG4” hereafter. This line responded to dexamethasone as expected (not shown). Next, the responsiveness of *AG* mRNA and *AG4-miRNA* precursor levels to ethanol was investigated.

First, the sensitivity of *AG* mRNA levels to ethanol was examined. Four days after plants had been treated with dexamethasone (see Section 2.5.2), they were treated with ethanol vapour and mock-treated for 3, 6, 12 and 24 hours, as described in Section 2.5.1. Tissue was harvested immediately after vapour treatments, total RNA was extracted, and cDNA was synthesized. Using qRT-PCR, I found that *AG* transcript levels were reduced to approximately 50% of those in mock-treated plants levels after three hours (Figure 3.6A). *AG* mRNA levels further decreased to ~30% and ~20% after 6 and 12 hours of exposure to ethanol vapour, respectively, until *AG* mRNA levels are just below 10% mock-treated levels after 24 hours ethanol vapour treatment. Given the sudden reduction of *AG* mRNA levels and the ethanol-induced phenotypes observed in Figure 3.4 and 3.5, it was hypothesized that *AG* mRNA levels might continue to fall even after the treatment was stopped. Plants that had been treated with dexamethasone 3 days earlier were treated with ethanol vapour or were mock-treated for 6 hours and tissue was harvested 18 hours later. *AG* mRNA levels were reduced to approximately 15% of those in mock-treated plants (Figure 3.6A). Therefore, *AG* mRNA continued to be degraded after ethanol treatment had ceased.

Next, the recovery of *AG* mRNA levels after amiRNA-mediated degradation was investigated. Plants that were treated with dexamethasone one day previously were treated with 100% ethanol vapour and water vapour for 6 hours. Tissue was harvested after 18 hours and then every 24 hours for up to 8 days after dexamethasone treatment. RNA was extracted from these tissues and cDNA was synthesized. *AG* mRNA levels dropped to ~10% of mock-treated levels 18 hours after ethanol treatment. After this treatment, the levels of *AG* mRNA in ethanol-treated plants increased every 24 hours, ultimately reaching 75% of mock-treated levels 7 days after the initial treatment. *AG* mRNA levels had recovered to ~50% between 4 and 5 day after dexamethasone-treatment.

The expression of the *AG4-amiRNA* precursor was also investigated. The same RNA used for the qRT-PCRs represented in Figure 3.6A and B was used to synthesize cDNA and primers specific to the amiRNA-backbone were used to estimate the levels of the *AG4-amiRNA* precursor (Figure 3.6C). After 3 hours of ethanol vapour treatment, precursor levels were increased dramatically above background (Figure 3.6C, blue line). Precursor levels dropped slightly after 6 hours before increasing again after 12 hours of ethanol treatment. *AG4-amiRNA* precursor levels peaked after 24 hours of ethanol treatment. The *AG4-amiRNA* precursor levels in mock-treated plants did not change over time (Figure 3.6C, red line).

Next, the expression of the *AG4-amiRNA* precursor once ethanol treatment had been removed was investigated. The same RNA used in Figure 3.6B was used to generate cDNA in order to measure *AG4-amiRNA* precursor levels. Although *AG4-amiRNA* precursor levels were higher than mock-treated plants when measured 18 hours after ethanol treatment, levels varied greatly (Figure 3.6D, blue line). *AG4-amiRNA* precursor levels were comparable in mock- and ethanol-treated plants 42 and 66 hours after ethanol treatment. These analyses show that the expression of the *AG4-amiRNA* precursor is highly responsive to, and tightly regulated by ethanol-treatment.

3.2.3 Generation of AG antibodies and of lines expressing epitope-tagged AG

As discussed in Section 1.5.3, highly specific antibodies against the protein of interest

must be available to perform ChIP assays. Such antibodies can either be directed against the endogenous protein or alternatively, they can target an epitope-tag that is fused to the protein of interest. For the present study, both of these strategies were used to maximize the chance of identifying suitable antibodies for ChIP experiments.

3.2.3.1 Translational fusion of a haemagglutinin (HA) epitope tag to AG

Initially, a tag consisting of six repeats of the haemagglutinin epitope tag (denoted as 6xHA) was translationally fused to the C-terminus of *AG* (Figure 3.7A) as described in Section 2.13.4. The resulting transformation vector, which conferred resistance to the herbicide ammonium-glufosinate (“BASTA”), was introduced into plants segregating the strong *ag-1* mutant allele. Six independent transformants were obtained and genotyped for the presence of both the transgene and the *ag-1* mutation. Two of the transformants were homozygous for the *ag-1* allele, two were heterozygous and two were wild type. The mutant phenotypes of the homozygous *ag-1* plants were partially rescued by the transgene. Stamen-like organs were present and immature carpels were observed (Figure 3.7A). These plants were not fertile and thus could not be propagated. The remaining transformants contained a functional copy of *AG* (Figure 3.7C, D), however, carpels were misshapen and organ transformations were observed (Figure 3.7C, D). Stamens often failed to elongate or dehisce and release pollen grains, and siliques were often bulged (not shown). The phenotypes observed in plants carrying the *AG-6xHA* transgene with an endogenous functional copy of *AG* suggested that the expression of the transgene had produced antimorphic defects (see Section 3.2.8.1). Therefore, no further experiments were performed using this transgene.

3.2.3.2 Translational fusion of an mGFP5 reporter to AG

Since expressing the translational fusion between *AG* and the 6xHA tag failed to rescue the *ag-1* mutant phenotype, an alternative strategy was pursued. To this end, an mGFP5 reporter gene was translationally fused to the C-terminus of the genomic locus of *AG*. The resulting construct was transformed into a population of plants segregating the *ag-1* mutation as described in Section 2.13.5. From five independent transformants, one was homozygous for *ag-1*; however, the transgene did not rescue the null mutant phenotype.

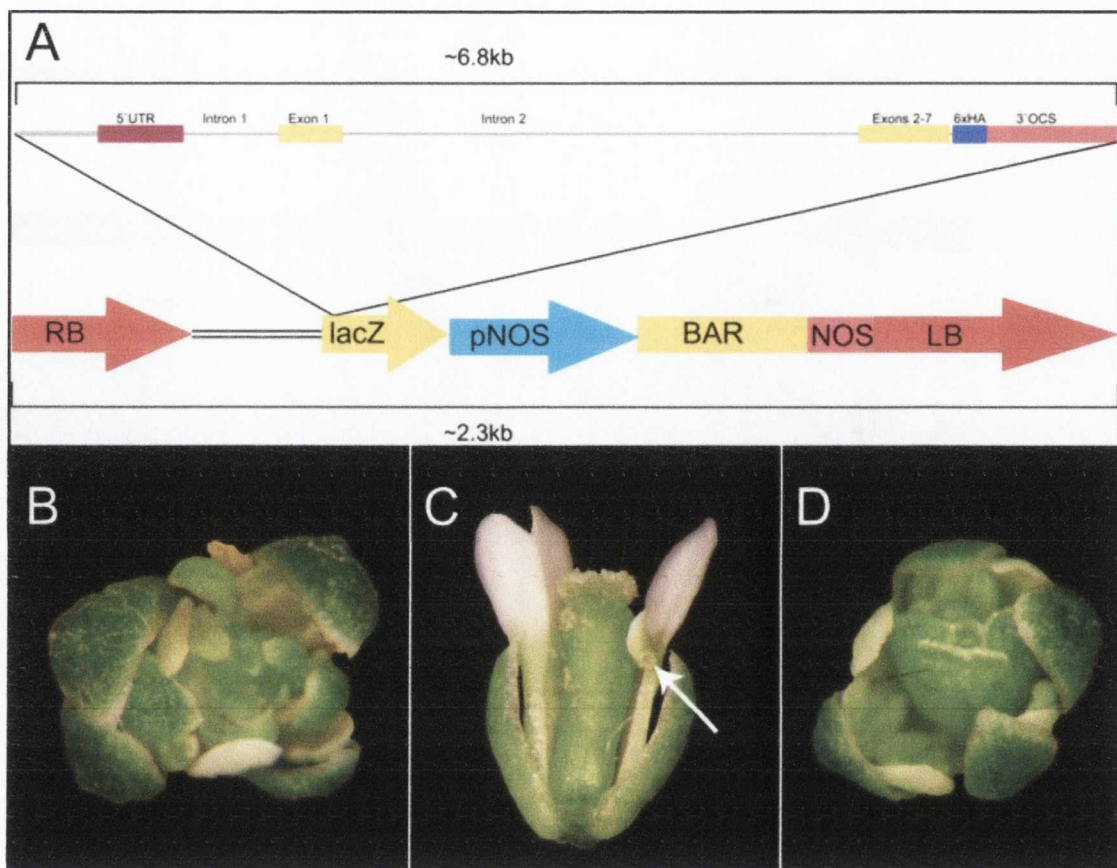


Figure 3.7 Floral phenotypes of *pAG::AG-6xHA* plants. (A) A schematic of the T-DNA mediating the expression of the *pAG::AG-6xHA* transgene. The binary vector pML-BART was used in conjunction with *Agrobacterium tumefaciens* to transform a population of *ag-1* segregating plants. For full details see section 2.13.4. (B) A flower of a homozygous *ag-1* plant (T2 generation) carrying at least one copy of the *pAG::AG-6xHA* transgene. Reiterations of stamenoid and carpelloid organs were observed in these flowers but stamens never fully matured. (C – D) Flowers of *ag-1/+* plants (T2 generation) carrying at least one copy of the *pAG::AG-6xHA* transgene. (C) A flower that contains a bulged carpel and a petaloid stamen (arrow). One sepal, one petal and two stamens were removed for better visibility of the inner whorl organs. (D) A flower with disrupted carpel identity and determinacy defects. UTR, untranslated region; 3'OCS, octopine synthase terminator sequence; RB, right border of T-DNA; LB, left border of T-DNA; lacZ, a modified version of the β -galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA® resistance gene; NOS, nopaline synthase terminator sequence.

Two other independent transformant plants were homozygous for the wild-type allele of *AG* and the remaining two transformants were heterozygous for *ag-1*. The seeds from the *ag-1/+* transformants were planted. Offspring were genotyped for the presence of the transgene and the *ag-1* allele. All plants from one of the two original transformants appeared phenotypically normal when the transgene was present, as determined by their

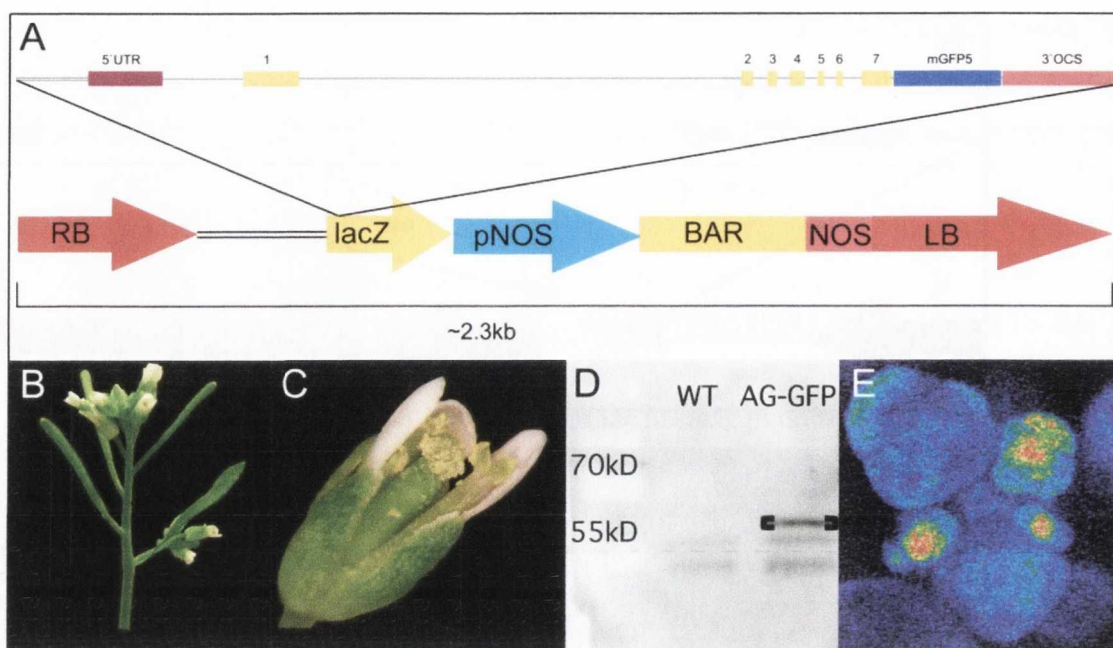


Figure 3.8 Characterization of *pAG::AG-GFP ag-1* plants. (A) A schematic of the T-DNA insertion carrying the *pAG::AG-GFP* transgene. The binary vector pML-BART was used in conjunction with *Agrobacterium tumefaciens* to transform a population of *ag-1* segregating plants. For full details see Section 2.13.5). (B – C) The phenotype of a plant (B) and an individual flower (C) carrying the *pAG::AG-GFP* transgene in a homozygous *ag-1* mutant background. (D) Results of Western blot analysis of nuclei extracts from inflorescences of plants expressing AG-GFP and plants that do not express AG-GFP (WT) using an anti-GFP antibody. A band at ~56 kDa is observed, which is the expected size of the AG-GFP fusion protein. This band was not present in plants that lack the corresponding transgene. (E) A confocal laser scanning microscopic image of a dissected *pAG::AG-GFP ag-1* inflorescence showing expression in the centre of floral meristems. UTR, untranslated region; 3' OCS, octopine synthase terminator sequence; RB, right border of T-DNA; LB, left border of T-DNA. lacZ, a modified version of the β -galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA® resistance gene; NOS, nopaline synthase terminator sequence.

herbicide resistance and by PCR-based genotyping. Plants that were homozygous for *ag-1* and contained the transgene appeared normal (Figure 3.8A, B). All floral organs were present and the plants were completely fertile. In addition, confocal microscopy revealed that the expression of the AG-GFP fusion protein occurred in a pattern that matched that for endogenous AG protein (Figure 3.8D). Moreover, Western blotting with an α -GFP antibody identified a band corresponding to the size expected (~56 kDa) for the AG-mGFP5 protein. This indicated that the AG-mGFP5 fusion protein remained intact after translation (Figure 3.8C). The genomic location of the T-DNA insertion could not be identified although

several attempts were made using Thermo Asymmetric Interlaced PCR (TAIL), therefore a homozygous line was identified by analyzing segregation patterns of plants resistant to BASTA and by subsequent genotyping and phenotyping (see section 2.13.6).

3.2.3.3 Generation of a floral induction system expressing *AG-mGFP5*

Plants harbouring the *AG-mGFP5* transgene and the *ag-1* allele were crossed with the 35S::*API-GR* floral induction system (denoted as *35S-FIS* hereafter) as described in Section 2.13.6. Once plants that were homozygous for all appropriate transgenes and mutations were identified, the lines themselves were characterized. When the inflorescence-like meristems of these plants were treated with dexamethasone (as described in Section 2.5.2), they failed to respond as expected. No response could be observed 5 days after treatment (Figure 3.9A, B).

Although treated plants produced flowers faster than non-treated plants (data not shown), the flowers produced lacked sepals and petals (Figure 3.9C). The accelerated flowering of dexamethasone-treated plants relative to mock-treated plants suggested that the *API-GR* protein was still active. Therefore, it seemed unlikely that the 35S::*API-GR* transgene had been post-transcriptionally silenced, however, this was investigated as a precaution. RNA was extracted from untreated inflorescence-like tissue and leaves of *ap1 cal*, *35S-FIS*, and the progeny of two independent *35S-FIS*, *AG-GFP* F₂ lines that were homozygous for all appropriate transgenes and mutations. Once cDNA was synthesized, PCR analysis using primers specific to the *GR* and *LIPASE* genes was performed (Figure 3.9D). DNA bands corresponding to the size of the *GR* amplicon (see Figure Legend 3.9C) were observed for the *35S-FIS* and the *35S-FIS*, *AG-GFP* inflorescence and leaf samples. This indicated that the 35S::*API-GR* transgene was being expressed in plants that also contained the *pAG::AG-GFP* transgene.

Therefore, I hypothesized that overexpression of the *API-GR* transgene may have led to elevated levels of the *AG-GFP* mRNA in comparison to wild type and/or misexpression of the *AG-GFP* mRNA. Given that the *AG* protein is known to interact with the *API* protein (Riechmann et al., 1996) and that *AG* and *API* have antagonistic functions during flower development (Mandel et al., 1992), this overexpression may have led to disruption of *API* function. Therefore pollen from the *35S-FIS*, *AG-GFP* flowers was used to fertilize

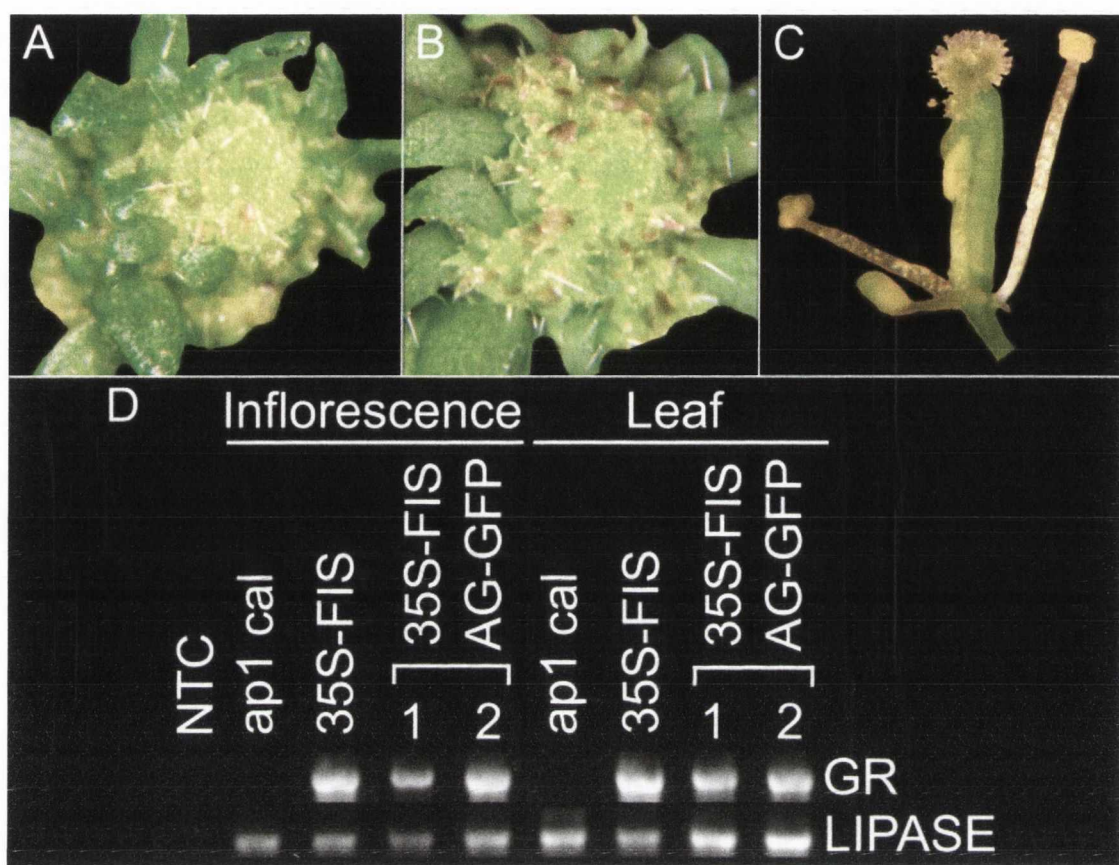


Figure 3.9 Characterization of *35S-FIS pAG::AG-GFP ag-1* transgenic plants. (A) An inflorescence-like meristem 5 days after being treated with a mock solution (see Section 2.7.2 for details). (B) An inflorescence-like meristem 5 days after being treated with a dexamethasone-containing solution. No synchronous induction of flowering was observed. (C) A flower produced from a dexamethasone-treated inflorescence-like meristem. (D) Detection of AP1-GR expression by RT-PCR. The cDNA used as a template for the reactions was generated from total RNA extracted from *ap1 cal*; *35S-FIS*; and *35S-FIS pAG::AG-GFP* inflorescences and leaves. For PCR amplification (30 cycles), primers complementary to a fragment of the rat glucocorticoid receptor were used. In addition, transcripts for a lipase gene were amplified. The lipase PCR assay could amplify sequences representing genomic and complimentary DNA. This was used to estimate genomic DNA contaminations of the RNA preparations. A no template control (NTC) was also included. Primer pairs DM-101/DM-102 (Primers from (Ito et al., 2004)) and DM-31/DM-68 were used to analyse the *LIPASE* (At4g27330) and *GR* cDNAs.

the gynoecium of the *pAP1::AP1-GR* floral induction system (referred to hereafter as *pAP1-FIS* (see Chapter 5)). This transgenic line should express the *AP1-GR* mRNA in a domain that is spatially and temporally separate from the *AG* mRNA. Plants were subsequently isolated that contained all appropriate transgenes and mutations (see Section 2.13.7). When the inflorescence-like meristems of these plants were treated with a dexamethasone-

containing solution, synchronous flowering was observed (Figure 3.10A, B). In addition, sepal and petal identity was restored to dexamethasone-treated flowers (Figure 3.10C, D). In order to verify that the AG-GFP protein was present in its appropriate domain, fluorescence microscopy was used to image dexamethasone-treated inflorescence-like meristems. No signal was observed in inflorescence-like meristems one or two days after the initial dexamethasone treatment (not shown), however, clear fluorescent signals were observed in the centre of flowers from 3 – 5 days after dexamethasone treatment (Figure 3.10E-G, arrowheads). This transgenic line was propagated and used in chromatin immunoprecipitation experiments as described in Section 3.2.7.

3.2.3.4 Generation of AG peptide antibodies

In order to generate antisera that specifically recognize AG, peptides were designed that were not homologous to other proteins in *Arabidopsis*. First, the entire AG protein was subject to BLAST analyses in order to identify regions that were unique in the amino acid sequence. Next, three-dimensional protein structure predictions were carried out to identify regions that would most likely appear on the surface of the AG protein. In addition, these regions were gauged for their relative hydrophilicity. Finally, the sequences that satisfied these criteria were analyzed for the likelihood of forming coiled-coil structures, which are often present on the surface of protein. The peptide sequences that matched these criteria were CKIAENERNNPSISL (Ab1) and CPPQTQSQPFDSRNY (Ab2). GenScript USA Inc. then synthesized these peptides and raised rabbit polyclonal antisera against them. The resulting antisera were affinity-purified against the peptides used for immunization in order to generate a more specific fraction of peptide antibodies. These antibodies were extensively tested by Western blotting and by ChIP-qPCR. No protein band corresponding to the size of the AG protein was ever observed (not shown) and no enrichment was observed for known AG binding regions when ChIP-qPCR was performed with these antibodies (not shown). Therefore, these peptide antibodies were not used in any further experiments.

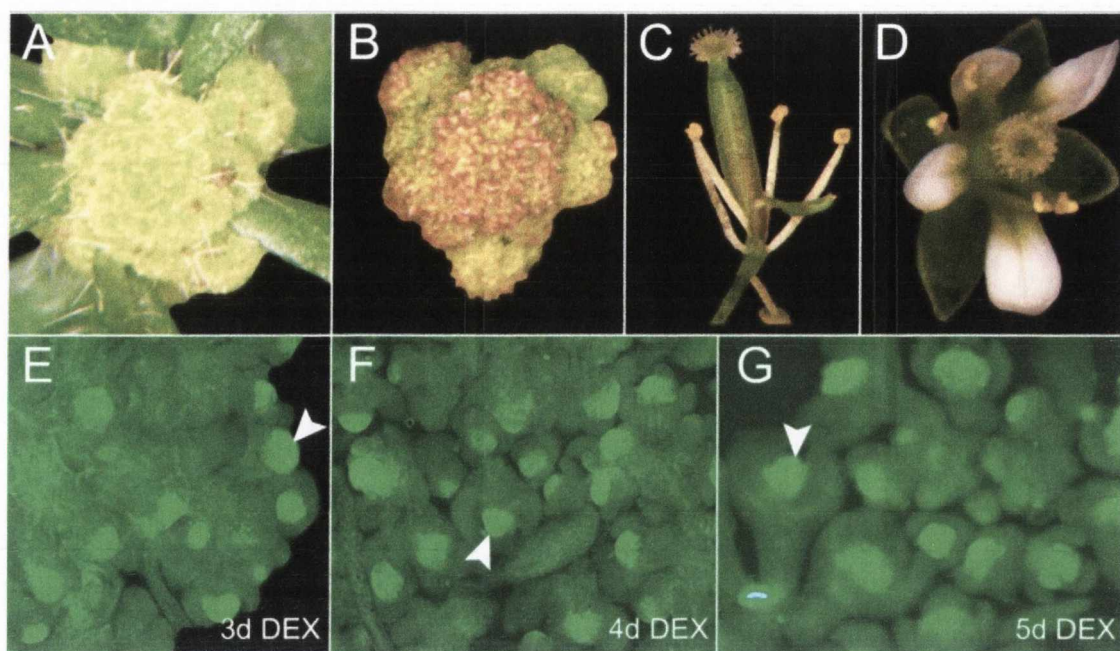


Figure 3.10 Characterization of *pAPI-FIS pAG::AG-GFP ag-1* transgenic plants. (A) An inflorescence-like meristem 5 days after being treated with a mock solution (see Section 2.5.2 for details). (B) An inflorescence-like meristem 5 days after being treated with a dexamethasone-containing solution. (C) A flower produced by a mock-treated inflorescence-like meristem. (D) A flower produced from a dexamethasone-treated inflorescence-like meristem. (E – G) Fluorescent microscopic images (see Section 2.4.2) of inflorescences 3 d (E), 4 d (F) and 5 d (G) after dexamethasone treatment. No GFP signal was observed 1 or 2 days after dexamethasone treatment. Arrowheads indicate GFP fluorescence.

3.2.4 Identification of genes acting downstream of AG through expression profiling of *ag* mutant flowers

In addition to the AG4-miRNA based ethanol-inducible system described below (Section 3.2.5), I used the *AG* mutant allele *ag-1* to identify, by microarray analysis, genes that respond to a perturbation of *AG* activity at different stages of early flower development. To this end, I first crossed the *ag-1* allele into the original, *35S* promoter-based floral induction system (Wellmer et al., 2006). Upon dexamethasone treatment, these plants produced the expected null mutant phenotypes, *i.e.* stamens and carpels were replaced by reiterations of petals and sepals, respectively (Figure 3.11A – C). The petals in this line were greenish and not as white as in wild-type flowers, suggesting that prolonged API activity might be required for proper petal maturation.

Because *ag* mutants are sterile due to a loss of all reproductive organs, the *35S::AP1-GR ap1-1 cal-1 ag-1* line had to be maintained as an *ag-1* heterozygote. Given that the *ag-1* mutation is recessive, only 25% of the offspring of this line would be homozygous for the *ag-1* allele. To allow the collection of sufficient numbers of *ag* mutant flowers for analysis by gene expression profiling, the following experimental strategy was adopted.

3.2.4.1 Experimental design and implementation

In order to compare the transcript profiles of *ag* mutant and wild-type flowers at different stages of early flower development, plants of the floral induction system that were segregating the *ag-1* mutant allele were used. Inflorescence tissue from individual plants was collected into separate tubes immediately before the induction of flower formation as well as 2.5 and 5 days after dexamethasone treatment. Subsequently, the axillary meristems of these plants were allowed to grow out and the resulting flowers were examined to identify plants that exhibited a loss of reproductive floral organs and thus were in all probability homozygous for the *ag-1* allele. Whenever the results of this phenotyping approach were inconclusive, PCR-based genotyping was performed as well to verify the genotype of the plants used for tissue collection. The tissue from approximately 20 individual plants identified as being homozygous for *ag-1* was then pooled for each time point taken, resulting in *ag* mutant flower samples. In addition, tissue from plants that formed wild-type flowers (these plants were either heterozygous for *ag-1* or were homozygous for the wild-type *AG* allele) was also combined, resulting in wild-type control samples for the different time points taken. This process of tissue collection was repeated several times and in total, four sets of biologically independent samples were obtained. Total RNA was extracted from the tissue samples, the RNA was amplified, labelled with fluorescent dyes and then hybridized to custom-designed Agilent whole-genome *Arabidopsis* microarrays as per manufacturers instructions.

After pre-processing and normalization of the microarray data (see Section 2.12.2), the expression levels of several genes known or suspected to be downstream of *AG* were inspected (Figure 3.11D). For example, *HEC1*, *HEC2*, *SHP1* and *STY1* are all involved in the formation of gynoecia (Gremski et al., 2007; Pinyopich et al., 2003; Sohlberg et al., 2006). In addition, *CRC*, *SPL*, *SHP2* and *KNU* have been proposed to be direct targets of

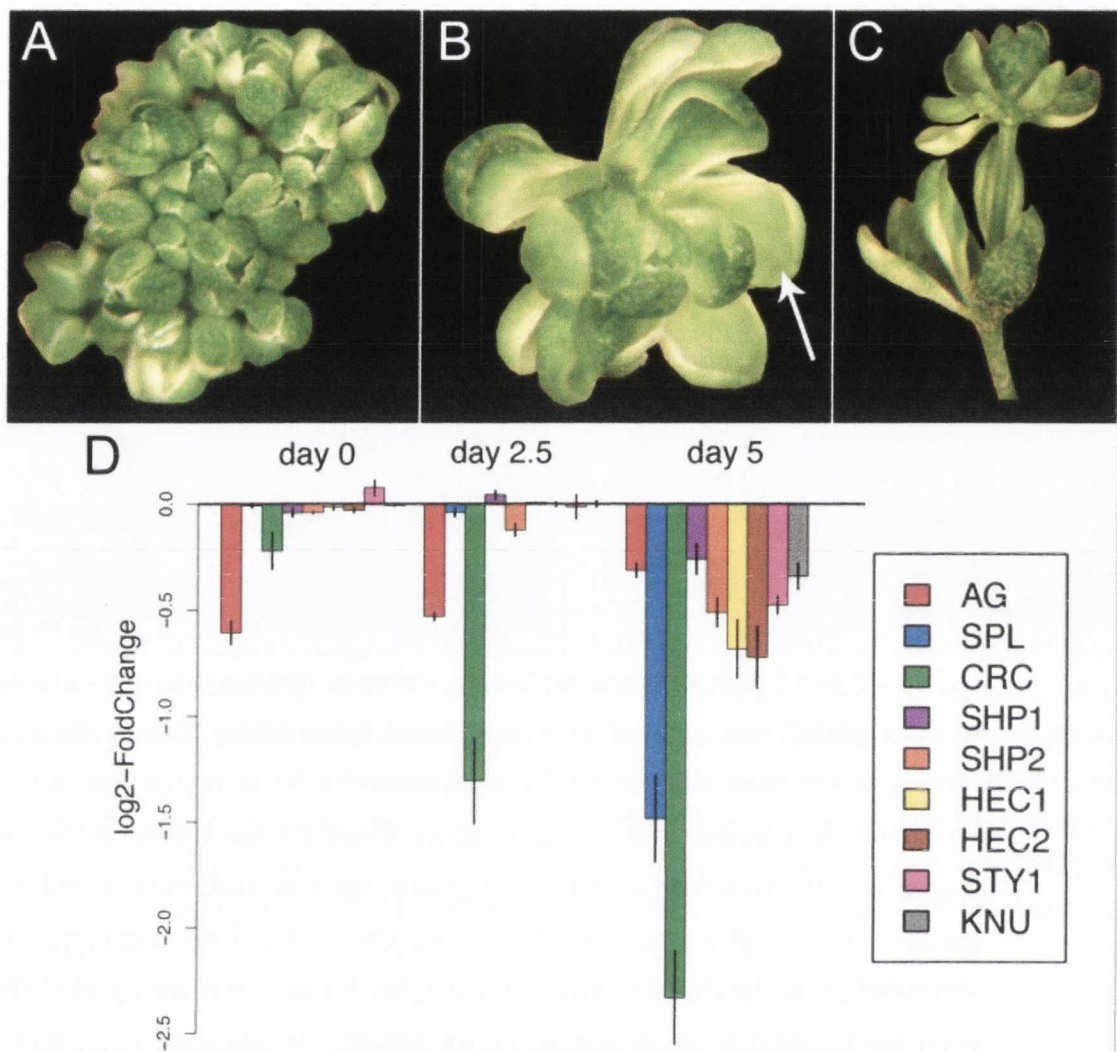


Figure 3.11 Phenotypes of 35S::AP1-GR *agl-1 cal-1 ag-1* plants and the expression of *AG*-dependent genes in *ag* mutant flowers. (A) An 35S::AP1-GR *agl-1 cal-1 ag-1* inflorescence 16 days after induction of flower formation by dexamethasone treatment. Floral buds have reached maturity. (B and C) Flowers produced from dexamethasone-treated inflorescence-like meristems. An arrow in (B) marks a greenish petal developing in the second whorl. (D) Expression of genes known or suspected to be *AG*-dependent in *ag* mutant flowers relative to wild-type flowers. Flowers were collected at different time points (as indicated) after induction of flowering by dexamethasone treatment and RNA isolated from these flowers was subjected to microarray analysis. Data shown are log₂-transformed fold change expression ratios (*ag* vs. wild type). Error bars indicate s.e.m.

AG (Gomez-Mena et al., 2005; Ito et al., 2004; Savidge et al., 1995; Sun et al., 2009). The expression of all of these genes was lower in *ag-1* mutant flowers than in the wild type at the 5 day time point (Figure 3.11D), while (with the exception of *CRC*) no or only minor expression differences were detected for these genes after 2.5 days. Overall, *CRC* and *SPL*

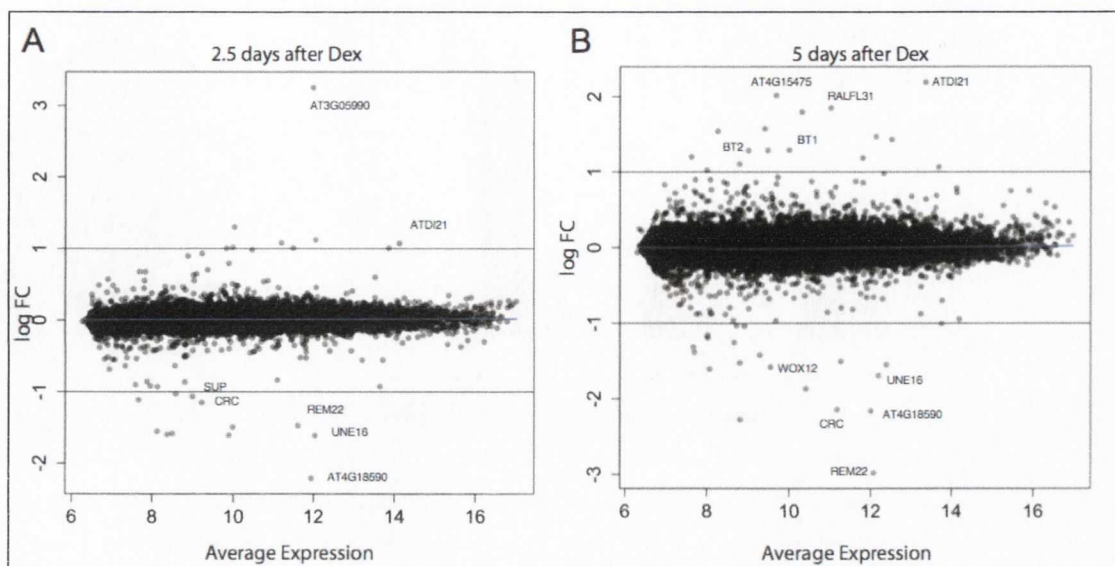


Figure 3.12 Microarray data from the comparison of the transcriptomes of *ag* mutant and wild-type flowers. **(A)** Data for the 2.5 day time point. **(B)** Data for the 5 day time point. Mean expression values (x-axis) are plotted against log₂-transformed fold change (FC) values (*ag* vs. wild type; y-axis). The data shown are based on the results of hybridizations of four sets of biologically independent tissue samples. Selected data points are annotated using the associated gene symbols or AGIs.

were the most strongly down-regulated genes at 5 days whereas the effects for the other genes mentioned above were weaker. The expression of *AG* itself was only slightly affected in agreement with data from a previous report showing that high levels of *AG* mRNA accumulate in *ag-1* and *ag-2* mutant flowers (Gustafson-Brown et al., 1994).

Although the overall response of these genes was in agreement with published data, the fold change expression differences were lower than expected. Therefore, a global overview of gene expression changes in the *ag* mutant experiments was obtained by averaging signal intensities and log₂-transformed fold changes over all four biological replicates and then by using these data to generate so-called MA-plots (Figure 3.12). These plots revealed that relatively few genes were strongly differentially expressed in any of the three time-points taken (the data for Day 0 are not shown in Figure 3.12). They also indicated that the expression differences were larger at day 5 than at day 2.5. An approximately equal number of genes appeared to be up- and down-regulated at day 2.5 (Figure 3.12A), whereas there appeared to be a slight shift towards down-regulated genes at day 5 (Figure 3.12B).

Experiments similar to the one described above for *ag* mutant flowers were performed

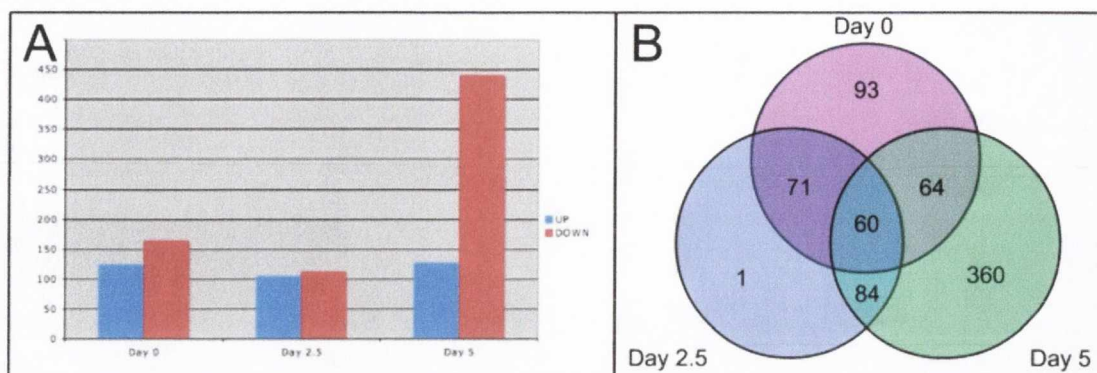


Figure 3.13 The directionality and the number of genes determined to be differentially expressed in *ag* mutant flowers. (A) The number and directionality of genes considered differentially expressed. **(B)** A Venn diagram indicating the overlap of differentially expressed genes between each time-point. Microarray experiments were done as described in the main text. A Bayesian approach and linear models (unadjusted p -value < 0.01) were applied using the Bioconductor package *limma* to infer differential gene expression. The number of up- (blue bars) and down-regulated (red bars) genes are shown for each time point (as indicated).

by Samuel Wuest using mutant alleles for *AP3* and *PI*. Using the data from these analyses and the *ag* mutant experiment, a Principal Component Analysis (PCA) was performed to identify the main sources of variation within and between each experiment (not shown). This revealed that inter-experimental variation stemmed largely from time points as opposed to genotypes (not shown). Given these small fold-change differences, using traditional analytical methods, such as fold change cut-off values would not result in the identification of many differentially expressed genes. Therefore, an alternative, process-oriented statistical approach was followed.

3.2.4.2 AG functions primarily as a transcriptional activator and controls a multitude of cellular and developmental processes

In order to identify genes differentially expressed in *ag* mutant flowers compared to the wild type, a linear model was fitted to each gene and an empirical Bayes method (Smyth *et al.*, 2004) was employed to identify genes that are differentially expressed (statistical analyses performed by Samuel Wuest). After testing for differential gene expression using this analytical method, a total of 733 individual genes were designated as differentially expressed (p -value < 0.01 ; unadjusted for multihypothesis testing) over the three time-points taken. At day 0, 288 genes were identified, whereas 216 and 568 genes were deemed

differentially expressed at day 2.5 and day 5, respectively (Figure 3.13A). The relatively high number of differentially expressed genes at day 0 is likely due to the precocious expression of *AG* in the FIS (data not shown). At day 2.5, a similar number of genes as in the 0 day time point were up and down-regulated, whereas the vast majority of genes were down-regulated at day 5. 93 genes are differentially expressed only at day 0, 1 gene is differentially expressed only at 2.5 and 360 genes are differentially expressed only at day 5 (Figure 3.13B). The number of genes that are differentially expressed at both day 2.5 and day 5 is 84, which is marginally higher than the 64 genes differentially expressed at both day 0 and day 5. 60 genes in total are differentially expressed at all three time-points (Figure 3.13B).

Next, the distribution of Gene Ontology (GO) terms among these differentially expressed genes was analysed to obtain a better understanding of their roles in flower development. To this end, more general (higher level) GO terms that were largely redundant with lower level terms were removed in order to reduce the required number of statistical tests and to simplify the output of the analysis. Biological processes that were under or over-represented in the microarray dataset were identified by comparing the occurrence of GO terms within the lists of differentially expressed genes to their genome-wide distribution using a two-sided Fisher's exact test. The resulting *p*-values were then adjusted to correct for multihypothesis testing using the method by Benjamini and Hochberg (Hochberg and Benjamini, 1990). GO terms that had an adjusted *p*-value < 0.05 were judged as having a statistically significant distribution in the dataset. This analysis revealed 54 GO terms that were enriched among the differentially expressed genes identified in the day 5 time point (see Table 3.2). In contrast, no GO terms were found to be enriched or under-represented at day 0 or day 2.5. The GO terms identified at day 5 included, for example, *Flower Development*, *Carpel Development*, *Gynoecium Development*, and *Ovule Development*, all of which correlate well with the known role of *AG* in flower development. I further categorised these GO terms according to the biological processes they represent (Table 3.2). This was done because many of the GO terms (such as *Post-Embryonic Morphogenesis*, for example) are somewhat vague and can hamper a meaningful interpretation of the data. GO terms were classified using the parent categories of *Flower Development*, *Epigenetic Regulation*, *Patterning*, *General Development* and *General Cellular Processes*.

In agreement with the role of *AG* in floral organ specification, 5 of the top 10 GO

Table 3.2 Gene Ontology terms significantly enriched among the genes identified as differentially expressed in *ag* mutant flowers. Data of the 5d time point were used for the analysis. An adjusted (Benjamini-Hochberg method for multihypothesis testing) *p*-value cut-off of <0.05 was used to identify over-represented Gene Ontology (GO) terms. The column “*p*-value Rank” separates the most significant *p*-value (Rank 1) from the least significant (Rank 54). The GO terms were categorised into five groups that broadly represent the biological processes involved.

Flower Development		General Cellular Processes	
p-value Rank	GO-term	p-value Rank	GO-term
1	gynoecium development	3	multicellular organismal development
2	carpel development	5	regulation of transcription
4	flower development	7	regulation of transcription, DNA-dependent
8	ovule development	19	cell cycle
9	organ formation	22	anatomical structure formation in morphogenesis
17	floral organ morphogenesis	27	negative regulation of cellular process
23	fruit development	28	negative regulation of transcription, DNA-dependent
24	floral organ formation	31	negative regulation of transcription
32	embryo sac development	33	regulation of cell cycle
38	pollen development	40	regulation of cellular macromolecule biosynthetic process
Patterning		43	negative regulation of cellular biosynthetic process
p-value Rank	GO-term	46	one-carbon metabolic process
10	regionalization	47	protein methylation
11	pattern specification process	52	cell differentiation
22	anatomical structure formation in morphogenesis	53	negative regulation of cellular metabolic process
42	formation of organ boundary	54	cell fate commitment
44	anatomical structure arrangement	General Development	
49	regulation of anatomical structure morphogenesis	p-value Rank	GO-term
Epigenetic Regulation		13	embryo development ending in seed dormancy
p-value Rank	GO-term	14	embryo development
6	chromatin modification	15	seed development
12	histone modification	16	shoot development
18	chromatin silencing	26	shoot morphogenesis
20	DNA methylation	29	biological process
21	DNA modification	30	post-embryonic morphogenesis
25	negative regulation of gene expression, epigenetic	36	organ morphogenesis
34	histone lysine methylation	39	meristem structural organization
35	histone methylation	41	embryonic meristem development
37	methylation	45	secondary shoot formation
51	regulation of gene expression, epigenetic	48	cotyledon development
		50	primary shoot apical meristem specification

terms (GO terms were ranked based on the associated *p*-values, the lowest *p*-value is ranked 1 whereas the highest *p*-value < 0.05 is ranked last) were part of the *Flower Development* category. Surprisingly, this category contained only one stamen-specific GO term namely *Pollen Development*. However, the terms *Stamen Development* and *Anther Development* were only marginally above the *p*-value cut-off of 0.05 (*p* = 0.059 and 0.065, respectively), indicating some degree of enrichment of genes involved in these processes in the microarray dataset. The second category, *General Cellular Processes*, contained three of the top 10 GO terms including the *Regulation of Transcription*. This result suggests that AG regulates organ development largely by controlling the expression of other transcription factors. A third category, *Patterning*, contained GO terms such as *Formation of Organ Boundary* in agreement with a previous study that suggested a role for AG in floral organ

boundary formation (Breuil-Broyer et al., 2004). The category *General Development* did not contain any of the top 10 GO terms. In fact, most of the GO terms within this category were comparatively general although *Meristem Structural Organisation* included genes such as *SHOOT MERISTEMLESS (STM)*, which is essential for meristem maintenance.

An unexpectedly large number of GO terms were assigned to the category *Epigenetic Regulation*. This category contained 10 GO terms altogether and one of the top 10 GO terms (*Chromatin Modification*). This finding suggested a direct role of AG in the epigenetic regulation of flower development (see Section 3.2.8.2 for a detailed discussion).

Although the GO term categories discussed above provided valuable information regarding the processes that AG is involved in, this type of analysis does not provide information about the directionality of the observed gene expression changes. For example, it does not say anything about whether genes involved in the epigenetic regulation of development are predominantly repressed or activated by AG. To address this point, a gene set enrichment analysis (Wu et al., 2010) was performed (Figure 3.14). This global approach allows the identification of coordinated shifts towards positive or negative fold-change values within predefined sets of genes. These gene sets can be defined by co-expression across many different transcriptome studies or, as in this case, are based on their involvement in the same biological process as defined by their assignment to the same GO term. Several of the gene sets that were analysed displayed clear directionality and all of them were down-regulated in the data from the *ag* microarray experiment. For example, the term *Formation of Organ Boundary* was down-regulated at the day 2.5 time point. At day 5, all GO terms included in the *Flower Development* category described above (with the exception of *Fruit Development*), were significantly down-regulated at day 5. Four of the five GO terms within the *Patterning* category were also considered significantly down-regulated (*Regionalization*, *Pattern Specific Process*, *Anatomical Structure Formation in Morphogenesis and Formation of Organ Boundary*). Furthermore, all GO terms except *Regulation of Gene Expression*, *Epigenetic* were repressed within the *Epigenetic Regulation* category. Within the category *General Cellular Processes* the GO term *Negative Regulation of Transcription* was down-regulated. This suggests that AG might positively regulate transcriptional repressors.

Several other terms that were not found to be over-represented in the initial GO-term analysis described above were judged significantly down-regulated in the gene set

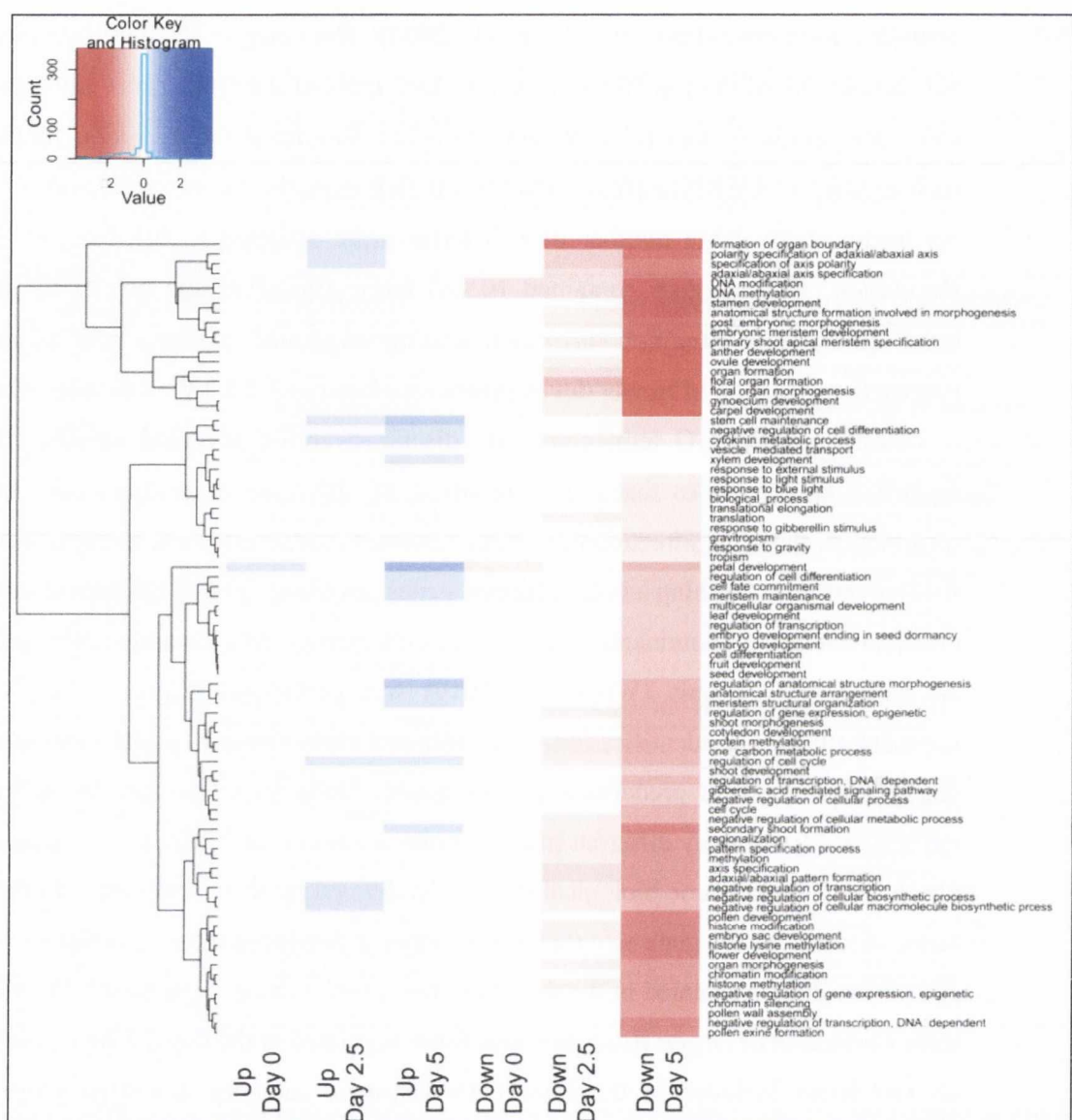


Figure 3.14 Gene set enrichment analysis using the data from the *ag* mutant experiments. The heatmap represents p -values obtained from a gene-set enrichment analysis (ROAST: rotation gene set test; (Wu et al., 2010)) of the microarray dataset stemming from the comparison of *ag* mutant and wild-type flowers at different stages of development. The method determined whether gene expression within the GO terms defined in Figure 3.10 is shifted towards positive (“up-regulation”, blue) or negative (“down-regulation”, red) fold-changes and resulted in the assignment of two p -values for each term (one p -value for “up-regulation” and one for “down-regulation”). The negative decadal logarithms of these p -values are plotted. GO terms were clustered based on euclidean distance calculations.

analysis. These terms included *Anther Development*, *Stamen Development* and *PollenWall Assembly*, as well as *Adaxial/Abaxial Axis Specification*, *Polarity Specification of Adaxial/Abaxial Axis* and *Adaxial/Abaxial Pattern Formation*. The latter terms relate to

Table 3.3 Protein families enriched among genes differentially expressed in *ag* mutant flowers. Protein families that were enriched in the *ag* microarray dataset, relative to their genome-wide distribution, were identified using a two-sided Fisher’s exact test. A Benjamini-Hochberg adjusted *p*-value cut-off of < 0.05 was used a significance level for protein family enrichment.

Transcription Factors	Gene Expression Regulation - Other
B3 domain	PAZ domain
Homeodomain Fold	SET domain
C2H2 Transcription Factor Family	Helicase conserved C-terminal domain
Other	SNF2 family N-terminal domain
Pentatricopeptide repeat	Cell Cycle
BTB/POZ domain	Cyclin
Protein Kinase Domain	

genes such as *CRC*, *INNER NO OUTER* and *KANADI2*, all of which have been shown to be important for carpel or ovule patterning (Eshed et al., 1999; Villanueva et al., 1999).

Taken together, the prevalence of down-regulated genes in the *ag* mutant microarray dataset suggests that AG acts largely as a transcriptional activator. Furthermore, the relatively large number of GO terms enriched in the *ag* mutant dataset implies that AG is involved in controlling a multitude of cellular and developmental processes.

In plants, gene families that are derived from gene or genome duplications are often involved in the same biological process. For example, all genes (with the exception of *AP2*) that are part of the ABC model of floral organ identity specification code for MADS-box transcription factors (Parenicova et al., 2003). Therefore, the *ag* microarray dataset was investigated for an over-representation of gene or protein families using the same approach as for the identification of enriched GO terms (Table 3.3). This analysis revealed 12 protein families that were significantly (Benjamini-Hochberg adjusted *p*-value < 0.05) over-represented. Of these 12 families, three contained transcription factors that include known floral regulators such as *VERDANDI* (a B3-domain transcription factor), *STM* (a homeodomain transcription factor) and *KNU* (a C2H2 zinc-finger transcription factor). Four other protein families that were enriched in the datasets also contain proteins involved in the regulation of gene expression, but do not encode transcription factors. These families included the PAZ domain protein family, the SET domain protein family, the Helicase conserved C-terminal domain family (Helicase-C) and the SNF2 domain family. The PAZ

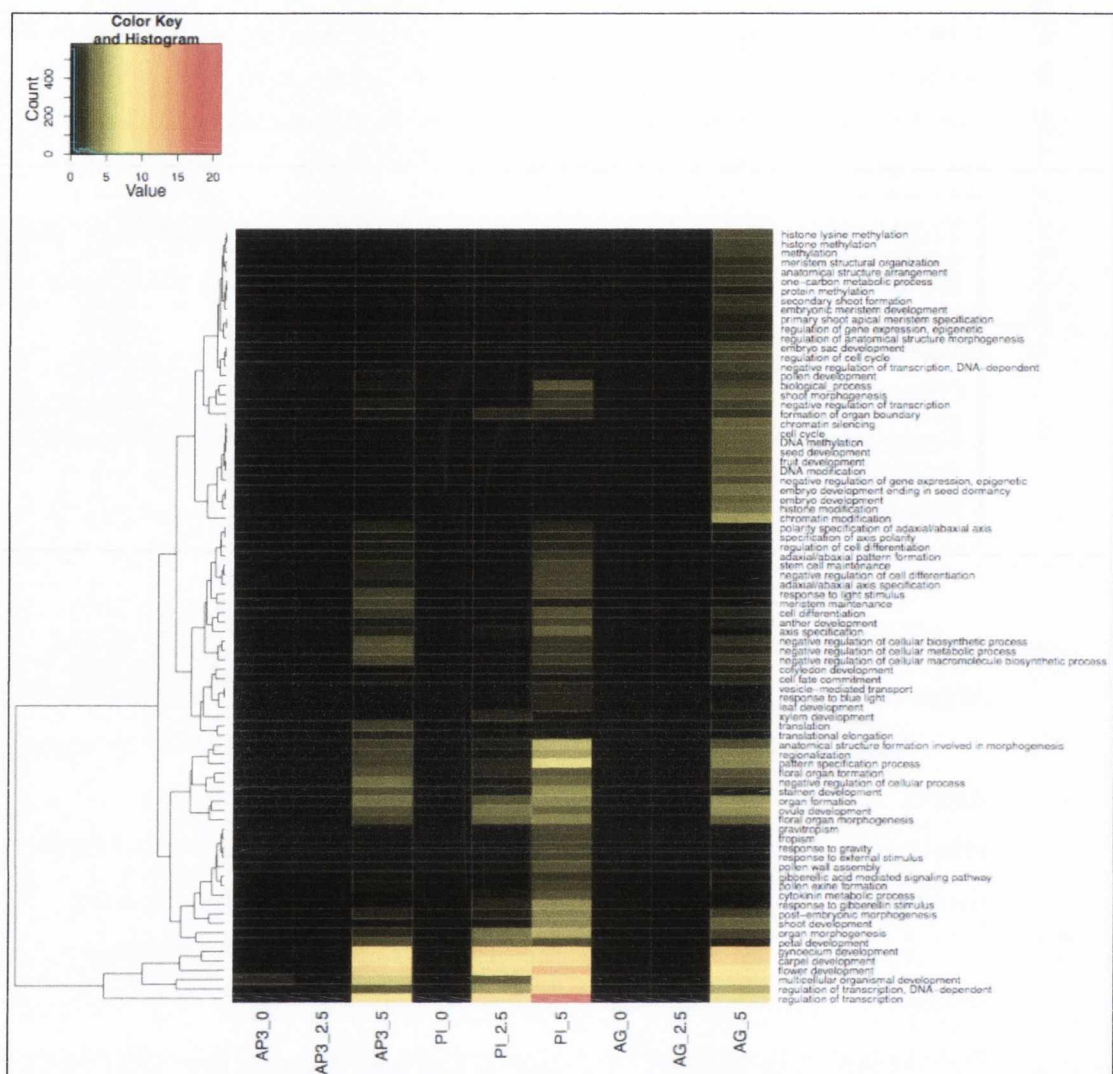


Figure 3.15 A heatmap of Gene Ontology (GO) terms found to be enriched among differentially expressed genes identified in *ap3*, *pi* and *ag* microarray experiments. The heatmap represents the significance levels for enrichment of GO-terms within the top differentially expressed genes in mutant-wildtype comparisons. The negative decadal logarithms of the adjusted *p*-values (Fisher's exact test) for each GO term and time point (as indicated) are expressed as coloured bars (see upper right corner of the figure for a colour key). Black indicates that a term was not over-represented, while yellow or red mark an increasingly significant enrichment.

domain protein family includes the *ARGONAUTE* proteins, which have been implicated in post-transcriptional gene-silencing and RNA-dependent DNA methylation (Havecker et al., 2010). The SET domain protein family are lysine methyltransferases involved in histone modification which is important for the regulation of chromatin structure and gene expression (Dillon et al., 2005; Thorstensen et al., 2011). The SNF2 protein family is

involved in chromatin modification and remodelling (Knizewski et al., 2008). Members of the Helicase-C protein family are involved in cytosine methylation and the DDM1 protein, which is down-regulated in the *ag* mut/WT contrasts, is Helicase-C homolog of the SNF2 family (Bourc'his and Bestor, 2002). Some core cell cycle proteins, namely members of the cyclin protein family, were also significantly enriched in the dataset. Moreover, protein kinase domain containing proteins were misrepresented. These proteins are involved in a variety of developmental processes and signalling pathways. One differentially expressed protein kinase identified in this study is *ERECTA-LIKE1*, which is functionally redundant with the *ERECTA* and *ERECTA-LIKE2*, and is important for cell-cell communication and organ growth (Shpak et al., 2003). Taken together, the results of this analysis indicate that *AG* function depends on the control of sets of related genes that mediate diverse processes such as the regulation of transcription, protein modification and epigenetic mechanisms.

3.2.4.3 Interplay between *AG* and *B* function genes

The data presented above have provided valuable information regarding the role of *AG* in flower development. In order to identify functions that *AG* has in common with the *B* function genes *AP3* and *PI*, or that are unique to *AG*, I compared the results of the GO term analysis of the *ag* mutant dataset (see Section 3.2.4.2 above) with those from the *ap3* and *pi* experiments conducted by Samuel Wuest (Wuest, 2011) using the same experimental set-up as for the analysis of *AG* function. The procedure for the GO term analysis was identical for all three transcriptome studies performed, ensuring their direct comparability. It revealed that certain GO terms were significantly (Benjamini-Hochberg adjusted *p*-value < 0.05) enriched in all mutants/wild type comparisons (Figure 3.15). These included *Flower Development*, *Regulation of Transcription* and *Organ Formation*, in agreement with the known role of the floral homeotic genes in flower development. Additional GO terms that were enriched in the datasets of all mutant experiments include *Gynoecium Development*, *Carpel Development* and *Ovule Development* as well as GO terms within the category *Patterning* (see Table 3.2). This suggests that *B* and *C* class genes regulate genes involved in female reproductive organ development and are involved in regulating boundaries between floral organs. In contrast to the GO terms discussed above, terms within the *Epigenetic Regulation* category (see Table 3.2) were specific to *AG*.

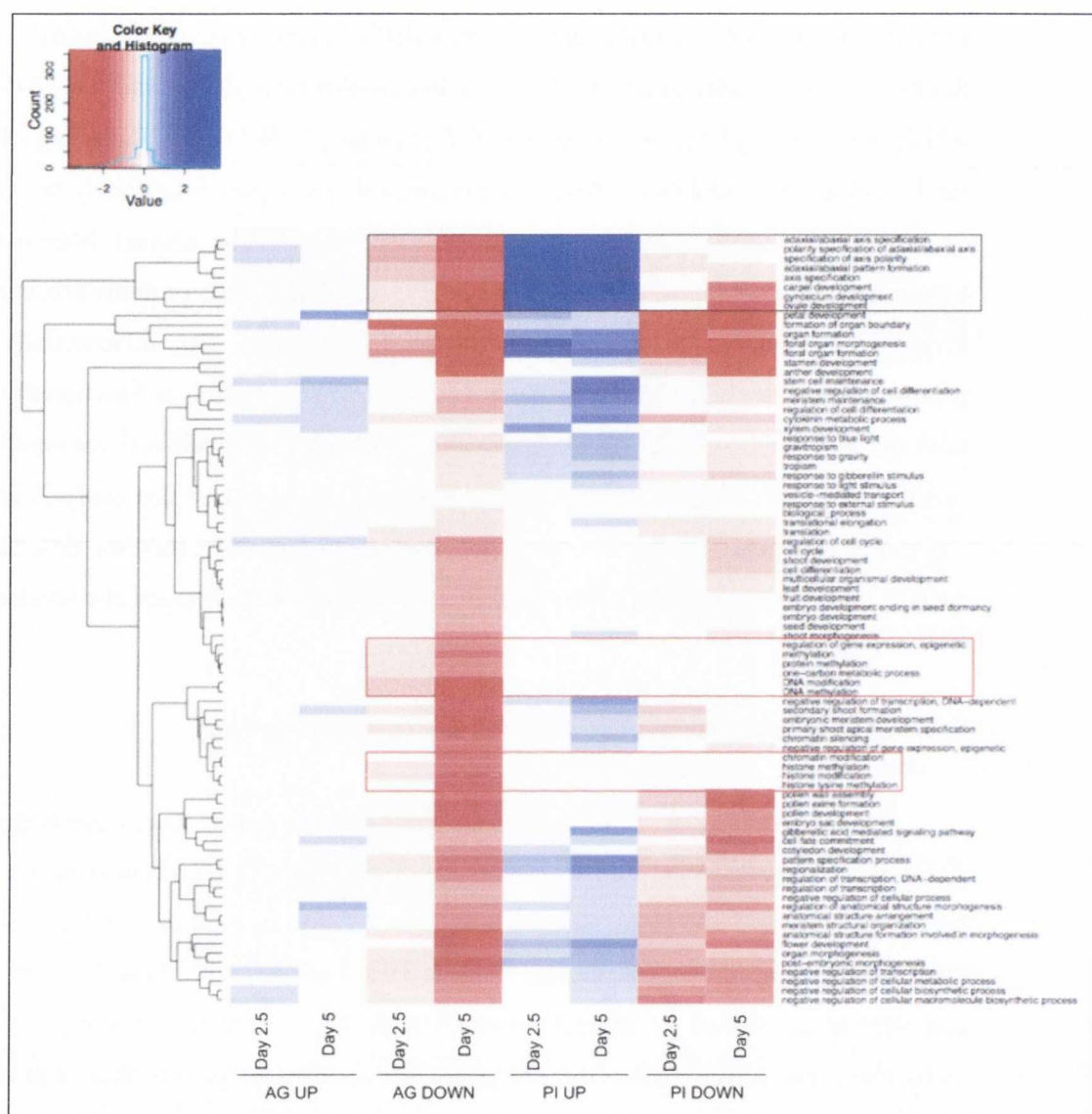


Figure 3.16 Gene set enrichment analysis for the datasets from the *ag* and *pi* microarray experiments.

The heatmap represents p -values obtained from a gene-set enrichment analysis (ROAST: rotation gene set test; (Wu et al., 2010) of the *ag* and *pi* microarray datasets. The method determined, for each mutant and time point separately, whether gene expression within a given GO term was shifted towards positive (“up-regulation”; indicated in blue) or negative (“down-regulation”; indicated in red) fold-changes and resulted in the assignment of two p -values for each term (one p -value for “up-regulation” and one for “down-regulation”). The negative decadal logarithms of these p -values are plotted. GO terms were clustered based on euclidean distance calculations.

As discussed above, the GO term analysis does not imply any directionality, therefore a gene set enrichment approach was applied to the data from the *pi* mutant experiment and the resulting dataset was compared with those of the *ag* dataset. The data of the *ap3*

experiments were not analyzed in this case as they appeared to be more variable than those of the *pi* experiments. The results of the gene set enrichment revealed that terms such as *Stamen Development* were significantly (unadjusted $p < 0.05$) down-regulated in both mutant datasets (Figure 3.16), in agreement with the known function of B and C function genes in the formation of this organ type. In contrast, terms associated with female reproductive development were down-regulated in the *ag* dataset whereas these terms were up-regulated in the *pi* experiments. This suggests that B function genes play an active role in the repression of genes involved in the carpel program in the 3rd whorl. The terms associated with organ boundary formation were down-regulated in both mutants; however, a bidirectionality of expression was observed in the case of the *pi* experiments. Terms associated with axis polarity specification were down-regulated in the *ag* experiments but were up-regulated in the *pi* dataset. Moreover, these analyses highlighted that B function equally promotes and suppresses developmental programs whereas *AG* predominately promotes gene expression. In summary, B and C function genes appear to jointly influence developmental programs involved in male and female reproductive development, organ boundary formation and axis polarity specification; however, their effects are in some cases antagonistic.

In a next step, the gene and protein families that are regulated by B and C function genes were compared. To this end, the data for the B function mutants were analysed as described above (Section 3.2.4.2) for *AG*. Figure 3.17 shows that both B and C function gene sets are significantly (Benjamini-Hochberg adjusted p -value < 0.05 ,) enriched for members of the B3 domain transcription factors, C2H2 zinc-finger and homeodomain transcription factors. Gene families specifically enriched in the *pi* dataset included several additional transcription factor families, while the *ag* dataset was enriched for members of the PAZ domain, SET domain, SNF2 domain, BTB/POZ domain and PPR protein domain families. In addition, core cell cycle genes were specifically enriched in the *ag* dataset.

Finally, gene set enrichment analysis was applied to specific gene and protein families in the datasets of the *ag* and *pi* mutant experiments (Figure 3.18). In the *ag* dataset, protein families were only significantly down-regulated at day 5 ($p < 0.05$, unadjusted). The families that were only represented in the *AG* gene set enrichment analysis included the PAZ domain protein family, the SNF2 protein family and the Cyclin protein family. The B3 transcription factor family was down-regulated in the *AG* mut/WT contrast but up- and down-regulated in the *PI* mut/WT contrast. The Zinc-Finger transcription factor family was

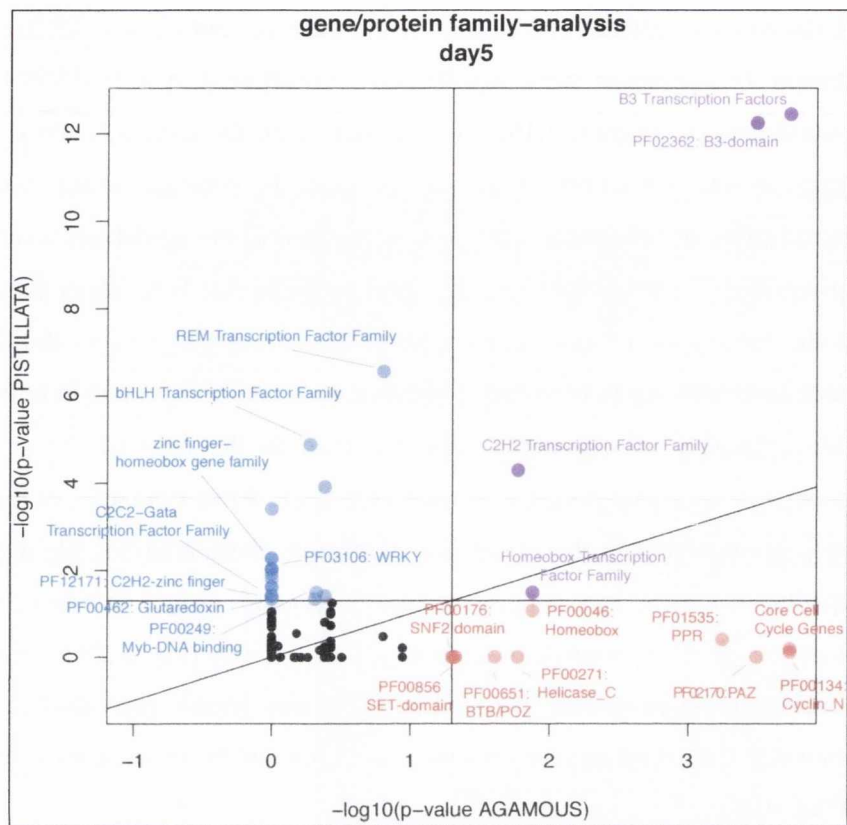


Figure 3.17 Comparison of the results of protein family analyses conducted for genes identified as differentially expressed in *pi* and *ag* mutant flowers. *P*-values (Fisher's exact test) for gene groups are shown for the 5 day time-point. *P*-values have been corrected for multihypothesis testing using the method by Benjamini and Hochberg and were $-\log_{10}$ -transformed. Data for the *AG* experiments are plotted on the *x*-axis, and data for the *PI* experiments on the *y*-axis. Colours indicate groups of significantly enriched protein families. Blue: significantly enriched in the *pi* experiment only; purple: significantly enriched in both experiments; red: significantly enriched in the *ag* experiment only. Selected data points are labelled.

up-regulated in the 2.5 day *pi* experiments but down-regulated in the day 5 *AG* experiment. Finally, the Rhomboid protein family, which are a class of proteases, were differentially represented at day 5 in both the *ag* and the *pi* experiments. This protein family was only present in the protein family analysis of the *pi* dataset.

Taken together, these analyses reveal a strong overlap in the processes that *AG* and *PI* regulate. In addition, these processes may be regulated by controlling the expression of distinct protein families. Whereas *AG* appears to support or activate developmental programs, *PI* appears to suppress and activate developmental programs. Whether these developmental programs are activated or repressed directly by the homeotic transcription

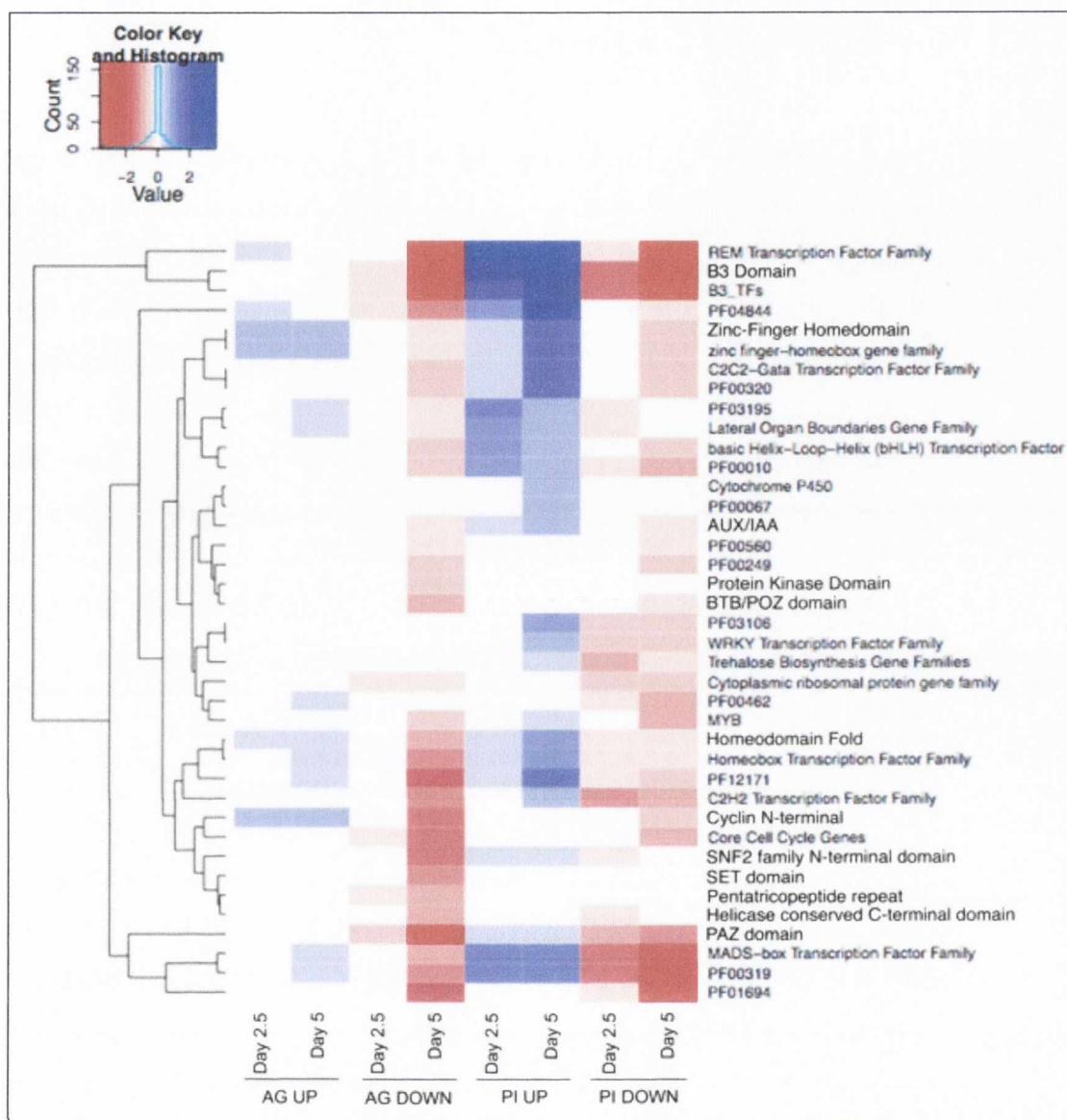


Figure 3.18 Comparison of gene- and protein families enriched in the datasets from the *ag* and *pi* microarray experiments. The heatmap represents gene set enrichment p -values for those gene/protein families that were identified as overrepresented amongst differentially expressed genes in both B- and C-class mutant experiments. P -values were $-\log_{10}$ (down-regulated) or $\log_{10}$ (up-regulated) transformed. Red indicates fold-change values that are shifted towards negative values and blue indicates fold-change values shifted towards positive values.

factors can be resolved by identifying the regions of the genome where AG, AP3 and PI bind.

3.2.5 Identification of genes acting downstream of AG through inducible gene perturbation

As discussed in Section 1.5, the use of an inducible perturbation system for the disruption of *AG* activity offers potentially several advantages over the *ag* mutant approach described in Section 3.2.4. In order to perform expression profiling experiments after a specific perturbation of *AG*, I first isolated an amiRNA (AG4-amiRNA) that leads to an efficient knockdown of *AG* mRNA levels (see Section 3.2.1) and to plants that form flowers resembling those of *ag* mutants (Section 3.2.2). This amiRNA was then introduced under the control of the *35S::AlcR/pAlcA* ethanol-inducible promoter systems into the floral induction system (Section 3.2.2.3) to allow the collection of large numbers of synchronized floral buds for genomic analysis. This line was subsequently used to identify genes that responded to a perturbation of *AG* activity. The experimental strategy and the results of the analyses of the microarray data are outlined below.

3.2.5.1 Experimental design and implementation

To analyse the effect of *AG* perturbation at early stages of flower development on a genome-wide scale, I used plants of the floral induction system containing the *35S::AlcR/pAlcA::AG4-amiRNA* transgene. These plants will be referred to as AG4-FIS hereafter. AG4-FIS plants were treated for 6 hours with ethanol vapour, or were mock-treated, and tissue from these plants was harvested 18 hours after the end of the treatment. The resulting samples constitute the 0 d time-point of the experiment. Subsequently, inflorescences of AG4-FIS plants (grown in parallel with the 0 d time-point) were treated with dexamethasone to induce flower formation. These plants were then EtOH or mock-treated as described above 1, 3, 5 and 7 days after the dexamethasone treatment (Figure 3.19). Tissue was harvested from these plants 18 hours after the end of the EtOH/mock treatment, resulting in the 2, 4, 6 and 8-day time-points. Total RNA was then extracted from each sample, cDNA was synthesized and qRT-PCRs were performed to confirm the ethanol-dependent

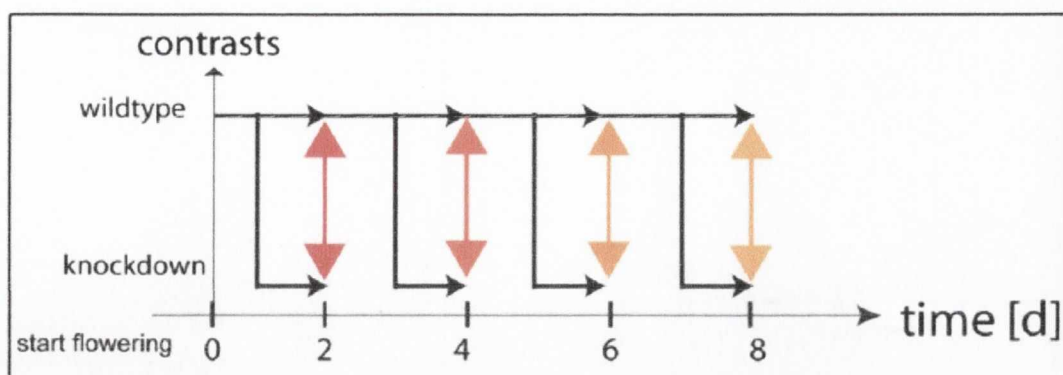


Figure 3.19 Experimental design of inducible *AG* perturbation experiments. Tissue was harvested 2, 4, 6 and 8 days after the initiation of flower development ('start flowering'). Prior to tissue collection, plants were either EtOH or mock-treated for 6 hours followed by an 18-hour recovery period. RNA isolated from EtOH and mock-treated plants was then co-hybridized to whole-genome microarrays.

knockdown of the *AG* mRNA (data not shown). Once the knockdown of the *AG* mRNA had been confirmed, RNA samples were amplified, labelled with fluorescent dyes and then hybridized to custom-designed Agilent whole-genome *Arabidopsis* microarrays as per manufacturer's instructions. Data for four biologically independent sets of samples were generated.

Mock-treated plants were used as a developmental control for EtOH-treated plants. Additionally, these plants provided a control for the presence of the *35S::AlcR/pAlcA::AG4-miRNA* transgenes, however, they did not control for any interactions between ethanol and the *AlcR* transcription factor. It is possible that the *AlcR* protein influences the expression of endogenous *Arabidopsis* genes when activated with ethanol. Therefore, a 21-nucleotide sequence was placed under the control of the *35S::AlcR/pAlcA* ethanol-inducible promoter system by Samuel Wuest (Wuest, 2011). This sequence (referred to as a DUMMY-miRNA hereafter) was specifically designed (using the WMD3-microRNA design tool as well as BLAST analyses) not to target any *Arabidopsis* transcripts. A selected DUMMY-miRNA line was crossed into the floral induction system and all mutations and transgenes were isolated in a homozygous state. Quantitation of *AlcR* expression levels in this line revealed that they were lower than in either the *pAlcA::AP3-miRNA* (Wuest, 2011) or *pAlcA::AG4* containing lines. Therefore, an

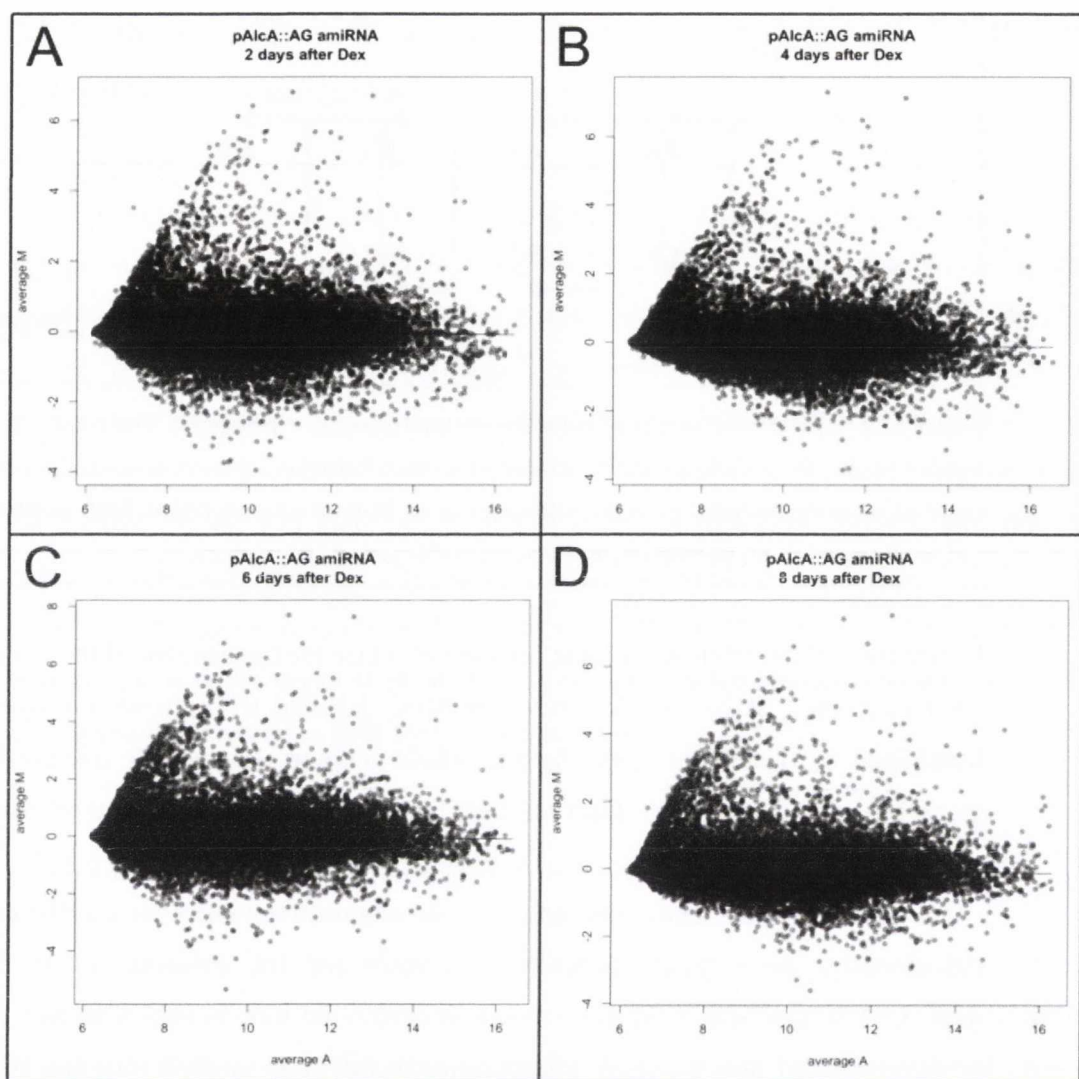


Figure 3.20 Microarray data of the AG4-amiRNA time course experiments. Mean expression values (x-axis) are plotted against log₂-transformed expression ratios (EtOH vs. mock; y-axis). The data shown are based on the results of hybridizations of four sets of biologically independent samples. **(A)** Data for the 2 day time point. **(B)** Data for the 4 day time point. **(C)** Data for the 6 day time point. **(D)** Data for the 8 day time point.

alternative DUMMY-amiRNA with high *AlcR* expression levels is being prepared. This line has not yet been isolated in a homozygous state. Therefore, a preliminary control experiment was conducted using the available DUMMY-amiRNA line. Plants of this line were treated with EtOH or were mock-treated for 24 hours one day after dexamethasone-treatment.

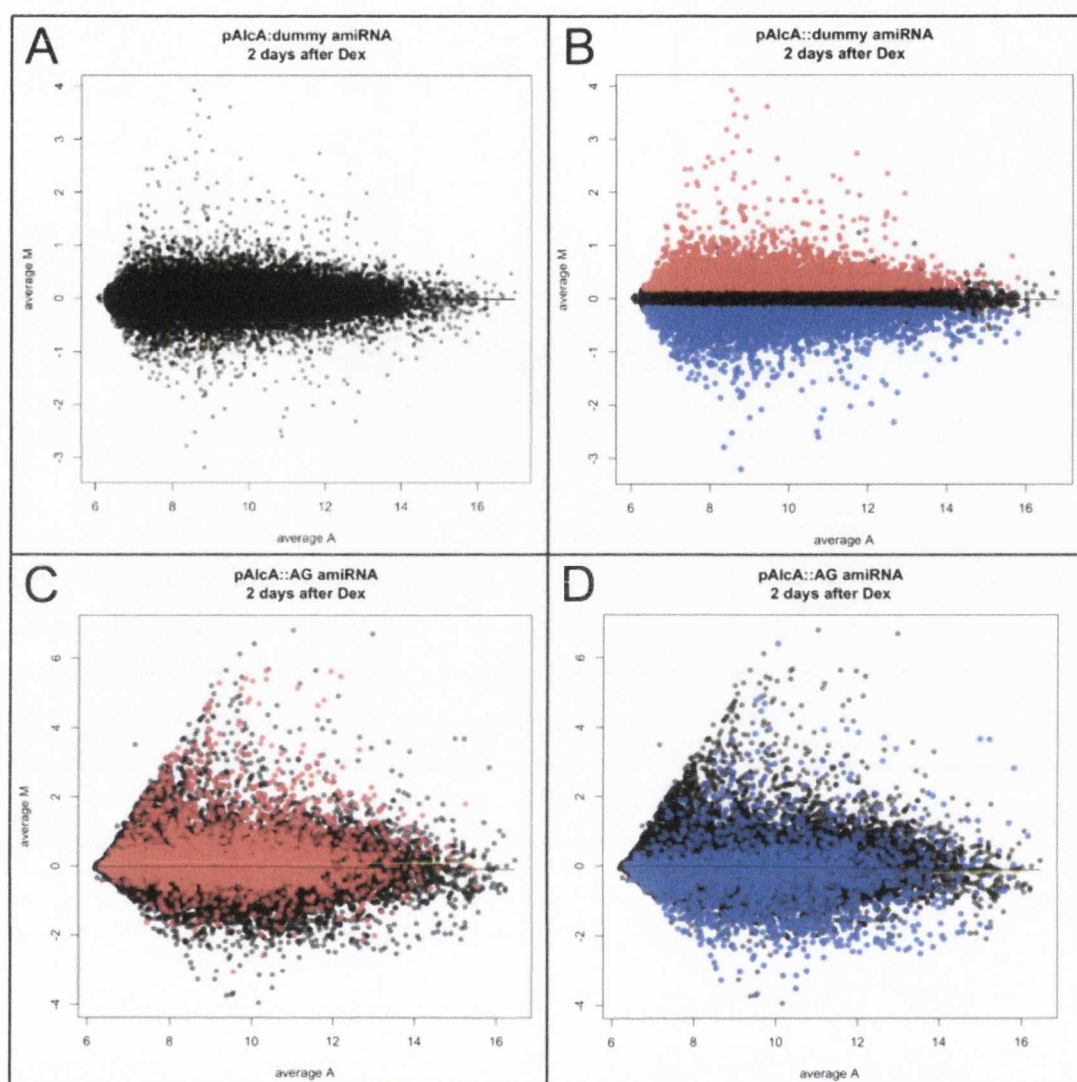


Figure 3.21 Filtering of microarray data from the inducible *AG* perturbation experiments through use of a DUMMY-amiRNA. (A) Microarray data for the 2 day time point of the DUMMY-amiRNA experiments. Mean expression values (x-axis) are plotted against log₂-transformed expression ratios (EtOH vs. mock; y-axis). (B) Plot as in (A), but genes that were differentially expressed in the DUMMY-amiRNA experiments (unadjusted *p*-value < 0.1 and more than 10% expression change) are highlighted. Red indicates up-regulation and blue down-regulation. (C – D) Genes that were considered up-regulated (C) and down-regulated (D) in the DUMMY-amiRNA experiments are projected onto the data of the *AG* knockdown experiments described in Figure 3.20. Data for the 2 d time point are shown. Locally weighted regression lines are drawn for all genes (black) or for genes that are up- or down-regulated in the DUMMY-amiRNA dataset (yellow), respectively.

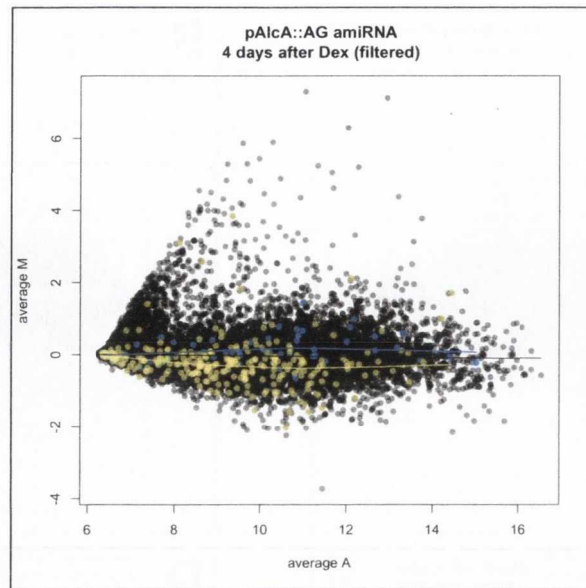


Figure 3.22 Comparison of microarray data from *ag* mutant and AG4-amiRNA knockdown experiments. The genes differentially expressed at day 5 in the *ag* mutant experiment were projected onto the microarray data from the 4 day time point of the AG4-amiRNA knockdown experiment. The latter data were filtered using the results of the DUMMY-amiRNA experiments and genes that were judged differentially expressed in both experiments were removed from the MA-plot. Genes considered up- (blue) or down-regulated (yellow) in the *ag* mutant experiment are highlighted. Locally weighted regression lines are drawn for all genes (black) or for genes that are up- (blue) or down-regulated (yellow) in the *ag* experiments. Average expression values (x-axis) are plotted against $\log_2$ -transformed mean fold-changes (y-axis) of four biological replicates.

Tissue was harvested immediately after the EtOH-treatment. Total RNA was then extracted, amplified, labelled with fluorescent dyes and hybridised to custom-designed Agilent whole-genome *Arabidopsis* microarrays. Two biologically independent sets of samples were generated.

To analyse the effect of *AG* perturbation at later stages of flower development on a genome-wide scale, I used inflorescences from plants containing the *pAlcA::AG4-amiRNA* transgene in a wild-type background. Plants were EtOH or mock-treated for 6 hours and tissue was harvested after an 18-hour recovery period. Older flowers were removed from the inflorescences in order to analyse up to stage 10 of flower development. Four independent biological replicates have been generated, processed and hybridised to custom-designed Agilent whole-genome *Arabidopsis* microarrays. The analysis of these data has not been completed. Below, I will present the results of the analyses of the FIS-AG4 and FIS-DUM experiments.

3.2.5.2 Characterization of AG activity using the inducible AG4-miRNA perturbation system

The analytical procedures were identical to the *AG* mut/WT analytic procedures in Section 3.2.4 (analysis performed by Samuel Wuest). Initially a global view of gene expression levels was obtained where signal intensities and log₂-transformed fold changes were averaged over all four biological replicates and then plotted (Figure 3.20). Strong ethanol-dependent effects were observed at all time-points and more genes appeared to be up-regulated than down-regulated. The fold-change values were higher for up-regulated than down-regulated genes. Moreover, up-regulated genes were shifted towards lower average expression values. These results were not consistent with those of the *ag* mutant experiments where relatively small fold-change differences were observed. Furthermore, the majority of genes considered differentially expressed in the *ag* mutant experiments were down-regulated and not up-regulated as in this case.

Next, a global view of gene expression level changes in the FIS-DUM experiments was observed (Figure 3.21). The gene expression changes observed indicated that many genes were responding to ethanol treatment in the FIS-DUM flowers (Figure 3.21A). Differentially expressed genes within the FIS-DUM data were then defined using the *limma* linear modelling approach (unadjusted $p < 0.1$ and more than 10% expression change). These differentially expressed genes (6,842 in total) are highlighted on the MA-plots of the FIS-AG4 2 d EtOH/mock datasets (Figure 3.21C, D). There was a clear overlap between up-regulated genes (Figure 3.21C) and down-regulated genes (Figure 3.21D) in both the FIS-AG4 and FIS-DUM datasets. Moreover, up- and down-regulated genes in the FIS-DUM exhibited a slight global fold-change shift, respectively, indicating that a systematic bias is present (Figure 3.21C, D, yellow lines). This suggested that many of the genes that were differentially expressed in the FIS-AG4 experiments were due to the combined effects of the *35S::AlcR/pAlcA* expression system in addition to the ethanol treatment.

The genes considered differentially expressed in the FIS-DUM experiments were then removed (“filtered”) from the AG4-FIS day 4 analyses. This filtered dataset was then compared with the genes considered differentially expressed in the day 5 time-point of the *ag* mutant experiments (Figure 3.22). The majority of genes identified as significantly up-

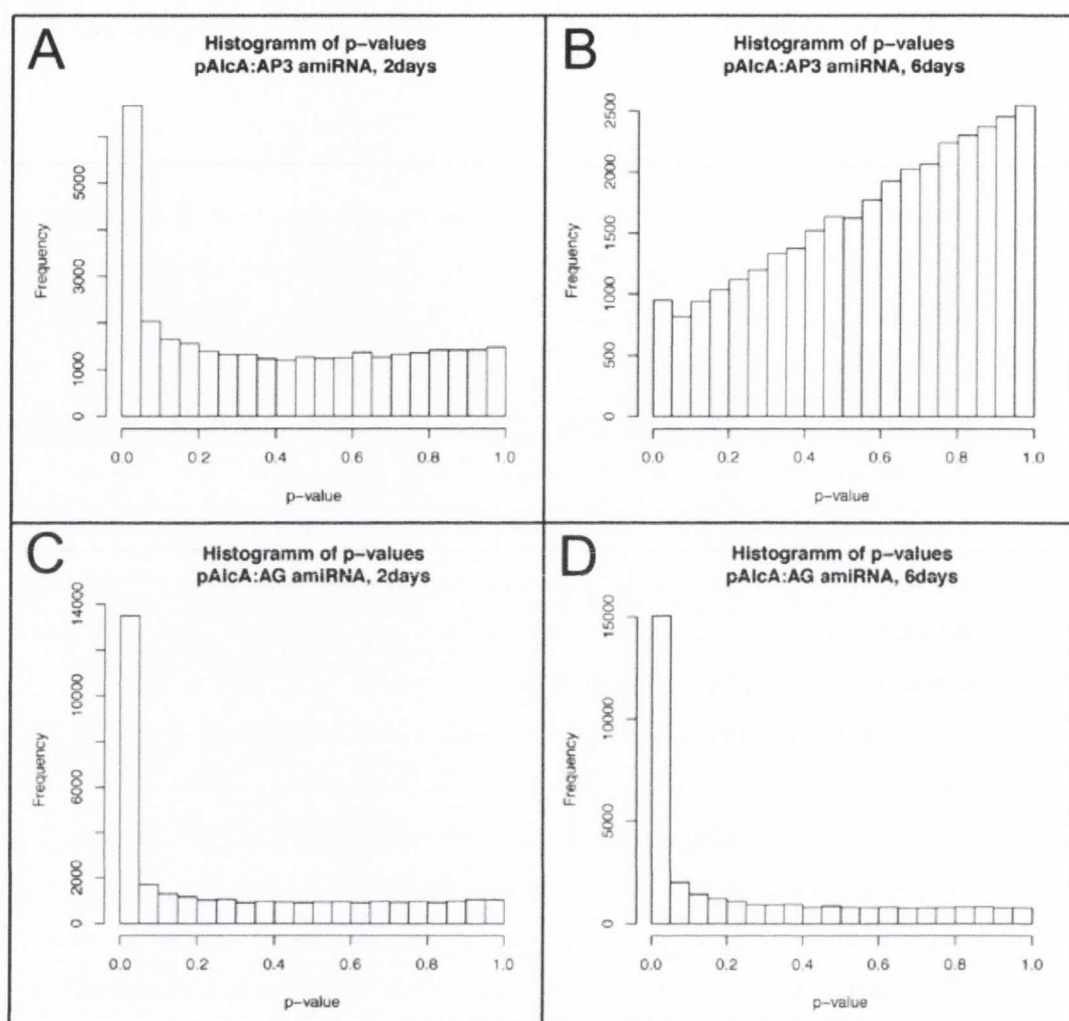


Figure 3.23 Distribution of p -values in the microarray datasets of the inducible *AP3* and *AG4-amiRNA* experiments. P -values were determined by testing for differential gene expression between EtOH and mock-treated plants using the linear modelling approach of Smyth et al. (2004) combined with Bayesian statistical analysis. P -value histograms are shown for the 2 d (**A** and **C**) and 6 d (**B** and **D**) time points of both the *AP3-amiRNA* (**A** and **B**) and the *AG4-amiRNA* (**C** and **D**) experiments.

or down-regulated in day 5 of the *ag* mutant experiments displayed the same directionality in the FIS-AG4 day 4 experiments. However, a number of genes that were considered down-regulated in the 5 day time point of the *ag* mutant experiments were up-regulated in the FIS-AG day 4 experiments (Figure 3.22). These results indicated that the ethanol-mediated knockdown of *AG* function resulted in the differential expression of many more genes than were identified in the *ag* mutant experiments.

Differentially expressed genes in the FIS-AG4 datasets were then identified by fitting a linear model to each gene and then applying the empirical Bayes method of Smyth *et al.* (2004). These analyses were performed using relaxed cut-off values ($p < 0.01$). Next, processes that were misrepresented were identified by GO term using stringent cut-off values (Benjamini and Hochberg adjusted p -value < 0.001). The *AP3-miRNA* knockdown experiments performed in the floral induction system (performed by Samuel Wuest) are also included here for comparison (denoted FIS-AP3 hereafter) (Wuest, 2011). P -value histograms were generated using the Bioconductor package *limma* (Figure 3.23). These histograms indicated that a large number of genes were differentially expressed in the FIS-AG4 experiments at the 2 d and 6 d time-points (Figure 3.23C, D). This trend was also seen at the 4 d and 8 d time-points (not shown). In the FIS-AP3 experiments, a similar degree of gene expression changes between EtOH and mock-treated plants was observed at the 2 d and 4 d time-points (Figure 3.23A and 4 d, not shown), but not after 6 (Figure 3.23B) or 8 d (not shown). Taken together, these results suggest that the ethanol treatment of these plants produced a large expression response even after filtering, except in the case of the later time-points of the FIS-AP3 experiments.

The GO terms that were identified as being enriched are visualised in Figure 3.24. The GO terms represented have been categorised using the parent terms *Developmental Processes*, *Stress and Pathogen Processes* and *General Processes*. The differentially represented FIS-AP3 GO terms are almost exclusively associated with *Stress and Pathogen Processes* whereas the differentially represented FIS-AG GO terms are associated with this parent term and the other two parent categories. Thirty-nine GO terms were assigned to the *Developmental Processes* category. Of these terms, 11 were also significantly enriched in the 5 d time-point of the *ag* mutant experiments. These terms included *Flower Development*, *Pattern Specification Process*, *Regionalisation*, *Post-Embryonic Morphogenesis*, *Regulation of Cell Cycle*, *Fruit Development*, *Seed Development*, *Cell Cycle*, *Shoot Development*, *Cell Differentiation* and *Embryo Development Ending in Seed Dormancy*. GO terms of interest contained in the *Developmental Processes* category that were not observed in the *ag* mutant experiments included *Regulation of Structural Meristem Organisation*, *Chromatin Assembly or Disassembly* and *Nucleosome Assembly*. The latter two terms possibly highlight the involvement of *AG* in epigenetic regulation as described in Section 3.2.4.2 and 3.2.4.3.

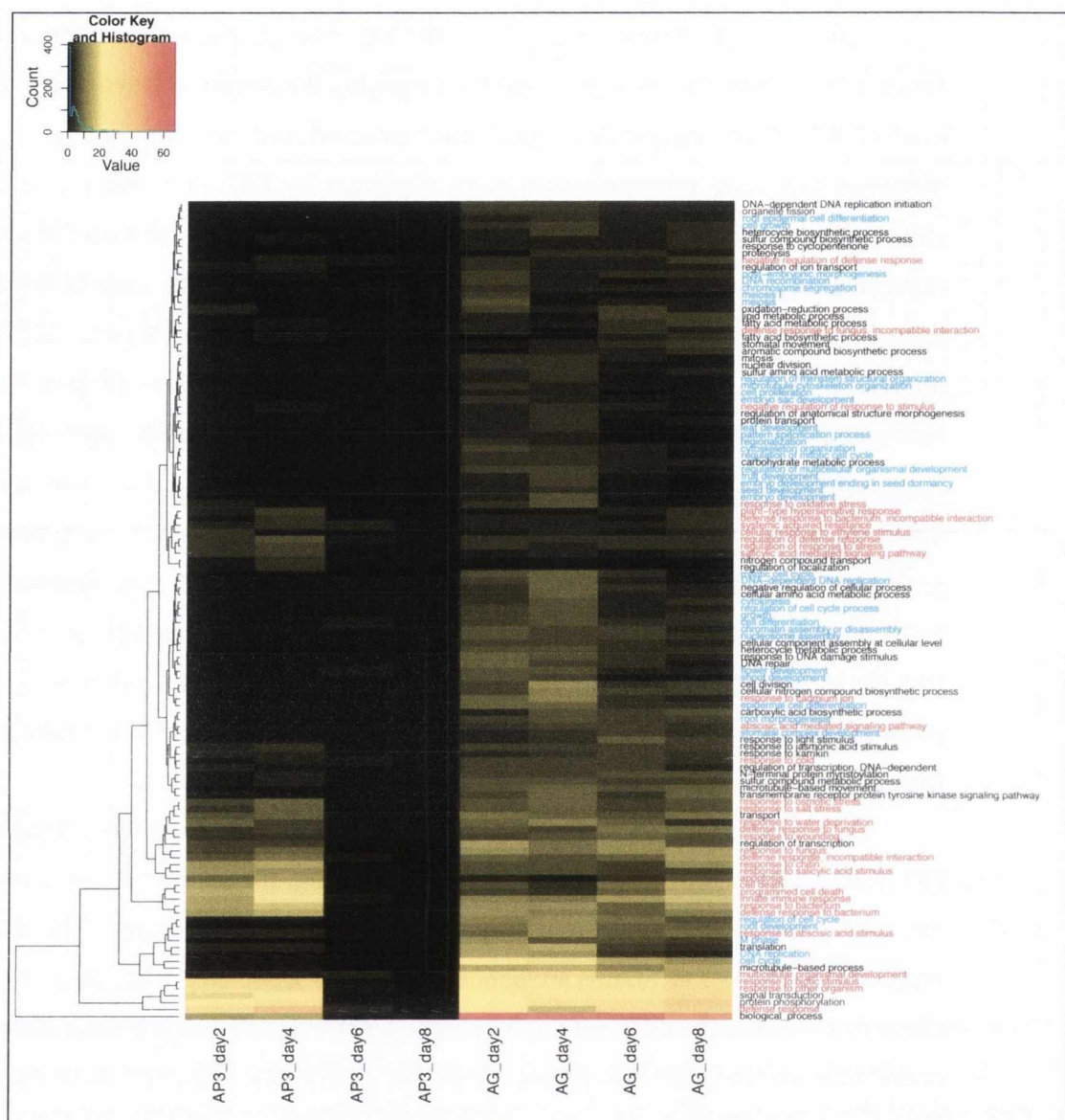


Figure 3.24 A heatmap of Gene Ontology (GO) terms found to be enriched among differentially expressed genes identified in the *AP3-amiRNA* and *AG4-amiRNA* microarray experiments. The heatmap represents the significance levels for an enrichment of different GO terms represented among the differentially expressed genes identified in the microarray experiments. The negative decadal logarithms of the adjusted *p*-values (Fisher's exact test) for each GO term and time point (as indicated) are indicated by coloured bars (see upper right corner of the figure for a colour key). Black indicates that a term was not over-represented, while yellow or red mark terms that are significantly enriched. GO terms have been assigned to one of the following categories: *Developmental Processes* (printed in blue), *Stress and Pathogen Processes* (in red) and *General Processes* (in black).

In summary, the combined effects of ethanol treatment and the activity of the AlcR transcription factor have resulted in the differential expression of many endogenous

Arabidopsis genes. The activated AlcR transcription factor mediates a systemic global shift of gene expression and appears to activate processes associated with stress response. We attempted to compensate for these effects by generating a transgenic line that drives a “DUMMY-miRNA” (see above) using the *35S::AlcR/pAlcA* expression system, however, the genes filtered using this control line were not sufficient to completely mask the non-specific effects of the activated AlcR transcription factor. Therefore, a more appropriate control line is currently being generated that expresses the AlcR transcription factor at similar levels as in the *AG4-miRNA* and *AP3-miRNA* containing lines.

Nevertheless, some information regarding the developmental processes regulated by *AG* has been garnered from this data. Some specific developmental processes identified as misrepresented in the *ag* mutant experiments have been confirmed in these experiments. Moreover, novel GO terms associated with epigenetic regulation have been identified. However, these data should be treated with caution until microarray data is available from an appropriate control line.

3.2.6 Establishment of a ChIP protocol

Chromatin immunoprecipitation (ChIP) is a powerful but complex technique used to identify the interactions between proteins and DNA (Kaufmann et al., 2010a). Because ChIP assays had not been previously performed in the Wellmer laboratory, I established and optimized a ChIP protocol that is suitable for the genome-wide detection of AG binding sites. This work was done in collaboration with Dr. Emmanuelle Graciet and Samuel Wuest.

Using epitope-tagged transgenic plant lines in the FIS, two separate protocols were tested and optimized (named ‘protocol 1’ and ‘protocol 2’ hereafter). Protocol 1 (Table 3.4; Section 2.10) was adapted from protocols by Dr. Adrian Bracken (Cancer Transcriptomics Laboratory, Smurfit Institute of Genetics, TCD) and by Ito et al. (1997), whereas the second protocol was based on one previously described by Kaufmann et al. (2010) (Table 3.5). Initially, inflorescence material from the *35S-FIS*, AG-GFP transgenic line was used to optimize protocol 1. Although the dexamethasone response of the *35S-FIS*, AG-GFP lines was not synchronous (see Section 3.2.3.3), the flowers produced did contain carpels and stamens, indicating that the AG-GFP protein was functional (Figure 3.9C). Moreover, Western blot analysis using tissue from this line confirmed the presence of the full length

Table 3.4 Optimizations performed and parameters chosen for ChIP Protocol 1. The optimization steps performed are summarized.

Step/Reagents	Parameters that could be optimized	Parameters chosen for Protocol 1	Parameters optimized
Fixation	Tissue type	Ground tissue	None
	Concentration of formaldehyde	1% formaldehyde	None
	Length of fixation	20 minutes	10, 15, 20 minutes fixation times
Nuclei lysis	Lysis conditions	Strong lysis conditions	Harsh vs. mild lysis conditions
Sonication	Number of cycles	10 cycles	2, 4, 6, 8, 10 cycles
	Intensity of sonication	High intensity sonication	None
IP buffer	Type of buffer	ChIP dilution buffer	None
Antibody	Type of antibody	Ab290 antibody	Ab290, A6455, Roche antibodies
	Concentration of antibody	4 uL of antibody	1, 2 and 4 uL of antibody
	Incubation time of antibody with chromatin	1 hour incubation of antibody with chromatin	Overnight vs. 1 hour incubation of antibodies with chromatin
Affinity purification beads	Type of beads	Protein A Sepharose beads	Protein A Sepharose vs. Protein A magnetic beads
	Incubation time of beads with antibody and chromatin mix	2 hour incubation of beads with antibody chromatin mix	None
Wash buffers	Type of buffers	SET 1 of buffers	None
	Incubation time of buffers with affinity purification beads	5 minute incubations between MMWB washes only	5 minute incubations for all wash steps or 5 minute incubations between MMWB washes only
	Type of buffer	Elution Buffer 1	None
Elution buffer	Incubation time of beads with affinity purification beads	2 x 2 hours incubation of elution buffer with beads at 65°C	None
	Reversing cross linking between protein and DNA	65°C incubation of chromatin o/n	None
DNA purification	Enzyme treatment times to remove proteins and RNA	Incubate chromatin with RNase for 1 hour (37°C), and 2 hours with Proteinase K (55°C)	None

AG-GFP protein (Figure 3.8D). Additional optimization experiments for Protocol 1 were conducted with a *35S-FIS*, AP3-GFP transgenic line (generated by Samuel Wuest), which responded to dexamethasone treatment with synchronized flower development and showed full rescue of the petal and stamen defects found in the null mutant *ap3-3*. Optimization steps for Protocol 2 were done using dexamethasone-treated inflorescence tissues from *35S-FIS*, AP3-GFP and *35S-FIS*, PI-GFP plants (generated by Samuel Wuest).

In order to determine how successful a particular ChIP experiment was, genomic DNA regions where AG, AP3 and PI are known to bind were used as positive controls. The AG protein has been shown to bind to CArG box-like motifs found in the regulatory regions of several floral regulatory genes. These include the *AG* gene itself (where AG binding occurs within the second intron), *CRABS CLAW (CRC)* (Gomez-Mena et al., 2005) and *SHATTERPROOF2 (SHP2)* (Savidge et al., 1995). The number of known target genes for

AP3 and PI is comparatively smaller, but these proteins have been shown to bind a CArG-box DNA motif upstream of the *AP3* locus (Sundstrom et al., 2006). These regions were used as positive controls to assess the performance of the various ChIP assays and optimization steps done. In addition, several negative control regions were used (i.e. *ACTIN*, *TUBULIN*, *REF1* and genomic regions close to the positive control binding regions). It was assumed that the floral homeotic proteins did not bind these regions. The average enrichment of these negative control regions was taken as a representative background signal for the ChIP assays.

The most important parameters to judge the success of the assay are the relative enrichment of genomic DNA (the enrichment at a position where the protein of interest binds in comparison to a region where the protein of interest does not bind (i.e. the signal-to-noise ratio)) and the total amount of DNA immunoprecipitated. The latter is estimated by measuring the amount of DNA after immunoprecipitation using a fluorometric assay (see Section 2.10.5). The relative enrichment of a specific region of genomic DNA can be calculated after quantifying the amount of DNA for specific positive and negative control regions in the input fraction (represents 100% of the possible chromatin available to precipitate) and in the fraction recovered after immunoprecipitation. The amount of DNA immunoprecipitated for a specific genomic region is expressed as a percentage of the amount of DNA for the same genomic region in the input. Finally, the relative values obtained for the positive and the negative control regions are compared to yield the relative enrichment at a specific genomic locus.

3.2.6.1 Optimisation of ChIP Protocol 1

As mentioned above, protocol 1 was adapted from protocols by Dr. Adrian Bracken (Cancer Transcriptomics Laboratory, Smurfit Institute of Genetics, TCD) and by Ito et al. (1997). The different parameters tested are summarised in Table 3.4. The first parameter optimised was the fragmentation of the chromatin into fragments ranging from 250 and 750 bp. To this end, we tested different fixation times, a different number of sonication cycles, combined with mild and harsh nuclei lysis conditions

Briefly, inflorescence tissue that had been treated with a dexamethasone-containing solution four days before collection was ground in liquid nitrogen until a powder formed.

This tissue powder was incubated in a solution containing 1% formaldehyde for 10 minutes at 4°C (the samples resulting from this step are referred to as “fixed samples” hereafter). In parallel, a second tube of inflorescence tissue was ground in liquid nitrogen and the resulting tissue powder was incubated in a solution that did not contain formaldehyde for 10 minutes at 4°C (referred to as “unfixed samples” hereafter). The fixation (or crosslinking) process was stopped by the addition of glycine. After the addition of glycine, intact nuclei were isolated using a series of centrifugation steps and wash buffers, followed by lysis in mild conditions. Both the fixed and unfixed samples were split into seven equal aliquots and were sonicated at high intensity for 2, 4, 6, 8 and 10 cycles. In addition, aliquots of chromatin that had not been sonicated were kept for both fixed and unfixed samples. Next, the crosslinking was reversed by incubation of the samples at 65°C overnight, followed by the addition of RNase and Proteinase K. Finally, the DNA was purified by phenol-chloroform extraction, followed by ethanol precipitation, and the samples were analysed on a 1.5% agarose gel (Figure 3.25A).

Chromatin isolated from unfixed tissue should shear into small fragments after very few cycles of sonication, whereas chromatin that has been crosslinked should require additional cycles of sonication (Lee et al., 2006). When mild nuclei lysis conditions were applied, the fixed and unfixed chromatin behaved similarly. In particular, a DNA band corresponding to intact chromatin was observed in fixed and unfixed samples even after 10 cycles of sonication (Figure 3.25A, arrowhead). Although fragmentation of DNA for all sonicated samples was observed, as evidenced by the presence of a smear, the fragments obtained ranged from 300 bp to 1 kb after 10 cycles of sonication, which is too large for ChIP. These results suggested that the cell lysis conditions were too mild to release the chromatin from the nuclei before sonication. Therefore, harsher cell lysis conditions were tested.

For this purpose, a similar experiment was performed to the one described above with the exception that harsh cell lysis conditions were used and an extra tissue sample that had been incubated with formaldehyde for 20 minutes at 4°C was also included. With the exception of the unsonicated material, no large genomic fragments corresponding to genomic DNA were observed after 10 cycles of sonication (Figure 3.25B). The unfixed chromatin was fragmented after only 2 cycles of sonication, whereas the fixed samples required additional cycles of sonication. For example, the samples that had been treated with formaldehyde for 10 minutes fragmented after approximately 4 – 6 cycles of

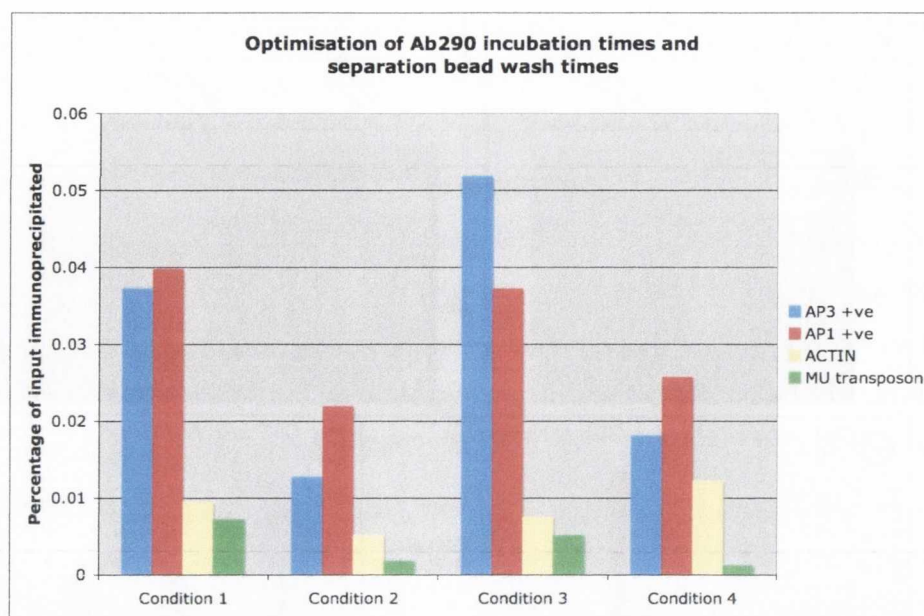


Figure 3.26 A ChIP-qPCR using tissue from 35S::API-GR pAP3::AP3-GFP *ap1 cal ap3* tissue. Incubation times of the antibody with the chromatin and incubation times of the wash buffers with the separation beads were tested here. Condition 1, Short separation bead washes with overnight antibody incubation; Condition 2, long separation bead wash with overnight antibody incubation; Condition 3, Short separation bead washes with one hour antibody incubation; Condition 4, Long separation bead washes with one hour antibody incubation. Binding regions for *API* (AP1 +ve) and *AP3* (AP3 +ve) were used as positive controls whereas the negative control binding regions used were *ACTIN* and the *mu* transposon. No error bars are presented because these values represent technical duplicates, however, C_p values were no more than 0.5 units apart (see Section 2.7). The primer pairs DM-111/ DM-112, DM-113/DM-114, DM-115/DM-116 and DM-122/DM-123 were used to amplify AP3 +ve, AP1 +ve, ACTIN, and MU transposon, respectively.

sonication, with a sharp decline in chromatin size already after 2 cycles of sonication. Samples that had been treated with formaldehyde for 20 minutes appeared to fragment more gradually depending on the number of sonication cycles. The DNA size distribution after 10 cycles was between 250 and 750 bp, as recommended for ChIP (Haring et al., 2007). Tissue samples were also treated with formaldehyde for 15 minutes, however, a similar sharp decline in fragment size was observed (data not shown). Therefore, 20 minutes of formaldehyde treatment, combined with harsh nuclei lysis conditions and 10 cycles of sonication was selected as the optimal parameters to prepare chromatin for immunoprecipitation.

Three commercial α -GFP antibodies (Ab290, A6455 and α -GFP Roche) were tested in order to identify one that would be suitable for use in a ChIP assay. Because pre-immune

serum was not available for these commercial antibodies, in order to control for non-specific antibody-epitope interactions, antibodies against β -glucuronidase (α -GUS, rabbit polyclonal) and luciferase (α -LUC, raised in mouse) were used as negative controls.

Of the three α -GFP antibodies tested, two were rabbit polyclonal antibodies (Ab290 (AbCam) and A6455 (Invitrogen)), whereas the third antibody (Roche) was a mouse monoclonal one.]Enrichment of genomic fragments stemming from the *AG* locus was observed with every antibody (Figure 3.25C, D and E). In contrast, enrichment of *CRC* genomic fragments was observed only when the Ab290 antibody was used. The best signal-to-noise ratio was observed at the *AG* and *CRC* loci when 1 μ L of the Ab290 antibody was used (Figure 3.25C). Although increasing the amount of the Ab290 antibody increased the relative amount of chromatin precipitated, it also increased the background, resulting in a lower signal-to-noise ratio. The relative amount of chromatin immunoprecipitated by the A6455 antibody was very high, however, this antibody resulted in a high background (Figure 3.25D). The signal-to-noise ratio for the Roche antibody was very low, as was the relative amount of chromatin immunoprecipitated (Figure 3.25E). Given these results, the Ab290 anti-GFP antibody appeared to offer the best signal-to-noise ratio, however, the relative amount of chromatin immunoprecipitated was low. The total yield of DNA, when using this antibody, ranged from 1 ng to 3 ng. Up to 10 rounds of ChIP may have to be performed to obtain the required 10 ng for Illumina ultra high-throughput sequencing.

The type of protein A beads to be used was also investigated using both protocols (see Table 3.4). Protein A sepharose beads and protein A magnetic beads (referred to as Dynabeads® hereafter) were compared by performing a ChIP without an antibody. This allowed the non-specific interactions between chromatin extracts and the two different types of protein A beads to be measured. It was found that the Dynabeads® produced a higher signal for all primer pairs tested than the protein A sepharose beads (Figure 3.25F). Therefore, protein A sepharose beads were chosen for use in all further ChIP experiments. Next, parameters that can affect the level of background were tested. These include incubation times between antibody and chromatin, as well as the incubation times during the washes required after binding of the chromatin/antibody complexes with the protein A beads. In this instance, tissue from AP3-GFP plants that had been treated with a dexamethasone-containing solution four days before collection was used. Chromatin was incubated with Ab290 for 1 hour and overnight (approximately 14 hours). The protein A

Table 3.5 - Optimizations performed and parameters chosen for ChIP Protocol 2. The optimization steps performed are summarized.

Step/Reagents	Parameters that could be optimized	Parameters chosen for Protocol 2	Parameters optimized
Fixation	Tissue type	Whole tissue	Whole tissue versus ground tissue
	Concentration of formaldehyde	1% formaldehyde	0.3%, 1% formaldehyde
	Length of fixation	60 minutes	15 vs. 60 minutes
Nuclei lysis	Lysis conditions	Harsh lysis conditions	None
Sonication	Number of cycles	12 cycles	8, 12 cycles
	Intensity of sonication	High intensity sonication	None
IP buffer	Type of buffer	IP buffer	ChIP dilution buffer vs. IP buffer
Antibody	Type of antibody	Ab290 antibody	None
	Concentration of antibody	4 uL of Ab290	None
	Incubation time of antibody with chromatin	1 hour incubation of antibody with chromatin	None
Affinity purification beads	Type of beads	Protein A Sepharose beads	None
	Incubation time of beads with antibody and chromatin mix	2 hour incubation of beads with antibody chromatin mix	None
	Type of buffers	IP buffer	IP buffer vs. SET 1 buffers
Wash buffers	Incubation time of buffers with Affinity purification beads	8 minutes between washes	None
	Type of buffer	Elution Buffer 2	Elution buffer 1 versus elution buffer 2
Elution buffer	Incubation time of beads with Affinity purification beads	Incubation of elution buffer with beads while shaking vigorously at 37°C	None
	Reversing cross linking between protein and DNA	Incorporated into purification	None
	Enzyme treatment times to remove proteins and RNA	Incubate chromatin with Proteinase K overnight (37°), 30 minutes RNase (37°C), Proteinase K > 6 hours (65°C)	None

beads were incubated with the wash buffers for 5 minutes after every wash step on ice or 5 minutes only for the first wash step with the Mixed Micelle Wash Buffer. It was found that the signal-to-noise ratio was highest when the separation beads were incubated for 5 minutes only for the first wash step, which was independent of the incubation time of the chromatin with the antibody (Figure 3.26). The signal-to-noise ratio appeared to be highest for samples that had been incubated with the antibody for 1 hour. Therefore, a slightly shorter separation bead wash step and antibody incubation step were used in all further ChIP experiments.

A ChIP-Seq experiment was performed on inflorescence tissue from AP3-GFP plants that had been treated with a dexamethasone-containing solution four days before collection. This ChIP-Seq experiment was successful, however, the number of binding sites identified was not as high as expected, and the background signal levels were high compared to the height of the peaks obtained. Despite these problems, this first ChIP-Seq

experiment allowed the identification of some novel binding regions of the AP3 protein, which were used in subsequent experiments to optimise protocol 2.

3.2.6.2 Optimisation of ChIP protocol 2

Given the low signal-to-noise ratios obtained from the ChIP-Seq experiments using protocol 1 and inflorescence tissue from AP3-GFP plants, the protocol detailed in Kaufmann et al. (2010) was tested. This protocol was published after protocol 1 had been optimised, and hence could not be tested before. The parameters optimised in protocol 2 are summarised in Table 3.5. One of the main differences between protocol 1 and protocol 2 is that whole tissue is incubated with formaldehyde under vacuum in protocol 2, whereas the tissue was ground in liquid nitrogen before formaldehyde fixation in protocol 1.

More specifically, in protocol 2, whole tissue was incubated with a solution containing 1% formaldehyde under vacuum on ice. This alone had the effect of increasing the signal-to-noise ratio by over seven fold. The effect of 2 different numbers of sonication cycles was tested (8 vs. 12 cycles). Although the sonication cycle number did not alter the signal-to-noise ratios, a more desirable DNA fragment range was obtained when 12 cycles were performed.

The other parameters that were tested were the buffers used for immunoprecipitation, the protein A beads washes and the conditions to elute the DNA/antibody complex from the protein A beads. Each of these parameters was changed step-wise from protocol 1 and had the effect of either increasing the signal-to-noise ratios or the total yield of DNA. In sum, protocol 2 produced better signal-to-noise ratios and yielded higher amounts of DNA after immunoprecipitation. Therefore, protocol 2 was used for subsequent ChIP-Seq experiments.

3.2.7 Global identification of genomic regions bound by AGAMOUS

3.2.7.1 Generation of an Illumina ChIP-Seq library from ChIP enriched material

In order to identify genomic regions bound by AG on a global scale, two independent large-scale ChIP experiments were performed (referred to as experiments 1 and 2 hereafter) using *pAPI-FIS*, AG-GFP transgenic plants and the α -GFP antibody Ab290 from Abcam (see Section 2.10.3). To this end, inflorescence-like meristems were harvested 4 days after treatment with a dexamethasone-containing solution and then processed using the ChIP protocol 2 outlined above (see Section 2.10). For experiments 1 and 2, 1.5 mL and 1 mL, respectively, of tissue were used as starting material. After the ChIP reactions, the precipitated genomic DNA was purified and quantified using a fluorescence-based assay. Experiment 1 resulted in a total DNA yield of 11.5 ng, while experiment 2 yielded 4.9 ng of genomic DNA. To test the specificity of the ChIP assays, aliquots of the DNA preparations were subjected to qPCR analysis using primers specific to positive (*CRC*, *AG* and *SHP2*) and negative (*ACTIN*, *TUBULIN* and *REF1*) control regions. All three positive controls were significantly enriched above background (Figure 3.27A). These results indicated that the ChIP assays were successful. Using the precipitated chromatin from the two experiments, two Illumina ChIP-Seq libraries were prepared (see Section 2.10.6).

In brief, the ends of the ChIP-enriched DNA fragments were first blunted to allow the addition of adenine overhangs to their ends. Following this, the Illumina-specific oligonucleotide adapter sequences were ligated to the modified ChIP DNA. Next, the adapter-DNA mix was size-selected on a 2% agarose gel to obtain a range of 300 – 500 bp fragments. This DNA was then purified and amplified using a high-fidelity DNA polymerase and adapter-specific oligonucleotides. The amplified material was subsequently size-selected for a second time in order to remove primer-dimers and adapter-dimers (see Section 2.10.6). Finally, the quality of the library preparations was examined by gel-electrophoresis and qPCR. The size-range of DNA amplified and extracted was in the range of 300 – 500 bp, as expected, and little or no primer- or adapter-dimers were observed (Figure 3.27C, D). An aliquot of these libraries was taken and qPCR analysis was

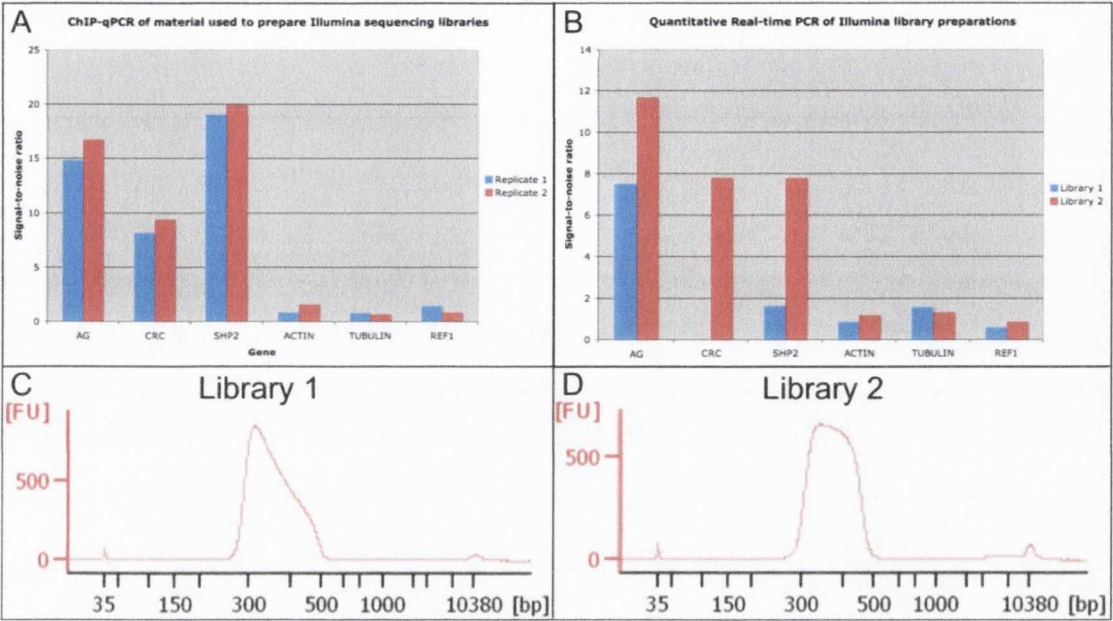


Figure 3.27 – Preparation of ChIP-Seq libraries for the genome-wide identification of AG binding sites. (A) Quantitative real-time PCR for selected genes (as indicated) using immunoprecipitated genomic DNA from dexamethasone-treated inflorescences of *pAPI-FIS*, AG-GFP plants. Two independent biological replicates were performed (Replicate 1 and 2). The signal-to-noise ratios were calculated as outlined in Section 2.9. (B) Quantitative real-time PCR of Illumina ChIP-Seq libraries prepared from the material used in (A). No PCR was performed using Library 1 material and *CRC*-specific primers. The signal-to-noise ratios were calculated as outlined in Section 2.9. Input DNA from both ChIP preparations were used to calculate the relative enrichments observed in the library preparations. (C – D) DNA size range of Illumina ChIP-Seq libraries from (B) used for ultrahigh-throughput sequencing. An Agilent 2100 Bioanalyser was used to analyse the size distribution of the DNA fragments. FU, florescence units; bp, base pairs. The primer pairs DM-103/DM-104, DM-107/DM-108, DM-117/DM-118, DM-115/DM-116, DM-119/DM-120 and DM-124/DM-125 to amplify *AG*, *CRC*, *SHP2*, *ACTIN* (AT5G09810), *TUBULIN* (AT5G62700) and *REF1*

performed. Input DNA from each ChIP was used to normalize the data, as described above. Only the *AG* and *SHP2* positive control regions were tested for enrichment in Library 1 (resulting from experiment 1 described above). The signal-to-noise ratios were lower when compared to the initial ChIP-enriched DNA. Similar results were obtained when Library 2 was tested, however, enrichment and the *CRC* positive control region was higher when compared to the initial ChIP-enriched DNA (Figure 3.27B). Although it remained unclear why the signal to noise ratios had been compressed during library preparation, these 2 libraries were sent for ultra high-throughput sequencing, which was carried out by Dr.

Amanda Lohan (B. Loftus Group, Conway Institute, University College Dublin) using an Illumina Genome Analyzer GAII.

In parallel to the 2 ChIP-Seq experiments carried out using inflorescence tissue from *pAPI-FIS*, AG-GFP plants, a ChIP using inflorescence tissue harvested from *35S-FIS* plants 4 days after treatment with a dexamethasone-containing solution was performed in the same conditions. While the ChIP was performed by Samuel Wuest, I prepared the Illumina ChIP-Seq library. The DNA fragments of the same size range as those used to prepare Libraries 1 and 2 were sent for sequencing (data not shown). The latter ChIP-Seq experiment was essential to determine the background and to correct for artefacts related to the antibody used. The sequencing results for these three libraries are presented below.

3.2.7.2 Quality assessment of the ChIP-Seq data

The libraries generated by ChIP from *pAPI-FIS* AG-GFP plants (denoted AG-1 and AG-2 hereafter) and the control library derived from *35S-FIS* plants (denoted NEG hereafter) were subjected to Illumina next-generation sequencing. The resulting data were analysed in parallel with three other ChIP-Seq libraries that were generated from the *35S-FIS* PI-GFP lines (2 independent biological replicates, denoted PI-1 and PI-2) and the *35S-FIS* AP3-GFP lines (denoted AP3-1 hereafter). In brief, the data produced by the Illumina sequencing pipeline included the sequence string output and a quality score. The quality score reflects the probability that a given base call at a given position is correct. These quality scores and sequence strings were uploaded to the online quality control tool, *FastQC* (www.bioinformatics.bbsrc.ac.uk). The *FastQC* software generated a box plot that illustrated the distribution of per-base qualities of all reads (Figure 3.28A, the results for the PI-1 library are shown as an example). The background colours indicate very good quality calls (green), calls of reasonable quality (orange), and calls of poor quality (red). The box-plots produced for all sequenced samples indicated that the sequencing data were generally of high quality. A moderate decrease in quality was observed towards the 3' end of reads. This trend was similar for the data from all six sequenced libraries.

The *FastQC* software analyses various other parameters of the sequencing run including the number of duplicated sequences. A high level of duplication indicates that there is some enrichment bias present. The expected duplication level for a sequencing run of the

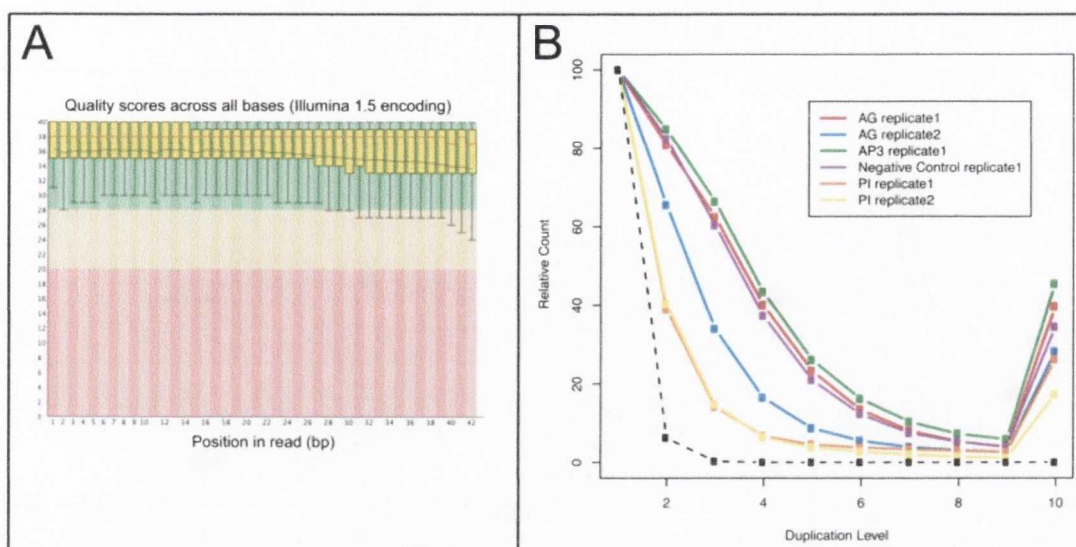


Figure 3.28 Results of ChIP-Seq quality controls. (A) Quality scores at a given position within the sequence of all reads from the PI-1 ChIP-seq experiment, as determined by the *FastQC* software. The box-plots show the distribution of per-base qualities for all reads in the PI-1 library. The average quality overall is high, and drops only moderately towards the 3' end of the 42 bp reads, indicating good overall quality of the dataset. (B) Duplication levels of the reads present in the ChIP-seq data from six libraries sequenced on the same Illumina GAIIIX analyzer flow cell. The number of single reads is represented as 100% and duplicated sequences are represented as a fraction of this. Duplication level 10 is equivalent to the percentage of sequences that were present 10 or more times. The PI-GFP and AP3-GFP libraries were included for comparison. All sequence data contain a substantial amount of multiple reads, with the PI libraries containing the lowest amount and the AG-1 and AP3-1 libraries containing a similar number of multiple reads as the control library. The AG-2 library contained an intermediate level of duplication. The black line indicates theoretically expected duplication levels, according to Poisson distribution densities with $\lambda = 30$ million (*i.e.* approximate reads per library) divided by 240 million (the approximate genome size, including information on strand orientation).

Arabidopsis genome was calculated (Figure 3.28B, black line) and compared to the number of observed duplicated sequences. A high level of sequence duplication was observed in all libraries sequenced (Figure 3.28). The AG-1 library contained an equal amount of sequence duplication as the AP3-1 and NEG libraries whereas the two PI libraries contained the lowest amount of sequence duplication. The AG-2 library contained intermediate levels of sequence duplication. Only single read sequences were retained to avoid false discoveries in peak-calling methods (Muino et al., 2011a). Given the large data volume resulting from ChIP-Seq experiments, these sequences can be readily removed without affecting the overall sequencing coverage.

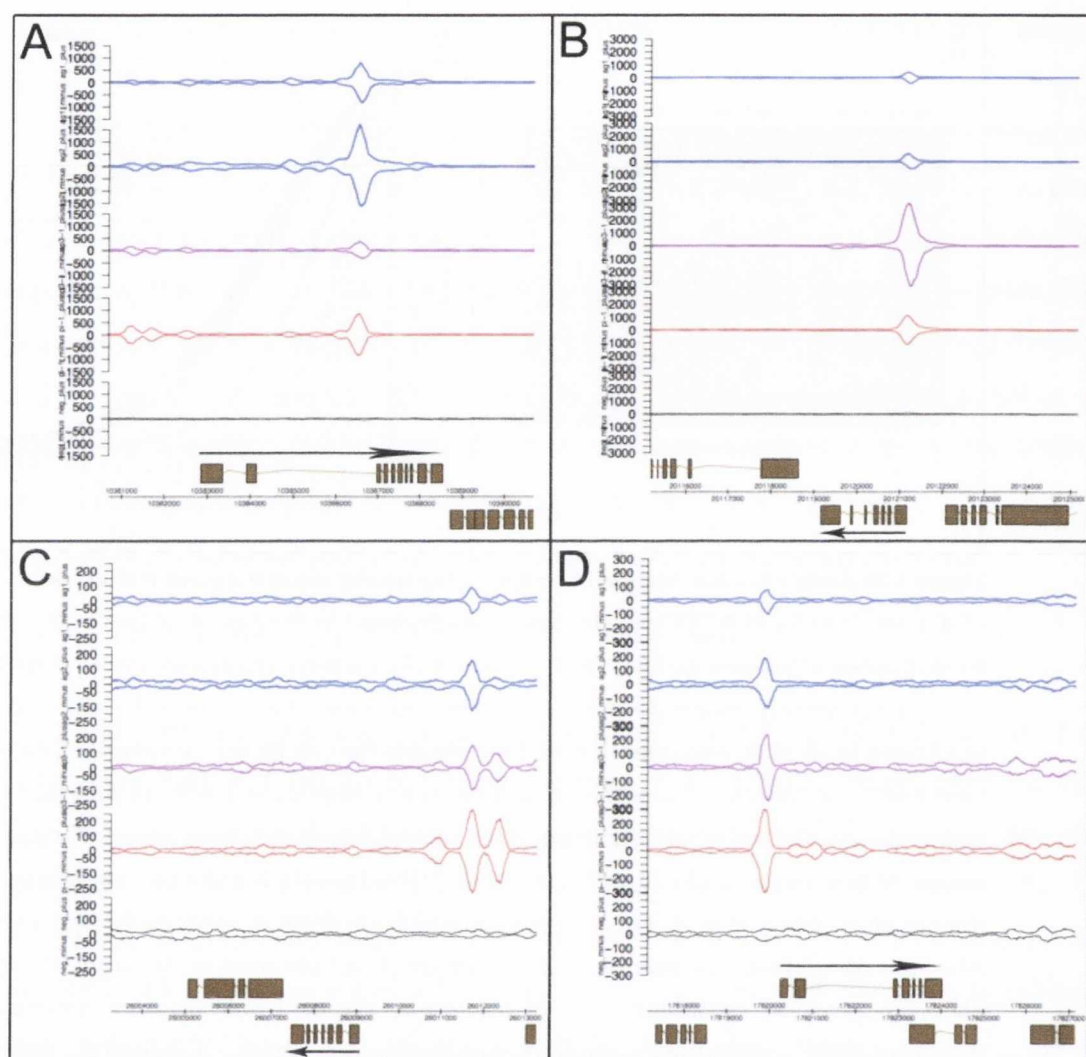


Figure 3.29 ChIP-Seq data for known AG targets. (A – D) Per-base pair coverage in the genomic regions surrounding the *AG* locus (A); the *AP3* locus (B); the *CRC* locus (C) and the *SHP2* locus (D). Per-base pair coverage was calculated by extending the 42 bp reads generated by Illumina sequencing to 200 bp in the 3' direction. Gene models are shown at the bottom of each graph (forward strand above the genome axis and reverse strand below). The same scale is used for every individual graph. The results for the AG-1 (top) and AG-2 libraries are shown in blue; the results for the AP3-1 library are shown in purple, and those for the PI-1 library in red. The bottom track (in black) represents the results for the control library (NEG). Images were generated using the *R* package *GenomeGraphs*, and annotation tracks were generated from information downloaded from the *BioMart* database using *biomaRt* functionality.

Next, the short oligonucleotide alignment program (SOAP2 (Li et al., 2009a)) was used to align the unique short read sequences obtained to the *Arabidopsis* genome (TAIR release 10, www.arabidopsis.org). For the AG-1 and AG-2 libraries, ~27.5 million and

~32.5 million unique reads were obtained and ~80% and ~89% of these reads aligned to the *Arabidopsis* genome. Similar numbers of unique reads and read mappability was observed for the other samples (Table 3.6). Together, the data presented above indicated that the ChIP-Seq samples generated were of suitable quality and sufficient coverage.

3.2.7.3 Detection of AG binding sites in the *Arabidopsis* genome

In order to identify genomic fragments bound by the AG transcription factor, the enrichments of reads that aligned to known AG binding sites were investigated first. Read lengths were extended to 200 bp in the 3' direction in order to reflect the size of the genomic fragments that were initially immunoprecipitated. Then the genome-wide per-base coverage from these extended reads was calculated. Previously known binding sites of AG (Gomez-Mena et al., 2005; Savidge, 1996) were analysed including *CRC*, *SHP2*, *AP3* and *AG* itself. The NEG, PI-1 and AP3-1 datasets were also included in this analysis for comparison (Figure 3.29). Strong per-base coverage enrichments were observed in the second intron of the *AG* locus in both the AG-1 and the AG-2 datasets although the enrichment observed for the AG-2 library was higher (Figure 3.29A). The degree of enrichment at this locus was very high, which is probably reflective of the extra copies of this genomic sequence in the *pAPI-FIS* AG-GFP line. Indeed, similar effects were observed at the *AP3* and *PI* loci in the AP3-1 (Figure 3.29B) and PI-1 datasets, respectively (Wuest, 2011). Enrichment was also observed at the *AP3* locus but the signal appeared comparatively low relative to the AP3 track due to scaling of the graph (both tracks are scaled to 3000 coverage units due to the high enrichment of AP3-GFP binding at the *AP3* locus since that genomic region is represented four times in the *35S-FIS*, AP3-GFP lines).

Table 3.6 Number of unique reads obtained from the differeny ChIP-Seq experiments (as indicated) and their alignment to the *Arabidopsis* genome.

Sample	No. reads	No. aligned	% aligned
AP3-1	31'863'205	29'242'357	91.77%
PI-1	31'660'640	29'124'266	91.99%
PI-2	27'679'860	25'591'938	92.46%
NEG	29'265'976	26'731'313	91.34%
AG-1	27'492'693	22'050'870	80.21%
AG-2	32'554'376	28'908'778	88.80%

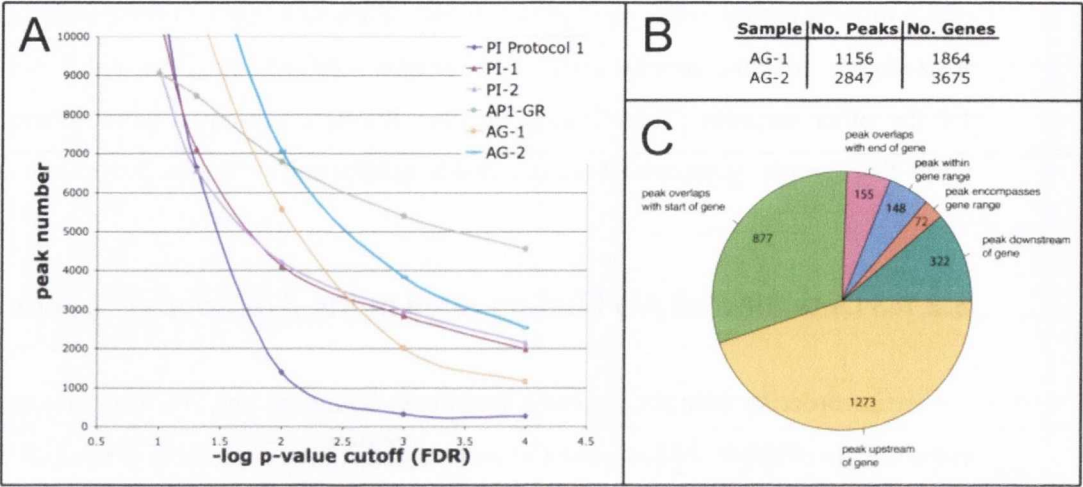


Figure 3.30 Detection of significantly enriched regions (peaks) in the ChIP-Seq datasets using a combination of two peak-calling methods. (A) The number of significant peaks at different p -value cut-offs as determined by the *CSAR* method. The analysis included the two AG datasets (AG-1, AG-2), the two PI datasets (PI-1, PI-2), the PI dataset generated using sub-optimal conditions (see Section 3.2.6.1), and the AP1 ChIP-Seq dataset by (Kaufmann et al., 2010b). (B) A table summarizing the number of peaks detected by the *CSAR* method in AG-1 and AG-2 ($FDR < 0.001$ and peak regions > 50 bp in width) and confirmed by the independent peak-calling method (BayesPeak, see Section 2.12.5). The number of genes with associated peaks (within ranges of 3 kb upstream of a translational initiation site to 1 kb downstream of a translational termination site) are also listed. (C) The distribution of peaks 3 kb upstream of a translational initiation site to 1 kb downstream of a translational termination site in AG-2. Peaks were assigned based on the region closest to a given peak.

Enrichments were also observed at the *CRC* and *SHP2* loci, however, as was observed at the *AG* locus, enrichments were higher for AG-2 than for the AG-1 dataset (Figure 3.29C – D). No enrichments were observed for the NEG dataset at any of the loci investigated.

Next, two separate peak-calling algorithms were used to identify AG binding site on a genome-wide scale. The first method (the Bioconductor package *CSAR* (Muino et al., 2011a)) removed duplicated reads from the datasets and calculated the per-base coverage in the AG-1, AG-2 samples and the negative control (NEG). After normalization of the data, a Poisson distribution-based statistic was calculated and used to determine the per-base enrichment over the negative control. The ChIP-Seq data for AP1 (Kaufmann et al., 2010) were also included in this analysis for comparison. The number of significantly enriched peaks were analysed at different significance cut-off values. Given the exponential increase in the number of binding sites observed at higher cut-off values for the AG datasets (Figure 3.30A), a False Discovery Rate (FDR) cut-off value of < 0.001 was chosen. To consider a

genomic region as enriched in a dataset, at least 50 consecutive base pair calls were required with an FDR < 0.001 per base. A second peak-calling method based on a Hidden Markov Model combined with a Bayesian statistical method, *BayesPeak* (see Section 2.12.5) (Spyrou et al., 2009), was used to confirm the findings of the CSAR analytical method. Only peaks first called by the CSAR method and then by the BayesPeak method were considered significant.

Using the two methods outlined above, 1,156 and 2,847 binding peaks were defined as being enriched above background in the AG-1 and AG-2 datasets, respectively (Figure 3.30B). Peaks that were within 3 kb upstream of a translational initiation site or within 1 kb downstream of a translational termination site of a given gene were considered to be associated with that gene. With these criteria, the number of gene-associated peaks in the AG-1 and AG-2 libraries was found to be 1,864 and 3,675, respectively (Figure 3.30B). In the AG-2 dataset, ~45% of peaks were within 3 kb upstream of a translational initiation site, ~31% of peaks overlapped with the translational initiation site and ~11% of peaks were within 1 kb downstream of the translational termination site (Figure 3.30C). These data are in agreement with previous studies (e.g. Kaufmann *et al.*, 2010) that showed that *Arabidopsis* transcription factors often bind to sequences in close proximity to the transcription initiation sites of genes.

Next, the overlap between the genes identified in the AG-1 and AG-2 libraries as putative AG targets was calculated to judge the reproducibility of the experiment. An overlap of 941 genes was observed between the AG-1 and AG-2 datasets whereas 856 and 2,617 genes were only identified in the AG-1 and AG-2 libraries, respectively (Figure 3.30C). This is consistent with the lower coverage enrichments observed at control loci for AG-1 (Figure 3.30A, C and D). In addition, the overlap of putative AG target genes with genes identified as differentially expressed in the 5-day time-point of the *ag* microarray experiments was determined. This analysis revealed an overlap of 93 and 142 genes that are present in the datasets from the AG-1 and AG-2 libraries, respectively, and that responded to a perturbation of AG activity (Figure 3.30C). Next, a Gene Ontology (GO) term analysis was performed using genes that were associated with binding peaks. This analysis was performed in the same way as in the microarray data analyses in Section 3.2.4 and 3.2.5 (see Section 2.12 for Methods). Using the Benjamini-Hochberg method for multi-hypothesis testing, an adjusted *p*-value cut-off of < 0.05 was employed to identify over-represented GO terms. This analysis yielded 74 GO terms in total (Figure 3.31A). These

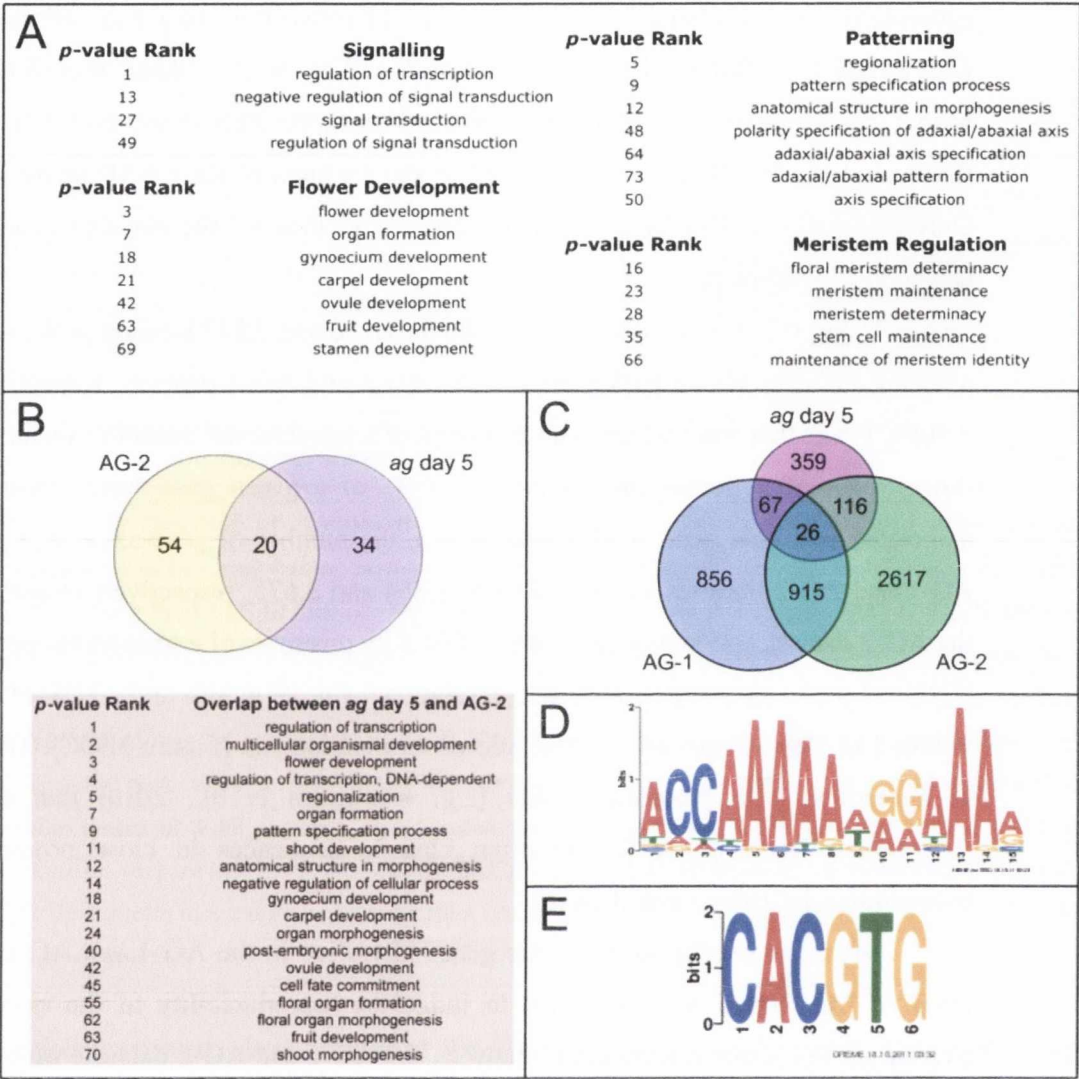


Figure 3.31 Analysis of the AG ChIP-Seq data. (A) An adjusted (Benjamini-Hochberg method for multi-hypothesis testing) *p*-value cut-off of < 0.05 was used to identify over-represented Gene Ontology (GO) terms among the genes bound by AG. The column “*p*-value Rank” separates the most significant *p*-value (Rank 1) from the least significant (Rank 54). The GO terms were categorised into four groups that broadly represent the biological processes involved. (B) A Venn diagram of the overlap between the significantly enriched GO terms in AG-2 and day 5 of the *ag* mutant time-course experiment (section 3.2.4). The list of GO terms enriched in both datasets is presented below the Venn diagram. (C) A Venn diagram of the overlap in genes associated with a significant binding peak in AG-1 and AG-2 and the overlap with genes significantly (unadjusted $p < 0.01$) differentially expressed in the *ag* mutant time-course at day 5. (D and E) The DNA motif discovery algorithm DREME (Bailey, 2011), was used to calculate the most abundant DNA motifs in the regions bound by AG-GFP. The software identified the CArG-box motif (D) and the G-box motif (E).

terms were further processed in that related GO terms were placed into one of the following parent categories: *Signalling*; *Flower Development*; *Patterning*; and *Meristem Regulation*. The GO terms *Regulation of Transcription* and *Signal Transduction* were among the highest ranked terms in the *Signalling* category. *Carpel Development*, *Ovule Development* and *Stamen Development* were among the top GO terms in the *Flower Development* category. *Regionalisation* and *Pattern Specification Processes* were significantly enriched in the *Patterning* category. Finally, *Meristem Maintenance* and *Floral Meristem Determinacy* were included in the *Meristem Regulation* category. The enrichment of all of these terms in the ChIP-Seq datasets is consistent with the known role of *AG* as a master regulator of floral organ identity and floral meristem determinacy.

Although these enriched GO terms indicated that *AG* binds to genes involved in the corresponding processes, no inferences of gene expression changes of the bound genes can be made based on these data. Therefore, the GO terms that were enriched in the ChIP-Seq dataset were compared with the GO terms enriched in the 5-day time-point of the *ag* mutant experiments (see Section 3.2.4). Twenty GO terms were found to be enriched in both the ChIP-Seq dataset and in the data from the *ag* mutant experiment (Figure 3.31B). These GO terms included *Regulation of Transcription*, *Regionalization* and *Gynoecium Development*. Again, these terms are consistent with the known role of *AG* during flower development.

It was observed that many genes involved in organ boundary formation were associated with *AG* binding peaks. Therefore, the binding enrichment coverage and the fold-change differences were visualized using the ChIP-Seq data as well as the gene expression data from the *ag* mutant experiments (Figure 3.32). Significant binding peaks were present in the *AG*-2 library for *API*, *AP2*, *SUP*, *CUC1*, *CUC2* and *RBE* (Figure 3.32A – D). In the *AG*-1 library, *AG* binding was also observed for *API*, *AP2*, *RBE* and *CUC2*. Although increased coverage was observed for the *CUC1* and *SUP* loci, no significant binding was observed. This reflects the relatively poor quality of the *AG*-1 dataset, which often contains binding coverage enrichments for known genes but these peaks do not meet the significance cut-offs of the peak-finding algorithms used. The expression levels of *API* and *RBE* were increased in the *ag* mutant experiments at the 2.5 and 5 day time-points, which is consistent with the antagonistic roles they play with *AG* in development. However, *API* was not considered significantly differentially expressed in the analysis and *RBE* was only considered significantly up-regulated at day 5 (unadjusted p -value < 0.01). *SUP*, *CUC1*, *CUC2* and *AP2* were all down-

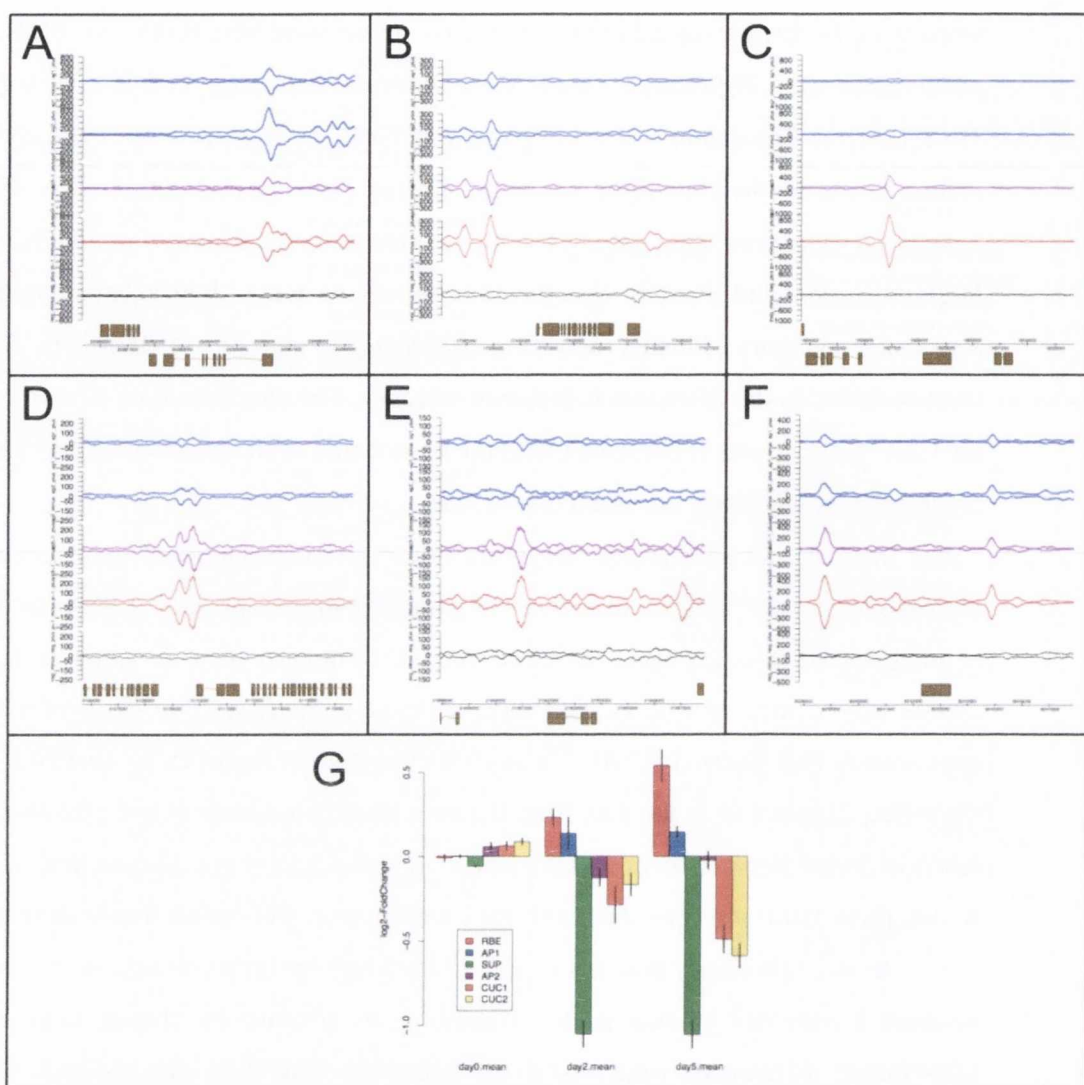


Figure 3.32 ChIP-Seq results for selected AG targets. (A – F) Per-base pair coverage in the genomic regions surrounding the *AP1* locus (A); the *AP2* locus (B); the *SUP* locus (C) the *CUC1* locus (D); the *CUC2* locus (E) and the *RBE* locus (F). The per-base pair coverage was calculated by extending the 42 bp reads generated by Illumina sequencing to 200 bp in the 3' direction. Gene models are shown at the bottom of each graph (forward strand above the genome axis and reverse strand below). The same scale is used for every individual graph. The results for the AG-1 (top) and AG-2 libraries are shown in blue; the results for the AP3-1 library are shown in purple, and those for the PI-1 library in red. The bottom track (in black) represents the results for the control library (NEG). Images were generated using the *R* package GenomeGraphs, and annotation tracks were generated from information downloaded from the *BioMart* database using *biomaRt* functionality. (G) Expression of genes known or suspected to be AG-dependent in *ag* mutant flowers relative to wild-type flowers. Flowers were collected at different time-points (as indicated) after induction of flowering by dexamethasone treatment and RNA isolated from these flowers was subjected to microarray analysis. Data shown are log₂-transformed fold change expression ratios (*ag* vs. wild type). Error bars indicate s.e.m.

regulated at the 2.5 day time-point although only *SUP* was considered significantly repressed. *SUP*, *CUC1* and *CUC2* were significantly down-regulated at day 5 whereas *AP2* did not show any differences in expression. The gene expression trends observed here suggests that *AG* maintains its own boundaries of expression during flower development. The non-significant expression of *API* can be explained by the use of the *35S::API-GR* gene for these experiments, which expresses *API* ectopically regardless of *AG* expression. The non-significant expression of *AP2* can be explained by the relatively mild expansion of the *AP2* expression domain in *ag-1* flowers (Wollmann et al., 2010). However, the trend towards down-regulation of *AP2* expression at day 2.5 was unexpected.

Finally, the DNA motif discovery algorithm *DREME* (Bailey, 2011) was used to identify DNA sequence motifs that are enriched in the regions bound by AG-GFP (Figure 3.31E and F). One of the most prevalent motifs found was the CArG-box DNA motif (5'-TDCC(W)₆GGH-3', where D = A, G or T; W = A or T; and H = A, C or T), which was identified by Shiraishi et al. (1993) as an *in vitro* binding site for AG (Shiraishi et al., 1993). Another motif that was identified as enriched in the genomic fragments bound by AG is the G-box motif, which serves as a binding site for basic Helix Loop Helix (bHLH) and basic leucine zipper (bZIP) transcription factors (Siberil et al., 2001). Several members of these transcription factor families are known to play important roles in flower development. For example, the *HECATE* genes encode bHLH transcription factors that are involved in formation of the gynoecium (Gremski et al., 2007), whereas the bZIP transcription factors bZIP21 and bZIP65 have been shown to be involved in anther development (Murmu et al., 2010). This raises the possibility that AG binds to DNA sequences in protein complexes with non-MADS-box proteins. A similar idea has been put forward by Kaufmann et al. (2009) for the MADS-domain protein SEP3.

In summary, published binding sites of AG were identified in the AG-2 library but not in the AG-1 library. Several processes directly under the control of AG have been highlighted by the ChIP-Seq data including its role in regulating floral organ boundary genes. These data can be further mined and combined with other datasets in order to identify processes and genes regulated by *AG*.

3.2.8 Discussion

I have performed studies that aimed at identifying genes and processes regulated by the floral homeotic transcription factor AG on a genome-wide scale. Using microarray analysis, genes that respond to a perturbation of AG activity at different stages of early flower development have been identified. Two different strategies were applied to disrupt AG function: 1) an *ag-1* mutant was used and 2) AG activity was perturbed at different stages of flower development using an inducible amiRNA delivery system. Additionally, I have identified the genome-wide binding sites of the AG transcription factor using chromatin immunoprecipitation combined with ultra high-throughput sequencing.

There are several avenues of discussion to explore and therefore, for the sake of simplicity, these discussions will be divided into two main sections. The first section will discuss technical aspects of transgenic line generation and experimental design whereas the second section will address the developmental insights that can be garnered from the data generated.

3.2.8.1 Transgenic line generation, experimental setup and implementation

In order to identify the most efficient method of perturbing AG activity, three different strategies were tested: amiRNAs and RNAi constructs complementary to the AG mRNA were designed and translational fusions of AG with transcriptional repressor domains were generated. In the following sections, I will discuss the results of the approaches used to disrupt AG function.

Identification of an amiRNA that efficiently targets AG

In order to isolate a functional amiRNA targeting the AG mRNA, 7 different amiRNAs were designed and constitutively over-expressed in wild-type plants. Only one of the amiRNAs (the AG4-amiRNA) led to a strong *ag* mutant phenotype (Figure 3.1D), whereas all other amiRNAs tested showed no or only weak floral defects, suggesting that

they did not efficiently perturb AG activity (Figure 3.1A-F). Although the *AG4-amiRNA* was initially predicted to specifically target the *AG* mRNA, an updated version of the amiRNA prediction software indicated that it might also target the mRNA of another gene. The predicted off-target of the *AG4-amiRNA* is a glucose-1-phosphate adenylyltransferase that is not highly expressed in flowers (according to the *Arabidopsis* eFP Browser; bar.utoronto.ca/welcome.htm). In order to determine if the mRNA levels of this gene were affected upon expression of the *AG4-amiRNA*, its levels were monitored in ethanol-treated flowers of 35S::AlcR/pAlcA::*AG4-amiRNA* plants. Its expression was found to be reduced to about 80% of wild-type levels. This weak response, as well as the nature of the gene affected, makes it unlikely that this off-target effect had any major influence on the experimental data obtained.

Another *AG*-specific amiRNA (the *AG5-amiRNA*) that I tested and for which I did not observe any phenotypic effects (Figure 3.1F) was recently over-expressed in wild-type plants by Wollmann et al. (2010), resulting in strong *ag*-like phenotypes. Because of the number of independent transformants isolated for this amiRNA in this study, it is unlikely that positional effects of the T-DNA insertion site have prevented the expression of the *AG5-amiRNA* to sufficiently high levels. It is currently unclear why there is a discrepancy between the effects observed for our *AG5-amiRNA* lines, compared to those of Wollmann et al., (2010). A possible explanation is that Wollmann et al. may have used tandem CaMV 35S promoters to drive expression of their amiRNA, thus resulting in higher expression levels.

Functional amiRNAs that efficiently target the mRNAs of different developmental regulators have been identified in the Wellmer laboratory (Table 3.7). These results indicate that amiRNAs are suitable for disrupting the functions of key floral regulators, but they also demonstrate that identification of an amiRNA that can efficiently knockdown the activity of its target gene can be difficult and time-consuming. So far, 29 different amiRNAs targeting 11 genes, with previously described mutant phenotypes, have been over-expressed in wild-type plants in our laboratory (see Table 3.7). Of these 29 amiRNAs, 10 have been shown to result in phenotypes similar to those of the corresponding mutants, while all others displayed no or only weak mutant phenotypes. Although relatively few genes have been investigated so far, the inability of many amiRNAs to efficiently silence target genes has been previously reported (Ossowski et al., 2008). Several explanations have been proposed to account for unsuccessful gene silencing, including the *in planta* production of amiRNAs

with unfavourable properties, inaccessibility of the complementary binding site in the target mRNA and difficulties in reducing steady-state target mRNA levels because of negative feedback regulation (Ossowski et al., 2008).

Very little is known of how plant miRNA precursors are processed, therefore it is possible that mature amiRNAs are produced without the intended sequence (Schwab et al., 2006). In the case of *AG* and *PI*, this is unlikely to be the situation as 7 and 10 different amiRNAs have been designed to target each gene product, respectively.

The inaccessibility of the complementary binding site may play a role in the efficiency of amiRNA silencing, as has been shown in mammals (Ossowski et al., 2008). The complementary region where the AG4-amiRNA binds to the *AG* mRNA overlaps with that of the RNAi construct designed to target the *AG* mRNA. Although the penetrance of the RNAi transgene was low, the phenotypes observed were much stronger than those of the other six AG-amiRNAs tested. This correlation suggests that the accessibility of amiRNA binding sites is an important determinant for the functionality of the different AG amiRNAs. Furthermore, nine different amiRNAs and an RNAi construct targeting *PI* have been designed and expressed in plants without efficient silencing being observed. None of these amiRNAs or the RNAi construct targeted the same region as the functional *PI-amiRNA* (not shown).

The increased silencing of a target gene by amiRNAs may lead to a higher rate of transcription of the target gene through a feed-back response (Ossowski et al., 2008). This possibility could be tested for certain genes by measuring the unspliced forms of transcripts in wild-type plants and those transformed with an amiRNA. In situations where this negative feedback exists, a dominant negative version of the transcription factor of interest would be a more suitable approach.

Another possible explanation for unsuccessful gene silencing is the position at which the amiRNAs target their complementary mRNAs. As described above, members of the Wellmer laboratory have expressed 29 different amiRNAs in plants targeting 11 genes (Table 3.7). All amiRNAs that successfully recapitulated their corresponding mutant phenotypes were complementary to coding sequences within their target mRNAs, as opposed to the untranslated regions. These results suggest that amiRNAs that target coding regions have an increased efficiency. This idea is supported by the observation that most endogenous *Arabidopsis* miRNAs target coding sequences as well (Chen and Rajewsky, 2007). However, the target region does not appear to be the only determinant for efficiency

Table 3.7 Summary of amiRNAs tested in the Wellmer laboratory by over-expression in *Arabidopsis* using the CaMV 35S promoter. The different amiRNAs designed for a given target genes are numbered (in parentheses; e.g. AGAMOUS (4) corresponds to AG4-amiRNA in the text). All of the amiRNAs that resulted in phenotypes resembling the corresponding mutants had complementary regions within the coding sequence of their target transcript.

Target Gene	Complementary binding site	Phenocopy
<i>KNOTTED-LIKE1 FROM ARABIDOPSIS THALIANA</i> 1 (1)	Exon	Yes
<i>APETALA 3</i> (1)	Exon	Yes
<i>APETALA 3</i> (2)	Exon	Yes
<i>AGAMOUS</i> (4)	Exon	Yes
<i>KNUCKLES</i> (1)	Exon	Yes
<i>JAGGED</i> (1)	3' UTR/Exon	Yes
<i>JAGGED/NUBBIN</i> (1)	JAG: Exon	Yes
	NUB: Exon	Yes
<i>APETALA 1</i> (1)	Exon	Yes
<i>PISTILLATA</i> (10)	Exon	Yes
<i>UNUSUAL FLORAL ORGANS</i> (1)	Exon	Yes
<i>AGAMOUS</i> (1)	3' UTR	No
<i>AGAMOUS</i> (2)	3' UTR	No
<i>AGAMOUS</i> (3)	3' UTR	No
<i>AGAMOUS</i> (5)	Exon	No
<i>AGAMOUS</i> (6)	Exon	No
<i>AGAMOUS</i> (7)	Exon	No
<i>PISTILLATA</i> (1)	3' UTR	No
<i>PISTILLATA</i> (2)	3' UTR	No
<i>PISTILLATA</i> (3)	3' UTR	No
<i>PISTILLATA</i> (4)	3' UTR	No
<i>PISTILLATA</i> (5)	Exon	No
<i>PISTILLATA</i> (6)	3' UTR	No
<i>PISTILLATA</i> (7)	Exon	No
<i>PISTILLATA</i> (8)	Exon	No
<i>PISTILLATA</i> (9)	Exon	No
<i>APETALA 2</i> (1)	Exon	No
<i>RABBIT EARS</i> (1)	Exon	No
<i>RABBIT EARS</i> (2)	Exon	No
<i>RABBIT EAR</i> (3)	Exon	No

of amiRNA function, as 11 other amiRNAs that were over-expressed in *Arabidopsis* target coding regions but did not result in phenotypic abnormalities. Therefore, it is likely that all of the scenarios discussed above contribute to ineffective silencing to varying degrees.

Recently, a new strategy for amiRNA design has been proposed that may improve the success-rate of amiRNA-mediated gene silencing. Park et al. (2009) showed that an amiRNA that is perfectly complementary to the *API* mRNA could silence *API* more efficiently than an amiRNA that was designed based on the principles proposed by Schwab et al., which favour amiRNAs with mismatches in certain positions of the amiRNA (Park et al., 2009). The mismatched and perfectly complementary amiRNAs tested by Park *et al.*, (2009) targeted identical regions of the *API* mRNA. Therefore, base-pair mismatches between an amiRNA and its target may reduce amiRNA-induced gene silencing. However, the risk of transitivity increases when perfectly complementary amiRNAs are used as they can prime the production of dsRNA that can be processed into siRNAs (Voinnet, 2008).

This may lead to the silencing of genes with similar nucleotide sequences similar to the target transcript.

To summarise, amiRNAs are an extremely valuable component of the molecular toolbox available to perturb the activity of genes of interest. The functional amiRNAs generated in the Wellmer laboratory recapitulate their corresponding mutant phenotypes with a high degree of accuracy. However, there was a disparity between the number of predicted functional amiRNAs and those that successfully phenocopied their mutant phenotypes when over-expressed in plants. Therefore, the process of identifying a functional amiRNA can be laborious and time-consuming.

The efficiency of RNA interference-mediated gene silencing

RNAi has been described as an efficient method to silence genes in plants (Wesley et al., 2001). The only RNAi construct tested in this study was designed to target *AG*. While some of the plants carrying this construct displayed phenotypes (Figure 3.2) similar to an *ag* mutant, the low penetrance of the mutant phenotype upon expression of the *AG*-RNAi construct (an *ag*-like phenotype was only seen in 4 independent transformants out of 52) implied that the isolation of an inducible RNAi construct would be difficult. Therefore, the *AG4*-amiRNA, whose over-expression resulted in plants with *ag*-mutant phenotypes at a higher frequency, appeared to be the better option for creating an inducible system for an *AG* knockdown.

Whether RNAi or amiRNA-based gene silencing is more effective is currently unclear (Ossowski et al., 2008). Although amiRNAs theoretically offer significant advantages in comparison to RNAi (see section 1.5.2.1), RNAi-based silencing could be used to identify complementary binding sites that are accessible to small interfering RNAs, and therefore should also be accessible to amiRNAs. Then amiRNAs could be designed that hybridize to these regions instead of designing many amiRNA constructs to tile the transcript of interest.

The over-expression of a dominant negative version of AG

The SRDX repressor domain, which is derived from the transcription factor SUPERMAN, was translationally fused to the coding region of *AG* and expressed in wild-type plants under the control of the constitutive 35S promoter. This approach has been described previously to result in a phenocopy of *ag* mutants (Mitsuda et al., 2006). However, the construct generated and expressed in this study failed to reproduce this observation (Figure 3.3B – D). Another construct was generated that expressed a fusion of AG to the WUSB domain, a dominant repressor domain from the transcription factor WUSCHEL. Although some plants that expressed the *AG-WUSB* fusion produced *ag*-like phenotypes (Figure 3.3E – G), some of the phenotypes observed in other transformants were similar to those found in the *35S::AG-SRDX* transgenic lines. Interestingly, the *AG-SRDX* and *AG-WUSB* constructs had an identical linker sequence (noted linker 1 above) between the AG moiety and the repressor domain. Hence, it was hypothesized that the linker region used in both constructs could be the cause of the non-*ag*-like phenotypes. The linker 1 (see Section 2.13.3) used for the fusion in the present study may cause the SRDX/WUSB domains to fold in such a way that it interferes with AG binding. Unfortunately, the sequence of the linker region used by Mitsuda et al. (2006) has not been published, which prevents us from testing if using the same sequence would result in *ag*-like phenotypes.

In order to test the effect of the linker on the function of the *AG-SRDX* fusion, an alternative, longer linker region (linker 2, see Section 2.13.3) was used to fuse AG to the SRDX domain. Several independent transgenic lines were isolated, but no mutant phenotypes were observed. Because relatively few transformants were obtained and the expression levels of the *AG-SRDX* transgene were not investigated in these lines, the effect of the linker region on the activity of the *AG-SRDX* protein is currently unclear.

Complementing the *ag-1* phenotype with an epitope-tagged version of AG

Populations segregating the *ag-1* allele were transformed with both *pAG::AG-GFP* (denoted *AG-GFP*) and *pAG::AG-6xHA* (denoted *AG-6xHA*) translational fusion constructs to test whether the epitope-tagged AG proteins could complement the *ag* mutant phenotype.

While the AG-GFP protein was able to compensate for the loss of the endogenous AG protein in an *ag-1* mutant (Figure 3.8), plants expressing AG-6xHA showed a neomorphic phenotype (Figure 3.7), suggesting that the epitope tag interfered with the function of the AG protein. It is possible that addition of a 6xHA tag at the C-terminus of AG affects the folding of this transcription factor, which may impair its function in different ways. For example, aberrant folding may cause the AG-6xHA fusion to interact with different proteins or bind different target genes than in normal circumstances. Alternatively, the epitope tag may simply obstruct the DNA binding domain or protein interaction domain of AG.

Because expression of the AG-GFP fusion protein rescued an *ag-1* mutant, the corresponding construct was introduced into the floral induction system (see Section 3.2.3.3), and the resulting plants were tested for their response to dexamethasone treatment. Unexpectedly, the 35S::AP1-GR pAG::AG-GFP *ap1-1 cal-1 ag-1* (denoted 35S-FIS AG-GFP hereafter) plants failed to respond in the same manner as the floral induction system without the *AG-GFP* transgene in the background (Figure 3.9). Therefore, this transgenic line was crossed with the improved pAP1::AP1-GR floral induction system generated in this study (see Section 5.2.2). When F₁ plants were treated with a dexamethasone containing solution, synchronous flowering was observed in patches of the inflorescence-like meristem (not shown). This showed that expression of the *pAP1::AP1-GR* transgene could partially circumvent the non-flowering phenotype of 35S-FIS AG-GFP plants upon treatment with a dexamethasone-containing solution. Interestingly, the *pAP1::AP1-GR* transgene could not fully restore synchronous flowering in the 35S-FIS AG-GFP background, which suggested that flowering was being actively disrupted (as opposed to a passive, non-functionality of the 35S::AP1-GR transgene). When pAP1::AP1-GR pAG::AG-GFP *ap1-1 cal-1 ag-1* (denoted *pAP1-FIS* AG-GFP hereafter) plants that did not contain the 35S::AP1-GR transgene were isolated and treated with dexamethasone, synchronous flower development as well as the formation of sepals and petals was restored (Figure 3.10).

It remains unclear why expression of the AP1-GR fusion from the 35S promoter (but not from the endogenous *AP1* promoter), combined with the expression of the AG-GFP fusion impaired induction of flower formation upon dexamethasone treatment. A possible explanation stems from the fact that AG and AP1 proteins can interact (Riechmann et al., 1996) and that AG has been shown to be expressed in the inflorescence-like meristems of *ap1 cal* plants before dexamethasone treatment (Wellmer et al., 2006). It is therefore possible that, prior to any dexamethasone treatment, the AG-GFP fusion is present in the

inflorescence-like meristems of the *35S-FIS* plants. Hence, AG-GFP could form a stable complex with the AP1-GR fusion and prevent it from activating flower formation upon dexamethasone treatment. In contrast, expression of AP1-GR from the endogenous *AP1* promoter may not result in early expression of the AG-GFP fusion (*i.e.* before dexamethasone treatment), thus preventing the adverse effect of AG-GFP on the function of the AP1-GR fusion.

Because the endogenous (untagged) AG protein does not have this adverse effect in the *35S-FIS*, this hypothesis would also imply that the GFP tag itself affects the interaction between AG and the AP1-GR fusion. Alternatively, the *AG-GFP* transgene may be expressed at higher levels than the endogenous *AG*. Indeed, if the GFP tag decreased the overall activity of AG, higher levels of *AG-GFP* would be required in order to produce phenotypically normal flowers. Therefore, when screening for a rescue of the *ag-1* phenotype, a transgene expressing *AG-GFP* at high levels may have been selected. This potential over-expression of *AG-GFP* may lead to inappropriate interactions between AG-GFP and AP1-GR, and it may be this interaction that blocks floral initiation by AP1.

Although no data is available to date to confirm or reject this hypothesis, it could be tested by determining if AG-GFP is present before dexamethasone treatment in the inflorescence-like meristems of the *pAP1-FIS*. The expression levels of the *AG-GFP* transgene and of the endogenous AG could also be compared between the *35S-FIS* and the *pAP1-FIS*. Furthermore, the interaction between AP1-GR and AG-GFP, as well as between AP1-GR and the endogenous (or untagged) AG could be compared.

The activated AlcR transcription factor can influence the expression of endogenous *Arabidopsis* genes

The *AG4-miRNA* was placed under the control of the *35S::AlcR/pAlcA* ethanol-inducible system, which has been described as a highly efficient, tightly regulated gene expression system (Deveaux et al., 2003; Roslan et al., 2001). Using the FIS containing the *pAlcA::AG4-miRNA* transgene (FIS-AG4), I performed genome-wide expression profiling analysis (Section 3.2.5), following dexamethasone and ethanol treatments. Upon ethanol induction, a substantial number of genes were differentially expressed in *amiRNA*-expressing line. A similar result was observed when an experiment was carried out using an *amiRNA* that targets *AP3* mRNA (Wuest, 2011). In order to estimate the effect that the

activated AlcR transcription factor has on global gene expression, an amiRNA (referred to as DUMMY-amiRNA) that was designed not to be complementary to any known *Arabidopsis* transcript was placed under the control of the *35S::AlcR/pAlcA* system. The resulting expression profiling experiments indicated that the DUMMY-amiRNA system was influencing the expression of many genes. Given that the sequence under the control of the *35S::AlcR/pAlcA* expression system is not predicted to target any known *Arabidopsis* transcripts, the differential gene expression observed is likely to be due to the activity of the activated AlcR transcription factor.

The genes identified as differentially expressed in this DUMMY-amiRNA line were removed (“filtered”) from the FIS-AG4 dataset. This filtered dataset was then analysed for the enrichment of particular Gene Ontology (GO) terms. As expected, among the genes that were differentially expressed upon expression of the *AG4-amiRNA*, there was an enrichment of GO terms associated with “Developmental Processes”. However, despite the use of the DUMMY-amiRNA to filter the list of differentially expressed genes, 34 GO-terms were associated with “Stress and Pathogen Processes”, which is unlikely to be the result of decreased levels of *AG* mRNA.

The effects of the activated AlcR transcription factor on global gene expression were unexpected. Although it had been observed that growth can become inhibited in ethanol-treated plants containing the AlcR transcription factor (Roslan et al., 2001), this effect can be circumvented by lowering the treatment times or the ethanol concentration (Roslan et al., 2001). The *AG4-amiRNA* ethanol-responsive line that was ultimately used (line #52, see Section 3.2.2) did not display any developmental defects, other than *ag*-related phenotypes, when treated with ethanol vapour for 6 hours, which is sufficient to reduce *AG* mRNA levels to 10 – 20 % of wild-type levels (Figure 3.6A). Therefore, it was deemed acceptable for use in a genome-wide expression profiling study. The same observations were made with plants expressing the *AP3-amiRNA* under the *35S::AlcR/pAlcA* system.

To my knowledge, no expression profiling experiments have been previously performed using the *35S::AlcR/pAlcA* expression system. Therefore, the effects of the activated AlcR transcription factor on endogenous transcript levels could not have been foreseen. Nevertheless, several strategies can be implemented in order to identify the genes that are influenced by the activated AlcR transcription factor. As mentioned in Section 3.4, an alternative line expressing the DUMMY-amiRNA is being prepared that expresses the *AlcR* mRNA at levels comparable to those of the FIS-AG4 and FIS-AP3 lines. Expression

profiling experiments can be used to identify genes that are differentially expressed in this new DUMMY-miRNA line. The latter can be compared to the list of differentially expressed genes in the FIS-AG4 and FIS-AP3 lines, so that a more accurate filter is applied.

Another method of identifying genes that are influenced by activation of AlcR is to compare the *35S::AlcR/pAlcA* miRNA expressing system to another inducible system expressing the same miRNAs. To this end, Andrea Raganelli (another member of the Wellmer laboratory) has placed the *AP3*-, *AG4*- and *PI-miRNAs* under the control of the dexamethasone-dependent *LhG4-GR/pOP* two-component system (Craft et al., 2005). To the best of my knowledge, no effects of LhG4-GR on global gene expression have been reported in the literature. Moreover, the use of dexamethasone to activate the GR receptor does not affect plant growth and many microarray studies using dexamethasone-activated protein fusions have been performed (Gomez-Mena et al., 2005; Ito et al., 2004; Wellmer et al., 2006). It is therefore likely that the overlap in the lists of differentially expressed genes obtained upon expression of the same miRNA from two different inducible promoter systems will yield valuable information about genes and processes downstream of the floral homeotic factors.

The most suitable control experiment that can be performed to identify, and thus discard, genes differentially expressed by the activated AlcR transcription factor is to generate an *AG4-miRNA*-resistant version of *AG*. A transgenic line containing an *AG-GFP* translational fusion that can restore the homeotic defects of *ag* mutants has been described here (Section 3.2.3.2). Therefore, the same genomic region used in this construct can be modified to produce synonymous substitutions in the region of the coding sequence that the *AG4-miRNA* targets. Once lines that express this mutated version of *AG* and rescue an *ag-I* mutant have been identified, this transgene can be introduced into the *AG4-miRNA* line used in this study. Genes influenced by activated AlcR can then be identified by microarray analysis. The advantage of this approach is that the same transgenic line carrying the *pAlcA::AG4-miRNA/35S::AlcR* transgene can be used, thereby directly comparing the effects of activated AlcR activity.

Ultimately, however, a different miRNA delivery system would need to be developed, so that network perturbation experiments can be carried out more easily. There are several other inducible expression systems available in plants (for review see (Moore et al., 2006) and preliminary comparisons of these systems has begun in the Wellmer laboratory.

Global expression profiling using the *ag-1* mutant resulted in relatively small fold-change differences of downstream genes

Although many genes known to be involved in flower development were identified as differentially expressed in the *ag* mutant microarray experiments, relatively low fold-change differences were obtained on average (Figure 3.11). While this might accurately reflect the differences between the expression profiles of *ag-1* and wild-type flowers early in development, the moderate response in gene expression could also be due to contamination of the *ag-1* mutant flower samples during preparation. The tissue for these experiments was collected individually (96 plants per time-point) and was subsequently pooled after phenotyping and genotyping. Therefore, the chance of unintentionally contaminating the *ag-1* sample with tissue from wild-type flowers was relatively high. This potential contamination would have the effect of reducing the expression changes that could be observed in this experiment. To test this possibility, I am currently replicating the *ag* mutant experiments on a smaller scale. All plants used for this experiment will be genotyped so that no cross-contamination of samples can occur. qRT-PCR analysis will then be employed to compare the expression in the different samples for selected genes.

3.2.8.2 The role of *AGAMOUS* in flower development

I have generated several genome-wide datasets that aimed at identifying the processes and genes regulated by the C-class homeotic transcription factor *AG*. Using a null mutant comparative approach, I have shown that *AG* is primarily a transcriptional activator of reproductive organ development (Figure 3.13A). Moreover, *AG* has been implicated in processes such as organ boundary regulation and the epigenetic regulation of development (Table 3.2). *AG* appears to initiate transcriptional networks and signalling pathways directly, as shown by chromatin immunoprecipitation combined with ultra high-throughput sequencing (ChIP-Seq) (Section 3.2.7), in order to regulate flower development in a modular manner. In addition to these molecular analyses, I have characterised the morphological effects of *AG* perturbation in a stage-specific manner (Section 3.2.2.2). These phenotypes have indicated that *AG* has distinct functions in early and late flower development.

Given the large volume of data generated in this study, much analysis remains to be done. However, several interesting observations were made that further unravel the gene regulatory network that *AG* controls. In the following sections, I will discuss the relevance of my findings in the context of the current state of the field of flower development.

AG promotes the expression of other transcription factors involved in reproductive organ development

It has been known for over 20 years that *AG* is a master regulator of reproductive organ development (Bowman et al., 1989). However, the manner in which it relays its developmental signal is currently not well understood. In this study, I have used ChIP-Seq experiments to identify *AG* binding sites on a genome-wide scale. Furthermore, I have employed gene expression profiling to detect genes whose expression is controlled, directly or indirectly, by the *AG* transcription factor. The combination of the datasets from these genomic approaches has led to novel insights into the gene regulatory network controlled by this master regulator.

We have found evidence that *AG* directly controls a large number of transcription factor-coding genes to relay its developmental signal in a modular manner. While some of these genes had already been proposed to be directly regulated by *AG* (see Section 1.3.2), many others had not been previously identified. Among the transcription factor-coding genes controlled by *AG* are, for example, *CRC*, the *SHPI/2* genes, the *HECI/2* genes and *VERDANDI* (*VDD*). All of these genes were associated with binding peaks in the *AG*-GFP ChIP-Seq dataset (*AG*-2) and are positively regulated by *AG*, as determined by their down-regulation in the *ag* mutant experiments. The expression of each of these genes is important for the development of different and overlapping aspects of the gynoecium.

CRC suppresses radial growth of the gynoecium while promoting its elongation (Bowman and Smyth, 1999). *SHPI/2* have been shown to redundantly control ovule development, stigma and style development, and medial tissue development in the *Arabidopsis* gynoecium in addition to controlling the differentiation of the dehiscence zone (Colombo et al., 2010; Pinyopich et al., 2003). *HECI/2* are important developmental regulators of the gynoecial style, transmitting tract and stigma (Gremski et al., 2007) while *VDD* is critical for the normal development of the female gametophyte (Matias-Hernandez et al., 2010).

It has been suggested that *VDD* is directly activated by the SHP1/2 and SEEDSTICK (STK) transcription factors (Matias-Hernandez et al., 2010). Given our data it is possible that AG is involved in a complex with these proteins to direct female gametogenesis. Cooperative binding between SHP1/2, STK and AG has been suggested previously (Pinyopich et al., 2003), however, no molecular evidence for this interaction has been provided. Indeed, the binding sites for STK, SEP3 and AG at the *VDD* locus overlap (Matias-Hernandez et al., 2010). This observation adds credence to the suggestion that the transcriptional activity of AG in different tissues is modulated by the transcriptional complexes to which it belongs in that tissue (Honma and Goto, 2001). However, direct interaction of AG with STK and SHP1/2 at the *VDD* locus has not been shown.

Patterning of the gynoecium is also controlled by *AG*. It is known that abaxial/adaxial polarity in carpel development is redundantly regulated by *CRC* and *KANADI* (*KAN*) (Pekker et al., 2005). Double mutant *kan1 kan2* gynoecia develop adaxial cell types on their abaxial axes (Eshed et al., 2001). Additionally, the ovules of *kan1 kan2* double mutants form aberrantly shaped outer integuments (Kelley et al., 2009). *KAN4* is also involved in ovule development (*KAN4* is also known as *ABERRANT TESTA SHAPE*). The inner and outer integuments of *kan4* ovules fail to develop as separate structures, leading to the development of heart-shaped seeds. This phenotype has been interpreted as a polarity defect in ovule development (Kelley et al., 2009). *KAN2* and *KAN4* were differentially expressed in the *ag* mutant experiments and differential binding was observed in the AG-GFP ChIP-Seq datasets at the *KAN2* and *KAN4* loci. Given this information, *AG* appears to promote the specification of the adaxial/abaxial axes of carpels and ovules in flower development by directly activating transcription factors.

AG also appears to promote gynoecium development by protein modification through direct activation of the receptor-like kinase gene *ERECTA-LIKE1* (*ERL1*). The *Arabidopsis* accession *L-er* (Landsberg *er-105*) contains a mutated version of the *ERECTA* (*ER*) gene, which is the underlying cause of their compacted inflorescences relative to Columbia-0 (*Col-0*) plants (Shpak et al., 2003). Triple mutant *er erl1 erl2* plants are defective in cell proliferation and the siliques of these plants contain wider valves and shorter, broader styles than wild-type *L-er* or *Col-0* plants (Shpak et al., 2004). *AG* binds to genomic regions near all three genes, however; only *ERL1* was found to be differentially expressed in the *ag* mutant experiments.

SPOROCTELESS (SPL) has been shown to directly regulate microsporogenesis during stamen development (Ito et al., 2004). Moreover, *SPL* expression is directly responsive to *AG* activity and *AG* has been shown to directly bind a region downstream of the *SPL* coding sequence (Ito et al., 2004). Direct binding of *AG* to the *SPOROCTELESS (SPL)* locus was not observed in the *AG*-GFP ChIP-Seq datasets. In this study, we found that *SPL* was down-regulated in the *ag* mutant experiments, however no binding enrichments were observed. Instead, binding of AP3-GFP and PI-GFP to a region upstream of the *SPL* coding sequence was observed (Wuest, 2011). The lack of differential binding of *AG*-GFP to the described *SPL* locus was unexpected as mutagenesis of the putative binding site at the 3' end of the *SPL* coding region resulted in the abolishment of binding in an electrophoretic mobility shift assay and severely decreased GUS staining in a transcriptional fusion between this region and *GUS*. The non-mutated sequence region was bound in the EMSA and strong GUS staining was observed in the non-mutated transcriptional fusion with *GUS*. It is possible that this binding peak was missed in this study, or *AG* could bind to this locus at later stages of development. Nonetheless, it appears that AP3/PI positively regulate *SPL* by directly activating its transcription whereas *AG* promotes the expression of *SPL*, but possibly not in a direct manner. Other known genes involved in stamen development include the basic leucine zipper (bZIP) transcription factors bZIP21 and bZIP65 and the C2H2-type zinc-finger transcription factors JAGGED and NUBBIN.

bZIP21 has been shown to redundantly regulate male gametogenesis together with *bZIP65*. Moreover, these genes appear to be activated by *SPL* and share mutant phenotypes with *roxy1 roxy2* double mutants (Murmu et al., 2010). The expression of *bZIP21* was found to be reduced in the *ag* mutant experiments and binding of the *AG*-GFP protein was observed in ChIP-Seq experiments, suggesting direct regulation.

The zinc-finger proteins JAGGED (*JAG*) and NUBBIN (*NUB*) are involved in regional cell division and organ growth. *jag nub* double-mutant plants have severe growth defects in the distal regions of carpels and stamens (Dinneny et al., 2006). *NUB* has been proposed as a direct target of *AG* (Gomez-Mena et al., 2005) whereas *SEP3* binding was observed near the *JAG* locus (Kaufmann et al., 2009). In this study, we have observed down-regulation of both *JAG* and *NUB* expression in the *ag* mutant experiments, whereas binding of the *AG*-GFP protein was observed in regions upstream and downstream of the *JAG* locus.

All of the genes discussed above, except *ER*, *ERL1* and *ERL2*, encode transcription factors, which strongly suggests that AG controls reproductive development in a modular manner by activating other transcription factor-coding genes. Furthermore, the mutant phenotypes of these genes show that formation of particular tissues within the developing reproductive organs is controlled by different transcription factors that have been activated by AG. This idea is consistent with the findings of the GO term analyses of both the AG-GFP ChIP-Seq and the *ag* mutant experiment datasets, where *Regulation of Transcription* is the top and fifth-highest ranked GO term, respectively (Table 3.2). Additionally, AG appears to predominately promote transcription, as over 80% of the genes differentially expressed at day 5 of the *ag* mutant experiments were down-regulated. Moreover, gene set enrichment analysis identified the GO term *Negative Regulation of Transcription* as being down-regulated in the *ag* mutant experiments. This suggests that AG activates both transcriptional activators and repressors.

The role of AG in activating transcription factor-coding genes was further highlighted by the findings of the protein family enrichment analyses. These analyses indicated that AG and AP3/PI co-regulate several transcription factor families including the B3 transcription factors and the C2H2-type zinc finger transcription factor families. Additionally, the homeodomain transcription factor family was over-represented in the AG protein family enrichment analysis. The protein family set enrichment analysis, which tested for a directionality in expression of groups of genes in the *ag* and *pi* mutant datasets implied that AG activates members of these families whereas PI both activates and represses members of these families.

In sum, the data outlined above suggest that AG predominantly functions by activating transcription factor-coding genes that then mediate diverse processes during the development and the patterning of the reproductive floral organs. While several of these AG target genes have been functionally characterized, many other await a detailed characterization using the powerful reverse genetic approaches available for *Arabidopsis*.

AG directly regulates genes involved in floral organ boundary formation

The compartmentalisation of floral organs in a defined pattern is a hallmark of eudicot flower development (Breuil-Broyer et al., 2004). In *Arabidopsis*, flowers contain 4

outer sepals, 4 second-whorled petals, six third whorl stamens and 2 fused carpels in the centre of the flower (Bowman et al., 1989). There are several mutants described, apart from the floral homeotic mutants, where floral organs form in an inappropriate whorl at the expense of another organ type or where organs are fused to each other. These genes include *RABBIT EARS (RBE)*, *SUPERMAN (SUP)* and *CUP-SHAPED COTYLEDONS 1 (CUC1)* and *CUC2*. These genes do not appear to specify floral organ fate directly, instead they either repress homeotic gene expression or slow the rate of cell division at a given position between developing organs, or potentially both. Homeotic genes can also play a cadastral role in flower development with the mutual repression of the A class genes, *API* and *AP2*, and *AG* forming a central theme of the ABC model of floral organ identity specification (Gustafson-Brown et al., 1994; Wollmann et al., 2010).

In *rbe* mutants petals are often replaced with stamens or stamen-like organs. In addition, first whorl sepals are often fused to each other (Krizek et al., 2006; Takeda et al., 2004). The petal to stamens transformations observed in *rbe* mutants were attributed to the expanded expression of *AG* into the second whorl (Krizek et al., 2006). Therefore, *RBE* has been proposed to negatively regulate *AG* expression. In addition, fused first whorl sepals were proposed to form due an inhibition of growth in first whorl cells adjacent to *RBE* expressing cells (Krizek et al., 2006). The mechanism underlying this inhibition was proposed to be non-cell autonomous.

Fused first whorl sepals are also observed in *cuc1 cuc2* double mutants (Ishida et al., 2000). *CUC1* and *CUC2* are involved in boundary maintenance between whorls 1, 2 and 3 (Breuil-Broyer et al., 2004). Apart from fused sepals, *cuc1 cuc2* double-mutant plants display a range of other defects including unfused septa, fused stamens and stamens that are fused to carpels (Aida et al., 1997). *CUC2* is expressed in the boundary regions of organ primordia and in the septum (Ishida et al., 2000). The *CUC* genes also appear to inhibit cell proliferation between developing organs (Breuil-Broyer et al., 2004)

sup mutant flowers also contain stamen-carpel fusions (Bowman et al., 1992). The *SUP* gene was originally thought to negatively regulate *AP3* and *PI* expression in the fourth whorl, however, it has been subsequently shown that *SUP* acts to maintain the boundaries between whorls three and four (Sakai et al., 2000). This is likely accomplished by regulating cell proliferation within these whorls (Breuil-Broyer et al., 2004).

In this study, we found that genomic regions associated with the *API*, *AP2*, *RBE*, *SUP* and *CUC1/2* loci were directly bound by the AG-GFP fusion protein. Additionally, the

expression of *RBE* was increased in the *ag* mutant experiments, indicating a direct negative repression of *RBE* expression by AG. The *SUP*, *CUC1/2* genes were down-regulated in the *ag* mutant experiments, which indicates that AG directly activates their expression. Finally, although genomic regions associated with *API* and *AP2* were bound by AG-GFP, no significant expression differences were observed. For *API*, the use of the 35S CaMV promoter to drive *API-GR* expression may have masked any expansion of *API* into the centre two whorls. Whereas the expansion of *AP2* into the AG domain in *ag-1* mutants is quite subtle (Wollmann et al., 2010), therefore, large expression differences might not be expected in this experimental design.

Taken together, these results indicate that AG actively controls the formation of organ boundaries by controlling the expression of several transcription factor-coding genes with known roles in this process.

AG is involved in the epigenetic regulation of flower development

In germ cells and early embryos of *Arabidopsis*, epigenetic modifications occur on a genome-wide scale (Feng et al., 2010). These modifications occur to suppress the expression of potentially deleterious non-genic sequences such as transposable elements (Law and Jacobsen, 2010). Examples of these modifications include histone deacetylation, histone methylation and cytosine methylation, which are used to alter the transcriptional state of a given gene. Moreover, these modifications can be inherited from generation to generation (Feng et al., 2010).

Methylation of lysine 9 of histone H3 (H3K9) is correlated with transcriptional repression (Rangwala and Richards, 2004). A loss of H3K9 methylation in the egg cells has been attributed to a mutation in *CHROMOMETHYLASE 3* (*CMT3*). Moreover, these mutant plants contain embryos that develop aberrantly (Pillot et al., 2010). *CMT3* is a DNA methyltransferase which, together with the histone methyltransferase, *KRYPTONITE*, maintains cytosine methylation patterns in the DNA context of CHG (where H = A, T or C) and a variable reduction of CHG methylation is observed in *cmt3* and *kyp* mutant plants (Feng et al., 2010). The lack of CHG methylation in ovules could lead to ovule abortion or the transmission of mutations to offspring. It is therefore essential that these methylation patterns be maintained.

Interestingly, *Arabidopsis* pollen has reduced expression levels of *CMT3* and *KYP* (Feng et al., 2010). In contrast, *DECREASE IN DNA METHYLATION 1* (*DDMI*) is exclusively expressed in pollen at a high level. *DDMI* plays an important role in chromatin remodelling and is required for DNA and histone methylation, and transposon silencing. The mature pollen grain consists of two sperm cells, which contains the paternal genetic information, and a vegetative nucleus, which migrates with the sperm cells during pollen tube growth but does not contribute any genetic information to the zygote or the endosperm (McCormick, 1993). *DDMI* was found to be highly expressed in the sperm cells but not in the vegetative nucleus (Slotkin et al., 2009). Recently, a model was proposed where hypomethylation of the genomic content of the vegetative nucleus results in the activation of transposable elements and subsequently a production of short interfering RNAs (siRNAs) from transposon transcripts that migrate to the sperm cells of pollen (Feng et al., 2010; Slotkin et al., 2009). This mechanism was proposed to amplify the “protective” signal within sperm cells against potentially mutagenic sequences.

Another example of potentially non-cell autonomous transposon silencing occurs during female gametogenesis (Feng et al., 2010). siRNAs guide RNA-directed DNA methylation (RdDm) machinery to cognate DNA sequences via the ARGONAUTE (AGO) proteins (Law and Jacobsen, 2010). A reactivation of transposons in the ovule is observed in *ago9* mutant plants. However, *AGO9* is not expressed in the megaspore mother cells or the female gametophyte, it is expressed in surrounding companion cells (Olmedo-Monfil et al., 2010). The transposons targeted in the companion cells are similar to the reactivated transposons in the vegetative nucleus of pollen grains (see above). The AGO9 protein appears to interact with the RdRm machinery in order to silence these transposons (Feng et al., 2010). It is possible that the siRNAs that guide AGO9 activity translocate to the female gametophyte and associated cells; however, no evidence is available for this at present.

In this study, we found that *CMT3*, *KYP* (a SET-domain gene), *DDMI* (a Helicase-C gene), *AGO3-6* and *AGO9* were significantly down-regulated in the *ag* mutant experiments. Only genomic regions surrounding *AGO5*, which has no described mutant phenotype (Vaucheret, 2008), were significantly enriched in the ChIP-Seq dataset (AG-2 replicate only). Although no expression differences were observed, genomic regions near the *AGO1* and *AGO10* loci were enriched in the AG-GFP ChIP-Seq dataset. This is also in line with the protein family enrichment analysis of the *AG* mut/WT data that found the PAZ domain family, which includes the AGO proteins, is overrepresented. Furthermore, the protein

family enrichment analyses identified other SET-domain proteins and Helicase-C proteins as being differentially expressed.

These results strongly suggest that *AG* is a positive, but in most case, an indirect regulator of genes involved in transposon silencing and possible also of other epigenetic mechanisms. Because in contrast to the data from the *AG* experiments, the AP3/PI datasets (Wuest, 2011) were not enriched in genes or processes involved in the epigenetic regulation of development, it can be assumed that this function of *AG* is restricted to the fourth whorl and thus mostly affects female reproductive development.

***AG* controls distinct developmental processes at early and late stages of flower development**

I used an inducible perturbation system to dissect the phenotypes resulting from an *AG* knockdown at distinct stages of flower development (Figure 3.4 and 3.5). Several notable observations were made during this experiment. First, *AG* plays a role in the regulation of stamen elongation and gynoecium shape in late flower development in addition to its known role in organ identity specification. Second, stamens are relatively robust to *AG* perturbation whereas determinacy of the floral meristem and carpel identity are more readily affected by a reduction of *AG* activity. Third, the dynamic knockdowns in the *AG* perturbation experiments may have highlighted the stochastic nature of the flower development gene regulatory network. These topics are discussed below.

Activation of the *AG4-miRNA* during late flower development retarded the elongation of stamen filaments and disturbed the shape of the gynoecium. As outlined in Section 1.3.1, gibberellic acid (GA) and jasmonic acid are important for stamen elongation. *AG* has been proposed to activate the GA synthesis enzyme, *GA4* (Gomez-Mena et al., 2005). Although differential expression of, or binding to this gene was not observed here, two functional C₁₉-GA 2-oxidases (*ATGA2OX4*, *ATGA3OX3*), which function to limit the availability of bioactive GA, (Rieu et al., 2008) were up-regulated in the *ag* mutant experiments and genomic regions associated with these genes were differentially bound by AG-GFP. Therefore, *AG* appears to promote GA signalling by down-regulating negative regulators of GA signalling. Given the role described for GA in cell elongation and growth, it is likely that the potential lack of bioactive GA in ethanol-treated flowers contributed to lack of filament outgrowth. The underlying cause behind the aberrant shapes of carpels

defective at late stages for *AG* function is currently unknown. A more detailed analysis of the gynoecial tissues affected is required to make a hypothesis. Nevertheless these phenotypes and molecular analyses strongly suggest that *AG* control aspects of reproductive growth late in flower development.

The sensitivity of each reproductive organ and determinacy to *AG* perturbation was also noted. From previous studies (Ito et al., 2007; Sieburth et al., 1995) it was expected that floral meristem determinacy would be the most sensitive to *AG* perturbation, followed by carpel identity, with stamen identity requiring the least amount of *AG* activity and therefore being the least sensitive to *AG* perturbation. The study of Ito et al. (2007) showed that one treatment of a 35S::*AG*-GR *ag-1* line was sufficient to restore the development of two fused carpels in the centre of the flower, whereas four consecutive treatments (12 or 24 h apart) were required for full stamen development. Moreover, when treatments were separated by 48 h, stamens failed to mature fully. These results did not indicate the levels of *AG* that are required for stamen development, however, they did highlight the large developmental window in which *AG* is required in order to completely direct stamen development.

The effects of a single pulse of ethanol in the pAlcA::*AG4* line was sufficient to disrupt floral meristem termination and to strongly affect carpel development. Two pulses of ethanol abolished fourth whorl carpel identity and further disrupted determinacy whereas up to four pulses of ethanol over an 8-day period were required to transform stamens into petals. Additionally, when ethanol treatments were separated by 48 h, stamens were never completely transformed to petals, anther tissue was always present. However, carpel development was abolished in these plants and floral indeterminacy was still present. These results are largely consistent with published data.

The flowers of ethanol treated pAlcA::*AG4* plants were affected differently, depending on the stage of flower development a given flower was in when treated with ethanol. Flowers treated at late stages were defective in growth and flowers treated at early stages were defective in organ specification. After flowers with third and fourth whorl homeotic transformations had developed, subsequent flowers often produced fused sepals and petaloid sepals in the first whorl. The levels of *AG* mRNA were presumably recovering in these flowers. Moreover, extra organs of the perianth in the first and second whorls was only observed in these flowers. *AG* function may have been disrupted for a very brief period during the early developmental stages of these flowers to create a fluctuation of *AG* expression. Since *AG* is not expressed in the first two floral whorls, these phenotypes could

be explained by the stochastic and dynamic nature of the flower development gene regulatory network.

The influence of *AG* on the GRN of flower development is not a linear system; it is a complex dose-dependence relationship, which is exemplified by the dose dependencies of stamens, carpels and floral meristem determinacy on *AG* activity. It has been shown that bursts of expression of a transcription factor can result in a high level of variability of transcript levels under the control of that transcription factor, even if the mean levels of down-stream expression are equal (Raj and van Oudenaarden, 2008). Therefore, the dynamic knockdown of *AG* mediated by the *AG4-amiRNA* may result in high variability of downstream transcript expression. Given the role of *AG* in organ boundary regulation, this level of variation could filter through the GRN to produce a wave of highly variable gene expression that ultimately affects the development of first and second whorl organs. The *RBE* genes or/and the *CUC1/2* genes could be affected in this way as mutations in both genes result in flowers with fused sepals. The transformation of sepals to petals could be a consequence of a fluctuation in *AP2* mRNA levels. Over-expression of *microRNA-172* (which regulates the accumulation of *AP2* mRNA and *AP2* protein) results in conversion of sepals to petals in *Nicotiana benthamiana* (Mlotshwa et al., 2006). However, this has not been described in *Arabidopsis*.

In summary, the ethanol-induced dynamic knockdown of *AG* mRNA in the pAlcA::AG4 transgenic lines resulted in the full conversion of stamens and carpels to petals and sepals. This line can be used to further dissect the stage-specific roles of *AG* during flower development through detailed phenotypic analysis of the ethanol knockdowns and potentially by tissue-specific expression of the *AG4-amiRNA*.

3.2.9 Future Work

In this study, I have generated several transgenic lines that have allowed us to assess the role of *AG* during flower development. However, we also experienced some experimental problems, especially with the perturbation of *AG* function at different stages of early flower development using the FIS-AG4 amiRNA line. Possible solutions to these problems have been outlined in Section 3.2.8.1 and we will attempt to generate higher quality data for these experiments in the near future.

Additionally, the first biological replicate of the AG-GFP ChIP-Seq experiment (AG-1) was suboptimal in comparison to the second replicate (AG-2) in that binding enrichments were not observed for several control genes. Therefore, a third ChIP-Seq replicate will be generated in the near future to confirm the findings made based on the AG-2 dataset.

A new line of experimentation will aim at identifying stage-specific differences in the binding patterns of AG. To this end, we will conduct additional ChIP-Seq experiments using the FIS AG-GFP line generated in this study and flowers of different stages of development. The data from these experiments will then be combined with those from the inducible *AG* perturbation experiments described above and should allow us to correlate changes in gene expression with changes in AG binding patterns. This work should yield new, detailed insights into the dynamic behaviour of the gene regulatory network downstream of AG.

Chapter 4 – Investigating the role of KNUCKLES in flower development

4.1 Introduction

As discussed in Section 1.3.3, the C2H2 zinc-finger transcription factor KNUCKLES (KNU) has been implicated in the control of floral meristem determinacy through an AG-dependent pathway (Sun et al., 2009). In particular, *KNU* may act downstream of AG in the gene regulatory network controlling determinacy. The flowers of *knu-1* mutant plants contain ectopic stamens and carpels within the fourth whorl (Figure 4.1A), which results in bulging of the siliques and a knuckle-like appearance (Payne et al., 2004). Additional defects found in *knu-1* mutants include the formation of tricarpeloid siliques, defects in pollen development, extended gynophores and exaggerated nectaries (Payne et al., 2004).

Genomic fragments spanning regions up and downstream of the *KNU* coding sequence were shown to fully restore wild-type development of *knu-1* plants (Payne et al., 2004). Using translational fusions of the *KNU* coding sequence with that of the β -glucuronidase (*GUS*) reporter gene driven by these regulatory regions (Payne et al., 2004), it was found that the expression of *KNU* in flowers begins at the base of the developing gynoecium at stage 6 of flower development (Figure 4.1B) and continues there until stage 13. *GUS* expression was also observed in specific tissues of the female and male reproductive organs, such as in megaspore mother cells and the embryo sac (Figure 4.1C – F).

It has been proposed that the phenotypes observed in *knu* plants are due to the persistent expression of *WUS*, which leads to prolonged activity of stem cells within the floral meristem. Sun *et al.* (2009) have suggested that KNU antagonises *WUS* expression. They argued that the timing of *KNU* expression, which coincides with the down-regulation of *WUS*, is crucial for the termination of the floral meristem (Sun et al., 2009). Given the phenotypes observed in *knu* plants and its expression pattern, it is possible that *KNU* is also involved in processes other than meristem determinacy.

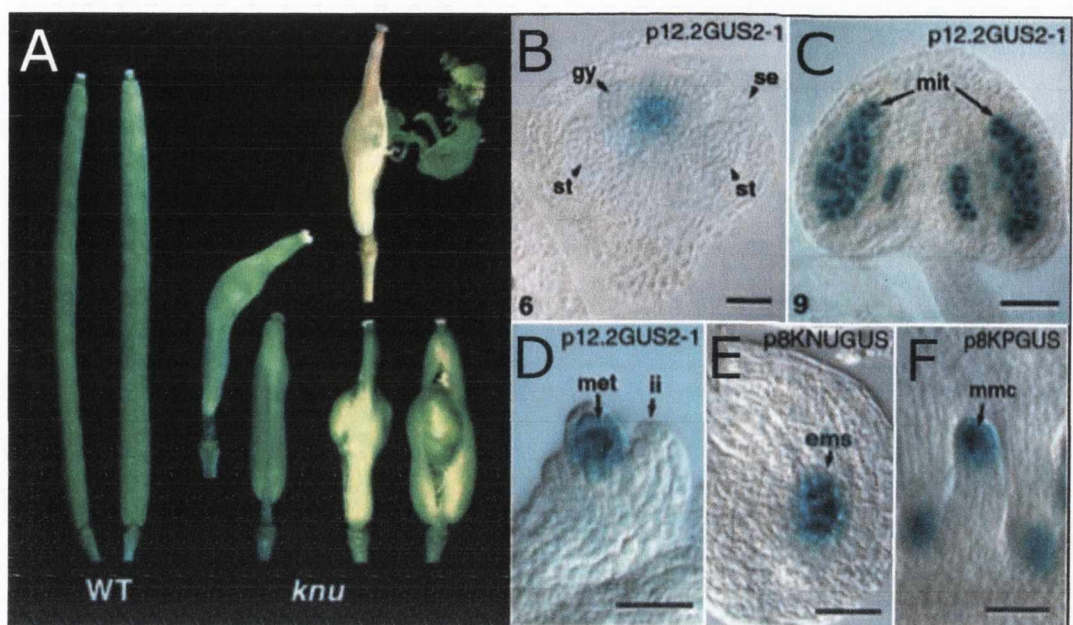


Figure 4.1 - Phenotypes observed in *knu-1* gynoecia and expression pattern of a *KNU-GUS* reporter. (A) *knu-1* flowers produce extra carpels and stamens within their gynoecia. (B – F) Expression of the *pKNU::KNU-GUS* reporter in wild-type flowers. GUS activity was detected in (B) the centre of stage 6 flowers; (C) the microspore tetrads cells of anthers (mit); (D) megaspore tetrads (met); (E) the embryo sac (ems) and (F) the megaspore mother cells (mmc). Figure adapted from (Payne et al., 2004).

In order to obtain detailed insights into the function of *KNU* in flower development and to identify genes whose expression is influenced by *KNU* activity, I aimed to generate transgenic lines that expressed an amiRNA specific for *KNU* mRNA. In addition to the use of amiRNAs, approaches such as ectopic expression of *KNU* were explored as well. Moreover, I attempted to isolate epitope-tagged *KNU* transgenic lines in order to identify the genomic regions bound by *KNU*. The results of this work are presented below.

4.2 Results

4.2.1 Identification of an amiRNA that efficiently perturbs *KNUCKLES* activity

In order to identify the processes that *KNU* activity influences, I aimed to generate

transgenic lines that would allow the monitoring of gene expression changes during early flower development in response to *KNU* perturbation. To this end, I used again the floral induction system (FIS) described above (see Section 1.6), as it allows the collection of large numbers of synchronized floral buds. For the perturbation of *KNU* activity, I decided not to use available *knu* mutant alleles (i.e. an experimental strategy that would be similar to the one outlined for the characterization of *AG* function) because they are either in a Columbia-0 (Col-0) (Sun et al., 2009) or in a Wassilewskija (Ws) (Payne et al., 2004) background, while the FIS has been generated in the *L-er* accession (Wellmer et al., 2006). Consequently, generating a version of the FIS that could be used to study gene expression changes in *knu* mutant flowers would have required time-consuming backcrosses. Therefore, I aimed to generate an amiRNA that could efficiently target the *KNU* mRNA for degradation or translational repression. This amiRNA overexpression line could be used instead of a null allele in the background of the floral induction system. The use of an amiRNA would also allow the application of an inducible perturbation approach, which could lead to the identification of genes directly responsive to *KNU* activity.

An amiRNA targeting the *KNU* mRNA was designed as described in Section 2.13.9. This amiRNA was placed under the control of the CaMV 35S promoter and expressed in wild-type plants (Figure 4.2A). Sixty-three independent transformants were isolated. Forty of these transformants displayed a *knu*-like mutant phenotype in the first generation, however, the penetrance of the phenotypes varied among individual lines. Although all 40 of these independent transformants produced at least one silique with a “knuckle” phenotype (Figure 4.2B), not every silique on a given transgenic plant was affected. Nevertheless, ectopic stamens and carpels were found within these “knuckle” siliques (Figure 4.2C). Locules and pollen were observed on the anthers of the ectopic stamens, and ovules appeared to mature correctly within the ectopic carpels (Figure 4.2C). In addition, a subset of the siliques were tricarpeloid (Figure 4.2D).

Although these mutant phenotypes indicated that this amiRNA was functional, their low penetrance made it difficult to determine how efficient the perturbation of *KNU* activity was. Therefore, the number of bulged siliques produced from the primary inflorescences of

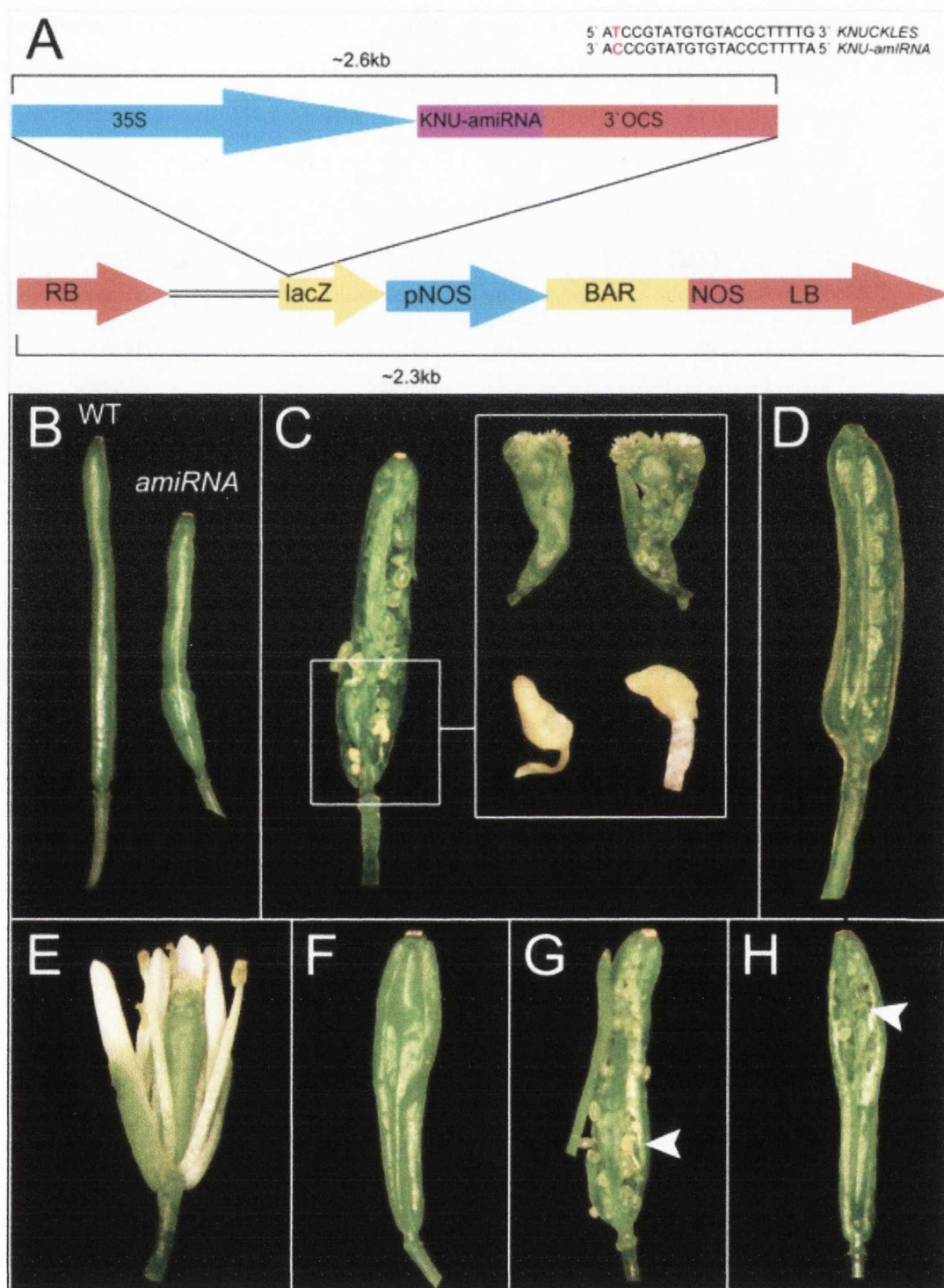


Figure 4.2 Over-expression of an amiRNA targeting *KNU*. (A) A schematic of the T-DNA mediating the expression of the *KNU-amiRNA* under the control of the 35S *CaMV* promoter. The sequence of the *KNU-amiRNA* and the region of the coding region it targets are shown. (B – H) Phenotype of plants constitutively expressing different *KNU-amiRNAs* in a wild-type background. (B – D) Phenotypes observed for line #28. In these flowers, levels of *KNU* mRNA were reduced to ~25% of wild-type levels. (B) Siliques of 35S::*KNU*-

plants was counted (Figure 4.3A, red bars). At most, 10% of siliques produced on a given T1 plant contained a “knuckle” (with the exception of line #6, see below). In order to determine if the phenotype correlated with the levels of the *KNU* mRNA, the axillary inflorescences from nine of the 40 primary transformants were harvested and RNA was extracted. Once cDNA synthesis was completed, the levels of *KNU* mRNA were assayed by qRT-PCR and compared with wild-type levels (Figure 4.3A, blue bars). Two (#18 and #23) of the 9 lines tested expressed *KNU* to 60-85% of wild-type levels. These lines showed mild *knu* phenotypes (*i.e.* shorter siliques with no ectopic organs, Figure 4.2A). In the other seven lines, the number of *knu*-like siliques per plant was between 3% and 10%. *KNU* mRNA levels in these seven lines were as low as 25% of wild-type levels (Figure 4.3A). Overall, the reduction of *KNU* mRNA levels correlated well with the strength of the phenotype, with the exception of line #25, which had only a mild reduction of *KNU* mRNA but produced approximately 6% of *knu*-like siliques (Figure 4.3A).

Given that the sample size in this preliminary experiment was small, the line (#28) that showed the strongest reduction of *KNU* mRNA and the highest penetrance of the *knu* phenotype (Figure 4.3A) was propagated and homozygous lines were isolated in order to explore further the relationship between *KNU* mRNA levels and the appearance of a *knu*-like phenotype. RNA was extracted from the primary inflorescences of 10 individual plants and then pooled. Line #2 was also included in the analysis as a control since the levels of *KNU* mRNA in this line were reduced to only ~50% in the first generation (not shown). Once cDNA had been generated from the RNA preparations of these lines, qRT-PCR was performed using primers specific to the *KNU* coding region. In addition, the numbers of

Figure 4.2 continued. *amiRNA* plants were shorter than wild-type siliques. **(C)** Ectopic reproductive organs were observed within bulged siliques. Ectopic carpels produced ovules. One to six stamens were observed within these siliques. **(D)** Tri-carpelloid siliques were also observed though they were present at a lower rate than bulged siliques that contained ectopic organs. **(E – H)** Phenotypes observed in plants of line #6. Levels of the *KNU* mRNA were reduced to ~10% of wild-type levels in these flowers (see Figure 4.3). **(E)** Flower with a short tri-carpelloid gynoecium. **(F)** All siliques of this plant were tri-carpelloid. **(G)** Stamens formed within siliques although they did not appear as frequently as in line #28. **(H)** Ectopic carpels were always observed within these siliques. Images are not to scale. 3'OCS, octopine synthase terminator sequence; RB, right border of T-DNA; LB, left border of T-DNA. lacZ, a modified version of the β -galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA® resistance gene; NOS, nopaline synthase terminator sequence.

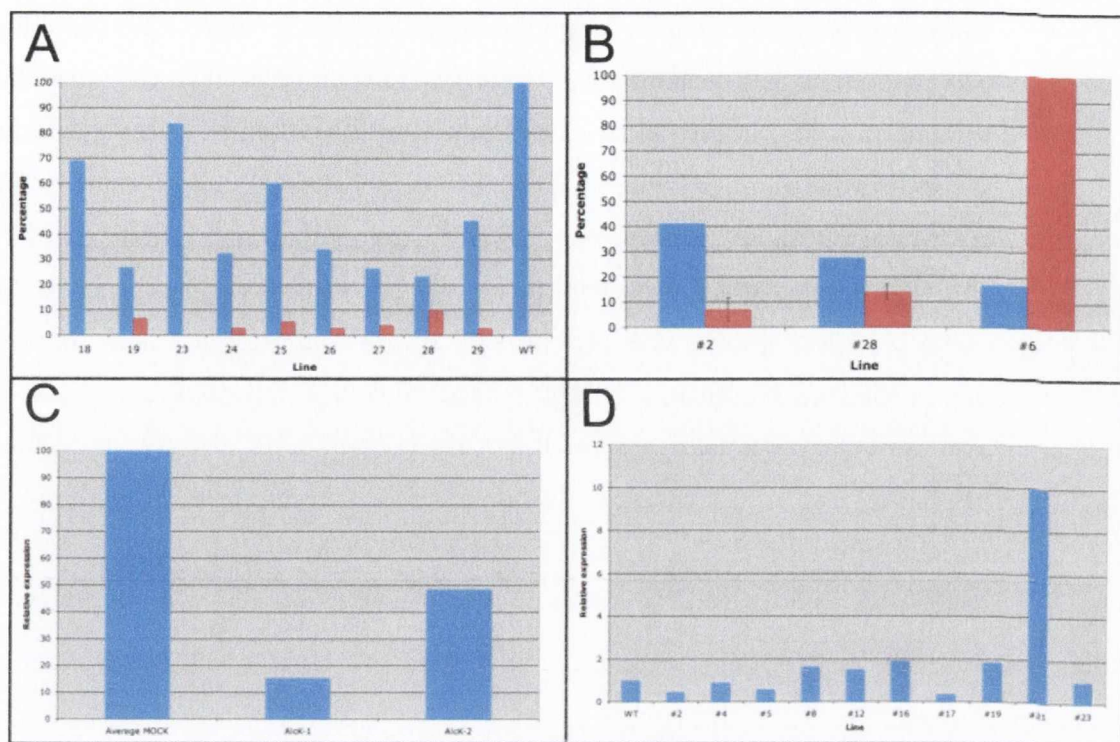


Figure 4.3 Accumulation of *KNU* transcripts in plants expressing the *KNU-amiRNA*. (A) Levels of *KNU* transcripts in T₁ plants constitutively expressing the *KNU-amiRNA* (blue bars) were measured by qRT-PCR and the percentage of bulged siliques relative to non-bulged siliques present on each plant (red bars; at least 15 siliques were counted per plant) was determined. (B) Levels of *KNU* transcripts in plants homozygous for the *35S::KNU-amiRNA* transgene (blue bars) and percentage of siliques that were bulged (red bars). Error bars are s.e.m. N = 9 (the siliques of 9 plants were counted for each line, at least 20 siliques per plant) (C) Levels of *KNU* transcripts in mock-treated (red bars) and ethanol-treated (blue bars) plants homozygous for the *35S::AlcR/AlcA::KNU-amiRNA* transgene. (D) Levels of *KNU* transcripts in F₁ plants that contained both the *35S::KNU-amiRNA* and the *pKNU::mKNU-GFP* transgene. No error bars are presented for expression assays because these values represent technical duplicates. C_p values of the duplicate measurements were with 0.5 units of each other. The primer pairs DM-94/DM-95, DM-94/DM-121 and DM-97/DM-98 were used to amplify the *KNU*, *KNU-GFP* and *REF2* (At4g34270) cDNAs, respectively.

bulged siliques in separate sibling populations grown in parallel were counted. *KNU* mRNA levels were reduced to ~25% and ~40% of wild-type levels in lines #28 and #2, respectively (Figure 4.3B). The mean percentage of bulged siliques relative to the total number of siliques was ~14% and ~7%, respectively. However, the degree of variation among plants of line #2 was comparatively high, as many plants did not produce any bulged siliques.

Another line (#6) containing the *35S::KNU-amiRNA* displayed a more severe phenotype than all of the lines discussed above (Figure 4.2E – H and Figure 4.3B). All

siliques produced by plants of this line were tricarpelloid (Figure 4.2E, F and Figure 4.3B) and in addition, they contained ectopic carpels. In contrast to other *KNU-amiRNA* lines, the ectopic carpels of line #6 developed in a more apical region of the gynoecium (Figure 4.2H, arrowhead). In addition, a subset of siliques contained ectopic stamens (Figure 4.2G, arrowhead). Because the phenotype of line #6 was unusual compared to that of other *KNU-amiRNA* lines or of the *knu-1* mutant, I investigated whether it was related to lower levels of *KNU* mRNA in these lines. To this end, I pooled inflorescence tissue from 8 T₂ plants that displayed the tricarpelloid phenotype, followed by RNA extraction and cDNA synthesis. qRT-PCR analysis showed that the levels of *KNU* mRNA were reduced to ~16% of wild-type levels (Figure 4.3B, blue bar) and thus, further than in any of the other *KNU-amiRNA* lines tested. Therefore, it appears that in line #6, the levels of the *KNU* mRNA were sufficiently reduced to ensure a full penetrance of a *knu*-like phenotype (see Section 4.5.1 for discussion).

4.2.2 Identification of a transgenic line that allows an inducible perturbation of KNU activity

The section above demonstrated that the expression of the *KNU-amiRNA* led to a considerable knockdown of *KNU* mRNA levels and to a disruption of its activity. To allow a specific perturbation of *KNU* function, the *KNU-amiRNA* was expressed under control of the 35S:AlcR/ pAlcA ethanol-inducible promoter system in wild-type plants. In total, 13 independent transformants were isolated and genotyped to verify the presence of the transgene (Figure 4.4A). The progeny of these plants was then treated with ethanol vapour for 48 hours and subsequently screened for the presence of *knu*-like phenotypes. A severe loss of apical dominance was observed in one line, while 10 other lines did not display any mutant phenotypes except mild and temporary

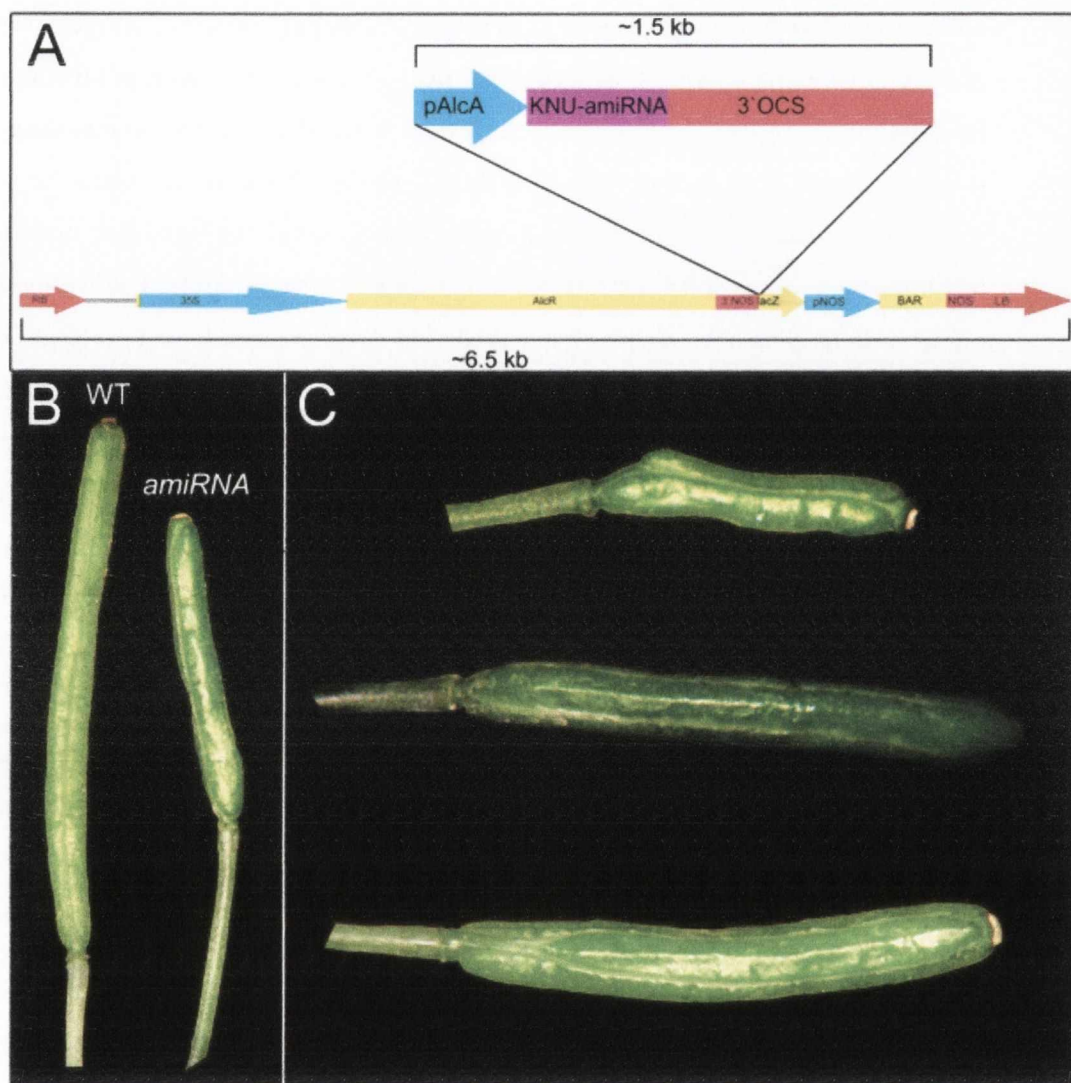


Figure 4.4 Phenotypes of 35S::AlcR /pAlcA::KNU-amiRNA plants. (A) A schematic of the T-DNA carrying the amiRNA under the control of the AlcA promoter. The binary vector pML-BART, which also carried a 35S::AlcR construct, was used in conjunction with *Agrobacterium tumefaciens* to transform wild-type plants. (B) Siliques of ethanol-treated plants were shorter than wild-type siliques. (C) “Knuckles” sometimes appeared on siliques (top of image), however, no ectopic organs were observed within these structures. Tri-carpelloid gynoecia (centre and bottom of image) were the most common mutant phenotypes observed. 3'OCS, octopine synthase terminator sequence; RB, right border of T-DNA; LB, left border of T-DNA. lacZ, a modified version of the β -galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA® resistance gene; NOS, nopaline synthase terminator sequence.

sterility. The remaining two lines (#1 and #2) displayed mild *knu*-like phenotypes (Figure 4.4B, C). Their siliques were shorter than those of the wild type (Figure 4.4B) and “knuckles” were sometimes present, although ectopic reproductive organs were never

observed within these structures (Figure 4.4C). Instead, tri-carpeloid siliques were the most common phenotype observed in both lines. Because these lines displayed the strongest knu-like phenotype in the T1 generation, they were isolated in a homozygous state for further characterization (Figure 4.4C). Homozygous lines were either mock-treated or were treated with ethanol vapour for 24 hours and inflorescence tissue was harvested immediately after the treatment. After RNA extraction and cDNA synthesis, qRT-PCR was performed in order to determine the extent of the KNU-amiRNA-mediated KNU mRNA knockdown (Figure 4.3C). This analysis showed that KNU mRNA levels were reduced to ~16% and ~50% in lines #1 and #2, respectively, compared to the mock-treated plants. Despite these differences in the degree to which KNU mRNA levels were decrease in these two lines, the strength of the resulting phenotypes did not appear to be dramatically different (Figure 4.4B, C). In particular, no ectopic reproductive organs within siliques were observed. Although the reasons for a lack of strong knu-like phenotypes in these lines are currently unclear, it can be speculated that they are due to insufficient treatment with ethanol vapour. In support of this idea, it was observed that 35S::AlcR/pAlcA::AG4-amiRNA plants required up to 4 days of ethanol treatment in order generate a phenotype resembling that of an *ag-1* null mutant (see Section 3.2.2). Therefore, similar durations of treatment may be necessary for the inducible KNU-amiRNA lines to show a full knu phenotype. Although this point has not been fully addressed at this stage, both ethanol-inducible KNU-amiRNA lines have been crossed into the *pAPI-FIS* line. Homozygous lines for all appropriate transgenes have been identified for line #2, whereas F1 seeds are available for line #1.

4.2.3 Generation of *KNU* over-expressing plants

Over-expression of a transcriptional regulator is a common method of perturbing a gene regulatory network and can be used to identify genes involved in a particular process. Because no *KNU* constitutive over-expression phenotype had been described, I generated transgenic lines that over-expressed *KNU* ectopically under the control of the the constitutive 35S promoter, as well as conditionally under the control of the 35S::AlcR/pAlcA ethanol-inducible promoter system (see Section 2.13.10). A large number of plants were transformed with the vector carrying the 35S::*KNU* construct, however, only 3 primary transformants were isolated that contained the transgene. None of these plants exhibited a

mutant phenotype in the first or second generation. It remains unclear why so few transformants could be obtained for this construct (see Section 4.3.3 for discussion).

For the inducible over-expression of *KNU*, the 3' end of the *KNU* coding region was translationally fused to six copies of the haemagglutinin tag (denoted 6xHA) to generate a *35S::AlcR/pAlcA::KNU-6xHA* construct. After transformation into wild-type plants, 16 primary transformants were identified that contained the transgene. Their progeny was submitted to different ethanol vapour treatments. One set of plants was treated continuously for 48 hours, while another set was treated twice for 24 hours, with a 24 hour recovery period in between treatments. No phenotypes were observed with either treatment regime. Homozygous lines containing these transgenes are currently being isolated and qRT-PCR will be used to identify lines that strongly over-express the *KNU-6xHA* transgene in an ethanol responsive manner. I then intend to screen for over-expression phenotypes using these homozygous lines.

4.2.4 Translational fusion of an mGFP5 reporter gene to *KNU*

The identification of a functional *KNU-miRNA* prompted the generation of an amiRNA-resistant, epitope-tagged version of *KNU* that could rescue the mutant phenotype observed in lines expressing *KNU-miRNA*. This line could be used in ChIP experiments in order to identify the genomic regions bound by *KNU*. To this end, an *mGFP5* reporter gene was translationally fused to the 3' end of the *KNU* coding region (see Section 2.13.11) and was placed under the control of the endogenous *KNU* promoter. A T-DNA construct carrying this fusion as well as a kanamycin resistance gene were transformed into homozygous *35S::KNU-miRNA* plants (line #2; see above). Transformants were identified based on their resistance to kanamycin and by PCR-based genotyping (see Section 2.13.11). Of 25 independent transformants it that contained the *pKNU::mKNU-mGFP5* transgene (these plants are referred to as mKNU-GFP plants hereafter), 10 did not produce any bulged siliques, suggesting that in these lines the *KNU-mGFP5* fusion protein could compensate for the amiRNA-mediated knockdown of endogenous *KNU* activity. These ten plants were then crossed with the *pAPI-FIS* (see Chapter 5). In order to carry out a preliminary characterization of these 10 mKNU-GFP lines, inflorescence tissue from F1 plants (none produced *knu*-like siliques), as well as from wild-type plants, was collected. After RNA

extraction and cDNA synthesis, qRT-PCRs were performed to determine *KNU* transcript levels. *KNU* mRNA levels in these lines varied from 35% of wild-type levels to a 10-fold increase in *KNU* mRNA when compared to wild-type plants. Although *KNU* mRNA levels in certain lines (#2, #4, #5 and #17) were lower than in the wild type, no *knu* phenotypes were observed, suggesting that the remaining *KNU* activity was functionally sufficient. Lines #12, #16 and #23 were selected for further analysis because *KNU* mRNA levels appeared to be similar to those in wild-type plants. Homozygous plants for line #12 and #23 have been isolated in the *pAPI-FIS* and await further characterization.

4.3 Discussion

KNU is an important developmental regulator involved in floral meristem termination (Sun et al., 2009). In order to elucidate the exact role that *KNU* plays in flower development, several transgenic lines were generated that can be used to identify genes controlled by *KNU* activity. The value of the lines obtained and the work remaining to validate these lines is discussed below.

4.3.1 Identification of a functional *KNU*-amiRNA

Over-expression of the *KNU*-amiRNA revealed that this particular amiRNA could perturb *KNU* activity efficiently. Phenotypes consistent with *knu-1* phenotypes were observed, however, the penetrance of these phenotypes was lower than what has been described for the *knu-1* allele (Payne et al., 2004). This result could be a consequence of the different accessions that provide the background for the known *knu* alleles and the *KNU*-amiRNA lines. In order to test this possibility, the *knu-1* (Payne et al., 2004) or *knu-2* (Sun et al., 2009) alleles could be crossed into the *L-er* accession (which was used for generating the amiRNA lines) and the phenotypes could be scored. Alternatively, the *KNU*-amiRNA could be transformed into Col-0 or *Ws* plants, *i.e.* the accessions in which the known *knu* alleles have been isolated.

Despite the relatively low penetrance of the *KNU*-amiRNA, I obtained lines that produced consistently *knu*-like phenotypes and exhibited a severe reduction of *KNU*

transcript levels (Figure 4.3A, B). These lines have now been crossed into the *pAPI-FIS*, and will allow experiments with the aim of identifying genes that are under *KNU* regulation.

As described in section 4.3.1, one line (#6) expressing the *KNU-amiRNA* showed an unexpected phenotype (Figure 4.2E – H), in that all siliques formed were tri-carpelloid and had ectopic organs. Notably, this line showed the strongest reduction of *KNU* mRNA levels from all lines tested, suggesting that the appearance of the tri-carpelloid siliques with ectopic organs could be related to a strong down-regulation of *KNU* activity. In contrast, only 11% of *knu-1* plants produced tri-carpelloid siliques according to Payne *et al.* (2004). The *knu-1* allele was generated by ethyl methanesulfonate (EMS) mutagenesis and the *knu-1* coding region contains a non-synonymous substitution that results in a cysteine residue being replaced by a tyrosine residue. This amino acid substitution affects one of the two cysteine residues required for zinc binding (Payne *et al.*, 2004) and thus, *knu-1* is likely a functional null allele. A second *knu* allele (*knu-2*), which has been described by Sun *et al.* (2009), produces “knuckle” siliques with ectopic organs forming within these structures. The *knu-2* allele contains a T-DNA insertion towards the 3' end of the coding region, but the exact location has not been described and it is currently unclear whether this allele is a hypomorph or an amorph.

In order to further understand the tri-carpelloid phenotype observed in the *KNU-amiRNA* line #6, future work will aim at the identification of additional lines that show a similar phenotype. This would rule out the possibility that mutant phenotypes of line #6 are a consequence of the insertion site of the T-DNA. In addition, line #6 will be crossed with the *mKNU-GFP* plants described above. If the phenotypes observed in plants of line #6 were indeed a result of *KNU-amiRNA* over-expression, then the F₁ progeny should display a wild-type phenotype.

The induction of *KNU-amiRNA* expression from the *35S::AlcR/pAlcA* ethanol-inducible promoter system resulted in a decrease of *KNU* transcript levels to as little as 16% of wild-type levels (Figure 4.3C). Tri-carpelloid siliques and shorter siliques were often present after ethanol treatment (Figure 4.4A, B), but were never observed in mock-treated plants. These data suggest that the *pAlcA::KNU-amiRNA* lines can be used to identify genes that are responsive to *KNU* activity. Several potentially suitable lines have already been crossed into the *pAPI-FIS* and F₁ seeds are available. In the F₂ generation, plants that are homozygous for all mutations and transgenes will be isolated. The levels of *KNU* mRNA will then be monitored upon treatment with ethanol vapour in order to validate the response

of the resulting lines and their suitability for knockdown experiments. Additionally, transcriptome analysis can also be performed using the inducible *KNU*-amiRNA lines in a wild-type background. This should lead to the identification of genes predominantly responsive to *KNU* activity at later stages of flower development.

4.3.2 Functionality of mKNU-GFP fusion protein

A *KNU*-amiRNA resistant, epitope tagged version of *KNU* was constructed to allow the identification of the genomic regions to which the *KNU* transcription factor binds. To this end, endogenous *KNU* activity will be knocked down using the *KNU*-amiRNA and the amiRNA-resistant version of *KNU* will then be used to complement for this loss of function. The addition of an epitope tag will ensure that commercially available antibodies that are of ChIP-quality can be employed for the global identification of binding sites. The mKNU-GFP line can be further used as a control for gene expression profiling experiment conducted with the inducible *KNU*-amiRNA lines. As described in Section 3.2.5, the *35S::AlcR/pAlcA* ethanol-inducible promoter system appears to lead to non-specific activation of gene expression. Therefore, appropriate control experiments have to be performed to distinguish true response genes from genes that are affected as a consequence of AlcR activation. For the inducible *KNU*-amiRNA lines such controls could be done by testing the effects that an induction of *KNU*-amiRNA expression has in the presence or absence of the *pKNU::mKNU-mGFP5* transgene. Genes that act downstream of *KNU* should only respond in the absence of the amiRNA-resistant *KNU* variant, while genes that respond under both conditions are likely influenced non-specifically by the inducible promoter system.

In order to test whether the mKNU-mGFP5 fusion protein can functionally replace endogenous *KNU*, I transformed one of the *35S::KNU-amiRNA*-containing plants with the *pKNU::mKNU-mGFP5* transgene. Several of the primary transformants obtained showed a wild-type phenotype and the further characterisation of these lines revealed that *KNU* mRNA levels were similar or above wild-type levels in the majority of them (Figure 4.3D). These results suggest that the mKNU-mGFP5 fusion protein is indeed functional and could be used for the experiments outlined above.

4.3.3 Inducible or constitutive over-expression of *KNU* did not result in novel phenotypes

In order to understand in more detail the role of *KNU* in flower development, plants were transformed with a *35S::KNU* transgene to allow a constitutive over-expression of the gene. Despite the fact that a large number of plants were transformed, only 3 independent transformants were identified and none of them produced an abnormal phenotype (not shown). Given the number of plants used for transformation of this T-DNA, many more transformants should normally have been obtained. One possibility to explain the low number of transgenic plants identified is that early developmental lethality is associated with expression of *KNU* from the 35S promoter. Indeed, *KNU* has been proposed to down-regulate *WUS* (Sun et al., 2009), which is central to meristem maintenance. Therefore over-expression of *KNU* could perturb *WUS* activity and result in embryo lethality or arrest.

Plants harbouring the *35S::AlcR/pAlcA::KNU-6xHA* transgene did not display any phenotypes after ethanol treatment (not shown) in contrast to a similar 5 α -androstan-17 β -ol-3-one (DHT) responsive over-expression line described by Sun *et al.*, (2009). The latter plants produced a *wus*-like phenotype when repeatedly treated with DHT. In addition, this transgene was able to retard meristem growth in *ag-1* mutant plants (Sun et al., 2009). These phenotypes were not observed here for unknown reasons. Further work will aim at validating the *pAlcA::KNU-6xHA* lines and, in particular, at identifying lines that exhibit an over-expression of the *KNU-6xHA* fusion protein.

4.4 Future Work

The lines described above were generated as part of a project to supplement the main goal of my thesis, *i.e.* to identify the genes and processes regulated by AG. The Wellmer group has recently been granted funding by Science Foundation Ireland to investigate the role of *KNU* in flower development in detail. Therefore, the lines generated in this work can be validated and used over the coming years. The experiments that can be performed using these lines (and their derivatives) include: (1) the identification of genes that are dependent upon *KNU* activity for expression (using the *35S::KNU-amiRNA* lines),

(2) the identification of genes responsive to KNU activity (using the *pAlcA::KNU-amiRNA* lines and potentially the inducible over-expression lines) and (3) the identification of genomic regions bound by the KNU protein (using the mKNU-GFP lines). These experiments can be performed in a wild-type background and in the FIS to investigate the role of KNU in early and late flower development.

Chapter 5 - Generation of alternative floral induction systems

5.1 Introduction

The *35S::API-GR ap1 cal* floral induction system described by Wellmer et al. (2006), which is based on the inducible activation of the AP1 transcription factor (see Section 1.4), presents the opportunity to collect large amounts of synchronized flower buds at early developmental stages. This greatly facilitates the study of the molecular processes that control flower development by genomic technologies. Although this system has already provided a large amount of interesting data, its versatility can be improved. For example, the generation of plants that carry another transgene (as described, for example, in Section 3.2.3.3) in the background of the floral induction system is onerous because transformation efficiencies for the *35S::API-GR ap1 cal* line are low. This problem is further exacerbated by the fact that the *35S::API-GR* transgene carries a Basta® herbicide resistance cassette. This cassette constitutes the most powerful selection marker currently available for the isolation of transgenic *Arabidopsis* plants as it allows the rapid screening of thousands of candidate plants. Therefore, secondary transgenes that are to be introduced into the floral induction system must carry a different selection marker. This selection marker is typically an antibiotic resistance gene, which makes the identification of primary transformants more difficult.

An additional, potentially more serious problem of the original version of the floral induction system is that it is based on the over-expression of an AP1-GR fusion protein. This over-expression might cause unwanted effects, and preliminary evidence suggested that the system is somewhat “leaky”, *i.e.* that it leads to a precocious induction of flower development when compared to *ap1 cal* control plants. Driving *API-GR* expression under the control of the *API* promoter might circumvent this problem since *API-GR* mRNA expression would be restricted to its normal expression domain and the transcript levels should be closer to that of the endogenous *API* gene. To address the potential problems outlined above, several alternative versions of the floral induction system were generated and characterized.

In a second line of experiments, the *CAULIFLOWER* (*CAL*) gene was translationally fused to a portion of the GR receptor and expressed in *ap1 cal* double-mutant plants. As described in Section 1.4, plants harbouring an *ap1-1* mutation produce flowers that lack sepals and petals (Bowman et al., 1993). In addition, floral meristems are partially converted into inflorescence meristems. In contrast, plants harbouring a *cal-1* mutation are almost indistinguishable from the wild type (Bowman et al., 1993). Double mutant *ap1-1 cal-1* plants do not flower for an extended period of time and instead undergo a massive over-proliferation of inflorescence-like meristems, indicating that *API* and *CAL* control the onset of flower development in a redundant manner (Bowman et al., 1993). However, while *CAL* can functionally replace *API* during the earliest stages of flower formation, it cannot compensate for a loss of *API* function at later developmental stages as indicated by the severe defects of *ap1* mutant flowers (Bowman et al., 1993). It thus seems that *API* and *CAL* act redundantly only during the onset of flower development.

The underlying mechanism for the functional differences between *API* and *CAL* remains unclear. The expression patterns of *API* and *CAL* do not overlap completely (Kempin et al., 1995), suggesting that functional differences could be due to different expression domains. However, the results of mis-expression of *CAL* in an *ap1-1* mutant background suggests that the functional differences between *API* and *CAL* are due to differences at the protein level (Alvarez-Buylla et al., 2006). In order to understand the functional differences between *API* and *CAL* in detail, we aim to identify the genes that are regulated by *CAL*. This work is done in collaboration with the laboratory of José Luis Riechmann in Barcelona. For this work, we intend to use an experimental approach, which relies on the expression of a *CAL-GR* fusion in an *ap1-1 cal-1* mutant background. The phenotypes of the *CAL*-based floral induction systems that could be used for expression profiling and ChIP experiments are presented below.

5.2 Results

5.2.1 Generation of a floral induction system with an alternative selection marker

A construct carrying *35S::API-GR* (Figure 5.1A) and a kanamycin-resistance gene (neomycin phosphotransferase) was transformed into *apl-1 cal-1* plants and transformants were identified by selecting for resistance to kanamycin (see Section 2.13.12). Of 13 independent transformants obtained, eight responded to dexamethasone treatment in a similar manner as the herbicide-resistant version described by Wellmer *et al.* (2006) (Figure 5.1B – H). These responses included synchronous induction of flowering (Figure 5.1G) and restoration of petal and sepal development (Figure 5.1I), which was not observed in mock-treated plants (Figure 5.1F, H). In summary, the kanamycin-resistant floral induction system behaved identically to the herbicide-resistant version.

Based on the results of a segregation analysis, seeds from T1 lines that appeared to contain only one T-DNA were selected. Among these lines, those that responded well to dexamethasone were kept to isolate plants that were homozygous for the *35S::API-GR* T-DNA. One homozygous line was selected as a reference and was propagated for use in further experiments.

5.2.2 Generation of an improved floral induction system driving *AP1-GR* under the control of the endogenous *AP1* promoter

In order to generate an improved floral induction system that behaved more similarly to wild-type plants than the original *35S* promoter-based floral induction system, the entire genomic locus of *AP1* was cloned and the last exon of *AP1* was fused to the GR coding sequence as described in Section 2.13.13 and depicted in Figure 5.2A. T-DNA constructs carrying either Basta® herbicide-resistance or kanamycin-resistance genes as selection markers were separately transformed into *apl-1 cal-1* plants as described in Section 2.3.

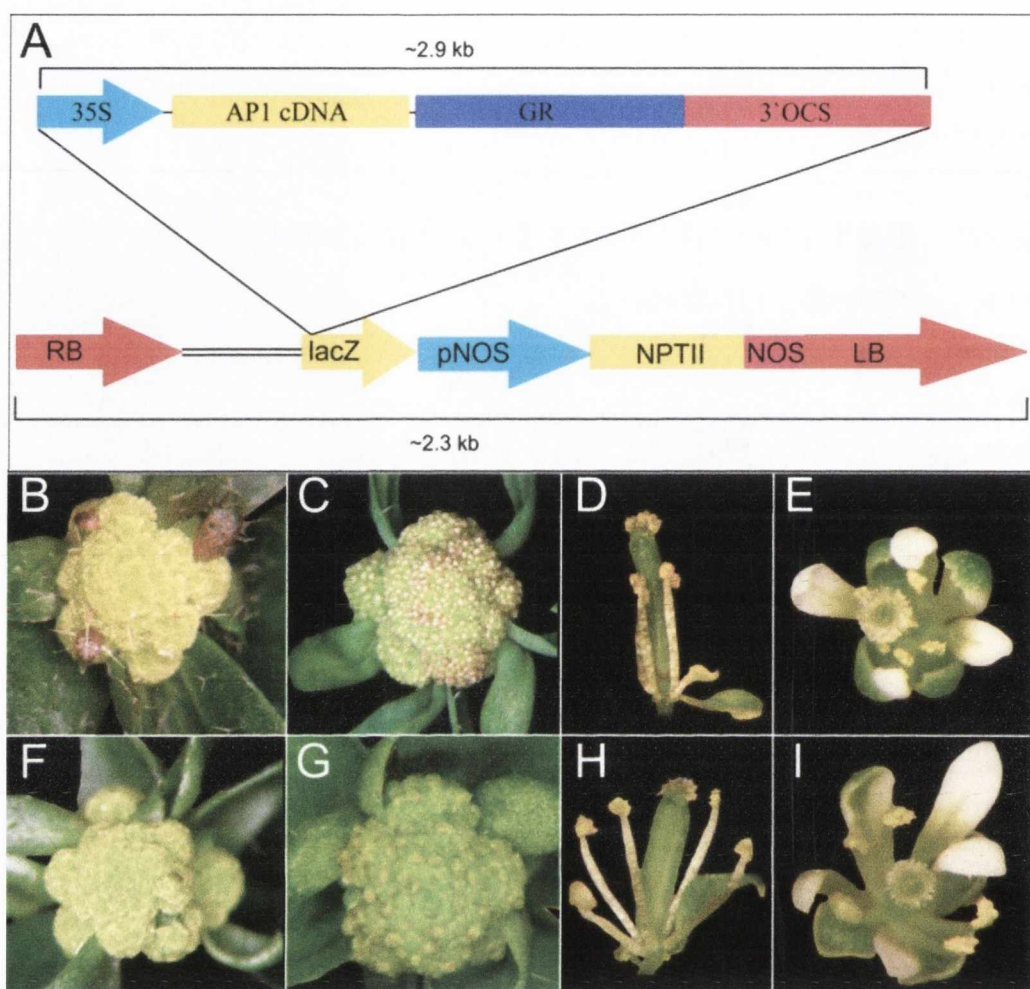


Figure 5.1 Phenotypes of floral induction systems based on AP1-GR over-expression. (A) A schematic of the T-DNA used to drive expression of the AP1-GR fusion protein. The binary vectors pGREEN (B – E) and pART27 (F – I) were used in conjunction with *Agrobacterium tumefaciens* for plant transformation. For full details, see Section 2.13.12. (B – E) BASTA herbicide-resistant plants harbouring the 35S::AP1-GR transgene described by Wellmer *et al.* (2006). (B) An inflorescence-like meristem 5 days after treatment with a mock solution (see Section 2.5.2 for details). (C) An inflorescence-like meristem 5 days after treatment with a dexamethasone-containing solution. (D) A flower produced by a mock-treated inflorescence-like meristem. (E) A flower produced from a dexamethasone-treated inflorescence-like meristem. (F–I) Kanamycin-resistant plants harbouring the 35S::AP1-GR transgene generated in this study. (F) An inflorescence-like meristem 5 days after being treated with a mock solution. (G) An inflorescence-like meristem 5 days after being treated with a dexamethasone-containing solution. (H) A flower produced by a mock-treated inflorescence-like meristem. (I) A flower produced by a dexamethasone-treated inflorescence-like meristem. 3'OCS, octopine synthase terminator sequence; RB, right border of T-DNA; LB, left border of T-DNA. lacZ, a modified version of the β -galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; NPTII, neomycin phosphotransferase (kanamycin resistance gene); NOS, nopaline synthase terminator sequence; GR, glucocorticoid receptor.

Eight herbicide-resistant and seven antibiotic-resistant transformants were isolated and treated with a dexamethasone-containing solution. The response to dexamethasone in these lines was similar to that of the *35S::API-GR* lines described above. The formation of synchronously developing flower buds was observed (Figure 5.2C, G). In addition, all types of floral organs were present (Figure 5.2E, I), suggesting that *API-GR* under the control of the *API* promoter mediates the meristematic and homeotic functions of the endogenous AP1 protein (Figure 5.2B – H). Mock-treated plants did not display synchronous induction of flower development (Figure 5.2B, D), nor did the inflorescence-like meristems produce flowers that contained petals or sepals (Figure 5.2F, H).

The development of flowers from *pAPI-FIS* plants was different to those of *35S-FIS* plants in some aspects. These differences included the morphology of floral organs and the transitioning of inflorescence-like meristems to flowering. First, the floral organs produced after dexamethasone treatment of inflorescence-like meristems were compared. Plants were treated with 10 μ M dexamethasone once the stems had bolted to a height of $\sim$ 1 cm and flowers produced were phenotyped as described in Section 2.5.2. No difference was observed between sepals (Figure 5.3F, I) and carpels (data not shown) of the two floral induction systems, however, petals and the anthers of stamens developed differently (Figure 5.3D, E, G, H). Petals of plants carrying the *pAPI::API-GR* transgene appeared to be broader and longer than those of plants expressing *API-GR* from the *35S* promoter (Figure 5.3E, H). However, the petals of the *pAPI-FIS* line were on average not as large as those of wild-type plants (Figure 5.3B). Although stamens appeared to develop correctly in both floral induction systems, the anthers of *35S-FIS* stamens produced less pollen than those of *pAPI-FIS* stamens (Figure 5.3D, G). The amount of pollen produced by the anthers of *pAPI-FIS* stamens appeared comparable to the wild type (Figure 5.3A).

Second, the transition time of the inflorescence-like meristems to flowering was quantified. It has been described previously that the inflorescence-like meristems of *ap1 cal* plants can transition to flowering in an *API*-independent manner (Bowman et al., 1993). I observed that the inflorescence-like meristems of *35S-FIS* plants appeared to transition earlier than those of *ap1 cal* plants in the absence of dexamethasone treatment. To determine if the *pAPI-FIS* line transitioned to flowering similarly to an *ap1 cal* mutant, the transitioning behaviour of *pAPI-FIS* inflorescence-like meristems was analysed in the absence of dexamethasone. To this end, plants were phenotyped at least every second day. In order to account for any other developmental differences between genotypes, the time of

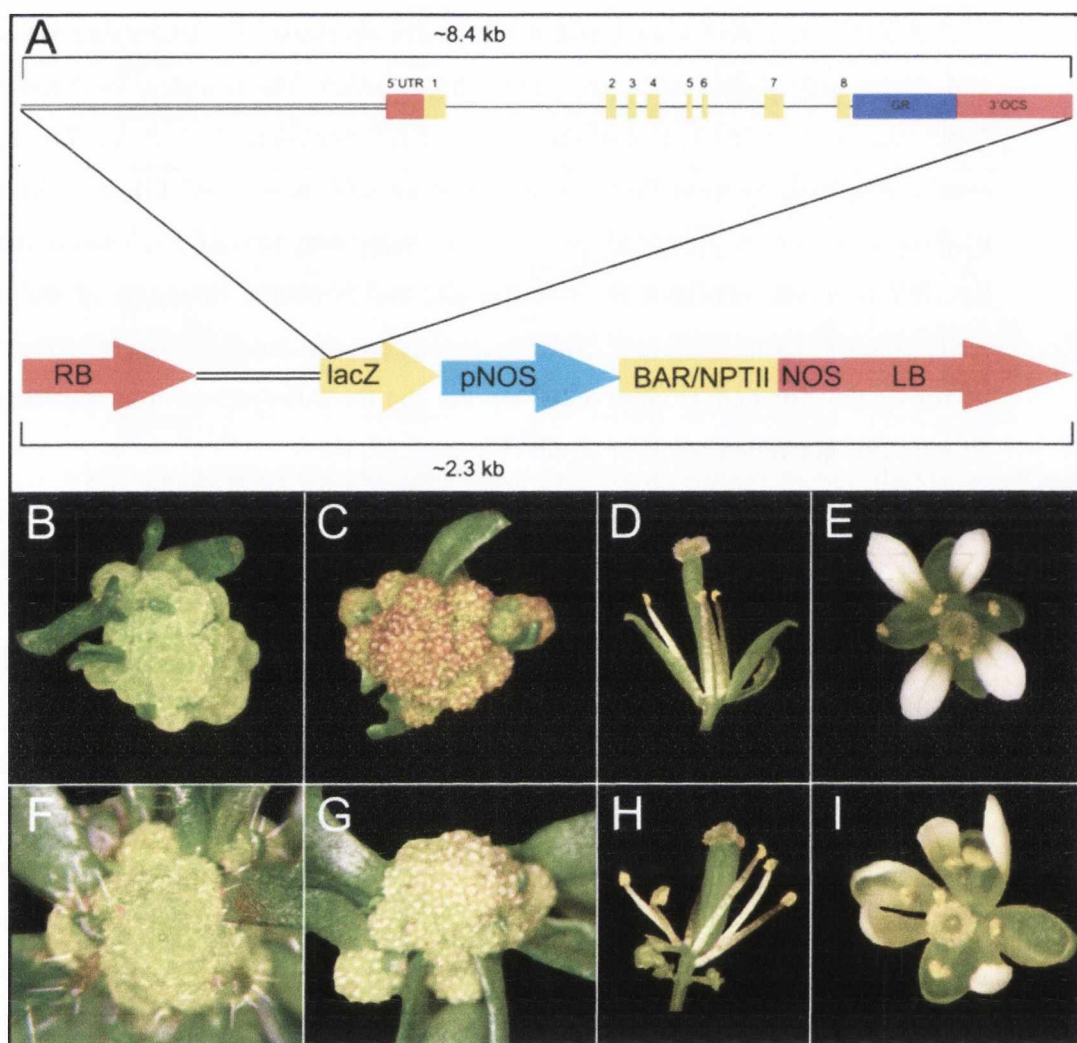


Figure 5.2 Phenotypes of floral induction systems based on expression of an *API*-GR fusion protein from the *API* promoter. (A) A schematic of the T-DNA insertion carrying the *pAPI::API-GR* construct, which was designed to drive expression of *API-GR* in an wild-type pattern. The binary vectors pML-BART (B – E) and pART27 (F – I) were used in conjunction with *Agrobacterium tumefaciens* for plant transformations. For full details see Section 2.13.13. Exons contained within the *API* locus are represented by yellow bars and are numbered. (B-E) The phenotypes of BASTA herbicide-resistant plants harbouring the *pAPI::API-GR* transgene. (B) An inflorescence-like meristem 5 days after being treated with a mock solution (see Section 2.7.2 for details). (C) An inflorescence-like meristem 5 days after being treated with a dexamethasone-containing solution. (D) A flower produced by a mock-treated inflorescence-like meristem. (E) A flower produced from a dexamethasone-treated inflorescence-like meristem. (F-I) Kanamycin-resistant plants harbouring the *pAPI::API-GR* transgene. (F) An inflorescence-like meristem 5 days after being treated with a mock solution. (G) An inflorescence-like meristem 5 days after being treated with a dexamethasone-containing solution. (H) A flower produced by a mock-treated inflorescence-like meristem. (I) A flower produced from a dexamethasone-treated inflorescence-like meristem. 3'OCS, octopine synthase terminator sequence; RB, right border of T-DNA; LB, left border of T-DNA. lacZ, a modified version of the β -

transitioning was determined relative to the time-point when the stems of plants had bolted to ~1 cm. In addition, the germination rates of all genotypes were determined and no discernable differences were observed (data not shown). The time of the onset of flower development was estimated by viewing the inflorescence-like meristems under a light microscope in order to detect the formation of flowers as early as possible.

Figure 5.4A shows that the *35S-FIS* plants transitioned to flowering earlier than either *pAPI-FIS* ($p < 0.001$, two-sided Student's T-test, equal variance assumed) or *ap1 cal* ($p < 0.001$) plants. Moreover, in comparison to the other two genotypes, the vegetative development of *35S-FIS* plants (*i.e.* the time from planting to bolting) was significantly shorter (*pAPI-FIS*; $p < 0.001$, *ap1 cal*; $p < 0.001$) (Figure 5.4B). The accelerated growth and floral transitioning of *35S-FIS* plants resulted in shorter total life cycles than both the *pAPI-FIS* ($p < 0.001$) and *ap1 cal* ($p < 0.001$) plants (Figure 5.4C).

Taken together the data of the experiments outlined above show that the *pAPI-FIS* plants transitioned to flowering significantly later than the *35S-FIS*, suggesting a reduced leakiness of the system. Furthermore, the flowers formed by *pAPI-FIS* plants resembled more close those of the wild type than the ones formed by the *35S-FIS*.

Finally, the localization of the *API-GR* mRNA was investigated in *pAPI::API-GR* plants (Figure 5.5) to ensure a correct spatial and temporal expression of the transgene. Briefly, the *pAPI-FIS* was crossed with wild-type plants and inflorescence tissue from the F_1 generation was prepared for *in situ* hybridisations as described in Section 2.11. Next, a digoxigenin-labeled antisense probe complementary to a portion of the *GR* coding sequence was generated. This probe was hybridised to tissue sections. Hybridisation signals were observed throughout young flower buds (Figure 5.5A), while expression was confined to sepals and petal primordia (Figure 5.5B) at more intermediate stages of flower development. This expression pattern resembles that for endogenous *API* (Gustafson-Brown et al., 1994), suggesting that the *pAPI::API-GR* construct contained all regulatory sequences required for proper expression of the gene.

Figure 5.2 continued. galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA® resistance gene; NPTII, neomycin phosphotransferase (kanamycin resistance gene); NOS, nopaline synthase terminator sequence; UTR, untranslated region; GR, glucocorticoid receptor.



Figure 5.3 – Morphology of floral organs from wild type, *35S-FIS* and *pAPI-FIS* flowers. (A – C) Floral organs from wild-type plants. (A) A stamen. (B) A petal. (C) A sepal. (D – I) Floral organs produced by dexamethasone-treated plants of the two different floral induction systems. (D – F) *35S-FIS* floral organs. (D) An example of the range of pollen production from the anthers of *35S-FIS* stamens. Stamens shown from left to right produced decreasing amounts of pollen. (E) Petal. (F) Sepal. (G – I) *pAPI-FIS* floral organs. (G) Stamen. (H) Petal. (I) Sepal. All organs are representative examples. Scale bar: 1 mm.

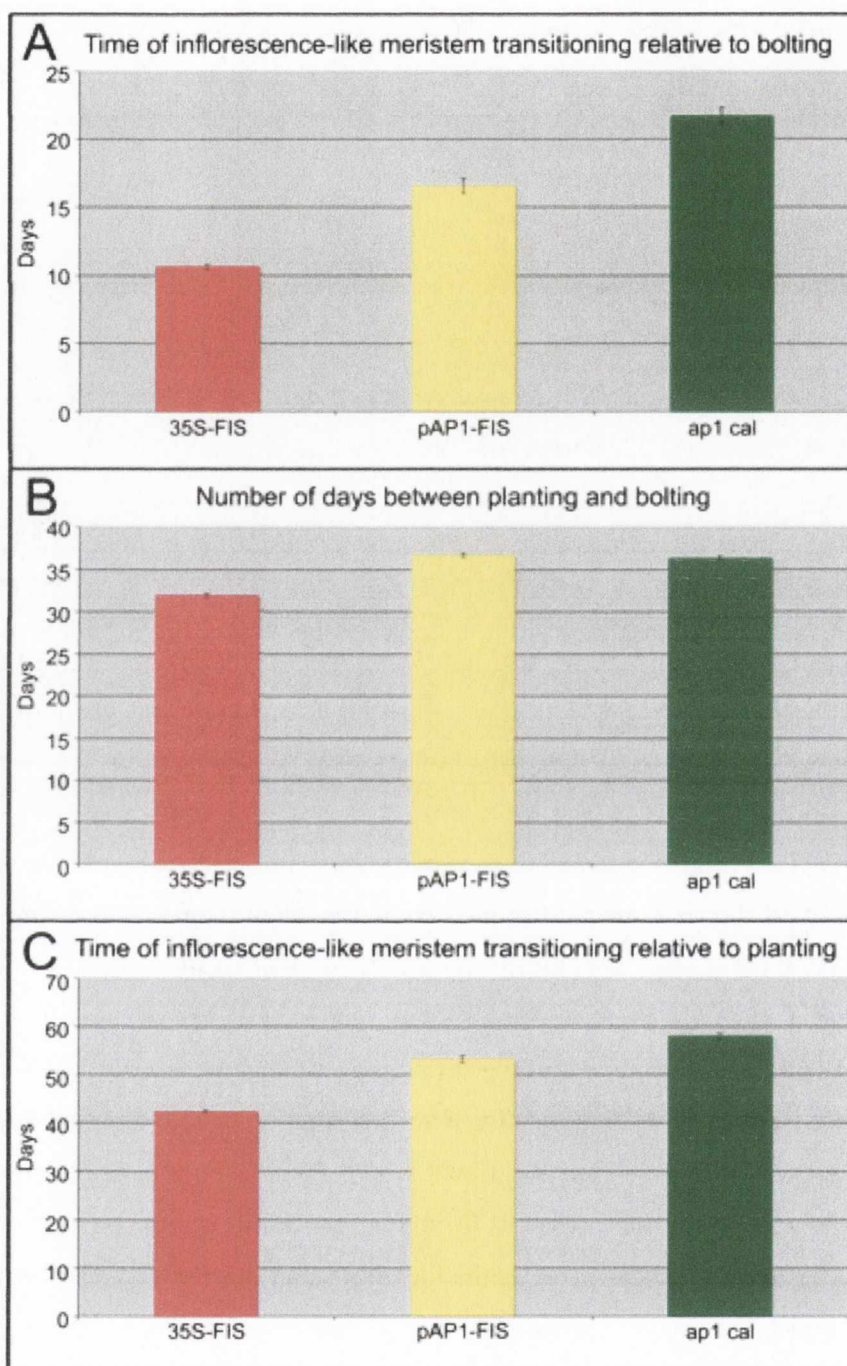


Figure 5.4 Transitioning time to flowering in the different floral induction systems. (A) Number of days between bolting of *35S-FIS* (red), *pAP1-FIS* (yellow) and *ap1 cal* (green) plants and the onset of flower formation is shown. (B) The number of days between planting of *35S-FIS*, *pAP1-FIS* and *ap1 cal* seeds and the onset of bolting is shown. (C) The number of days between planting of *35S-FIS*, *pAP1-FIS* and *ap1 cal* seeds and the onset of flower formation is indicated. The germination rate of all genotypes was similar (data not shown). Experiments were repeated three times with similar results. The results of one representative experiment are shown. Error bars indicate the standard error of the means. N = 24.

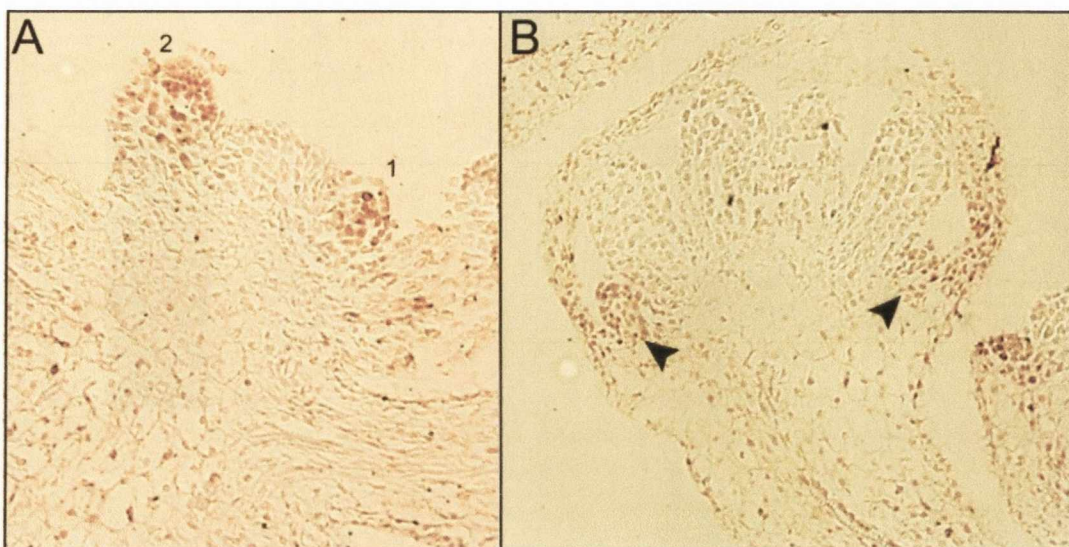


Figure 5.5 Pattern of *API-GR* mRNA expression in flowers of *pAPI::API-GR* plants. *In situ* hybridizations were done using longitudinal sections of *pAPI::API-GR* inflorescences and an anti-sense probe against the GR portion of *API-GR* fusion. **(A)** Expression of *API-GR* was detected throughout stage 1 and 2 floral buds. **(B)** Expression of *API-GR* was detected in petal primordia (arrowheads) and sepals in a stage 7 flower. Stages of flower development according to Smyth *et al.* (1990).

Two reference lines (designated pAPI-3 and pAPI-5) were chosen for the *pAPI-FIS* based on the response to dexamethasone treatment. Homozygous lines were identified for both of these lines and the positions of the T-DNA insertions within the genome were located by TAIL-PCR as described in Section 2.8. The primers used for genotyping are available in Appendix 1 and the genotyping strategy is described in Section 2.13.16. These lines were used in combination with the ethanol-inducible amiRNA and *AG-GFP* lines, respectively (see Section 3.2.5 and 3.2.7). Finally, one homozygous kanamycin-resistant *pAPI-FIS* line was isolated to allow super-transformation with T-DNA constructs conferring resistance to the herbicide Basta (see Section 2.13.16 above). Despite several attempts, the genomic position of the T-DNA insertion in this line could not be located by TAIL-PCR.

5.3 Generation of floral induction systems based on inducible activation of *CAL*

Two different dexamethasone-inducible versions of *CAL* were generated in order to identify the genes that are regulated by this transcription factor. First, the coding sequence of *CAL* was translationally fused to a portion of the GR receptor and placed under the control of the CaMV 35S promoter (Figure 5.6A, Section 2.13.14). Second, the entire genomic locus of *CAL* was cloned and the last exon was translationally fused to a portion of the GR receptor. The latter should allow expression of *CAL* under the control of its own promoter. I generated and sequenced both of these constructs, while Dr. Jiao Jin (Riechmann group, Catalan Institution for Research and Advanced Studies, Barcelona) used them to transform *ap1 cal* mutant plants. Plants containing the *35S::CAL-GR* transgene responded to dexamethasone in a manner similar to *35S-FIS* and *pAPI-FIS* plants. Synchronous flowering was observed 5 days after the inflorescence-like meristems were treated with a dexamethasone-containing solution (Figure 5.6B, C). In addition, the flowers that arose contained sepals and petals (Figure 5.6D). In summary, the meristematic and homeotic defects of *ap1 cal* plants were restored upon activation of the ectopically expressed *CAL-GR* fusion protein.

The inflorescence-like meristems of *ap1 cal* plants harbouring the *pCAL::CAL-GR* transgene formed synchronized flowers when treated with a dexamethasone-containing solution (Figure 5.7B, C). However, these flowers did not contain any petals or sepals (Figure 5.7D) unlike flowers of treated *35S::CAL-GR* plants. This phenotype was closer to the presumed role of *CAL* in flower development, *i.e.* that it is involved in controlling the onset of flower formation together with *API*, but does not take part in floral organ formation. Therefore, it appears that only the *pCAL::CAL-GR* transgenic plants are suitable for use in expression profiling experiments to identify genes that are regulated by *CAL* (see Section 5.3.2).

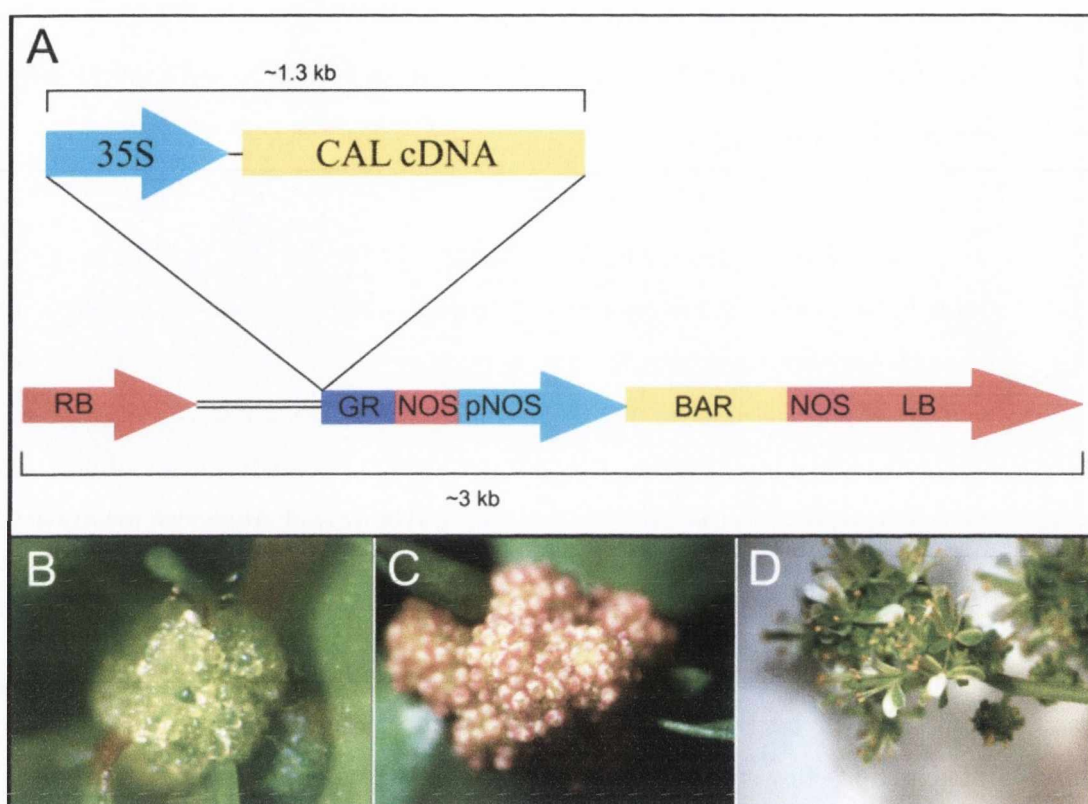


Figure 5.6 - Phenotypes of 35S::CAL-GR *apl cal* plants. (A) A schematic of the T-DNA used to express the CAL-GR fusion protein under the control of the *CaMV* 35S promoter. The binary vector pGREEN was used in conjunction with *Agrobacterium tumefaciens* for plant transformation. For full details, see Section 2.13.14. (B) An inflorescence-like meristem before treatment with a dexamethasone-containing solution. (C) An inflorescence-like meristem 5 days after treatment with a dexamethasone-containing solution. (D) An inflorescence-like meristem 13 days after treatment with a dexamethasone-containing solution. Images shown in panels (B – D) were kindly provided by Dr. Jiao Jin (Riechmann Group, Catalan Institution for Research and Advanced Studies, Barcelona). RB, right border of T-DNA; LB, left border of T-DNA. lacZ, a modified version of the β -galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA® resistance gene; NOS, nopaline synthase terminator sequence; GR, glucocorticoid receptor.

5.4 Discussion

5.4.1 Floral induction systems based on the activation of the AP1-GR protein

The initial FIS described by Wellmer *et al.* (2006) carried a transgene that conferred resistance to the herbicide BASTA (Wellmer *et al.*, 2006). The introduction of additional transgenes therefore required the use of an antibiotic selection marker, which is more difficult to select for in *Arabidopsis*. Hence, a FIS that confers resistance to the antibiotic kanamycin was generated. When the inflorescence-like meristems were treated with a dexamethasone-containing solution, this kanamycin-resistant line produced phenotypes identical to those of the original FIS. These plants can now be super-transformed with transgenes that confer resistance to the herbicide BASTA®, which allows a faster and a more economical identification of doubly transgenic plants. This line may also help to overcome problems related to the relatively low transformation efficiencies observed with the FIS, because it dramatically facilitates the screening for transgenic plants. A kanamycin-resistant version of the *pAP1-FIS* has been generated with similar reasoning.

The over-expression of the AP1-GR fusion protein creates the potential for interactions of the AP1 protein with proteins or sites in the genome with which the endogenous AP1 would normally not interact. In addition, the translocation of proteins fused to the GR receptor into the nucleus before steroid-hormone treatment has been described (Craft *et al.*, 2005; Lloyd *et al.*, 1994). Over-expression of the AP1-GR fusion may exacerbate this problem, as suggested by the precocious formation of flowers in *35S::AP1-GR ap1 cal* plants when compared to *ap1 cal* double mutants (Figure 5.4). The use of the *AP1* promoter to drive *AP1* expression may reduce the likelihood of such unwanted effects. In fact, the late transitioning of untreated *pAP1-FIS* plants to flowering in comparison to

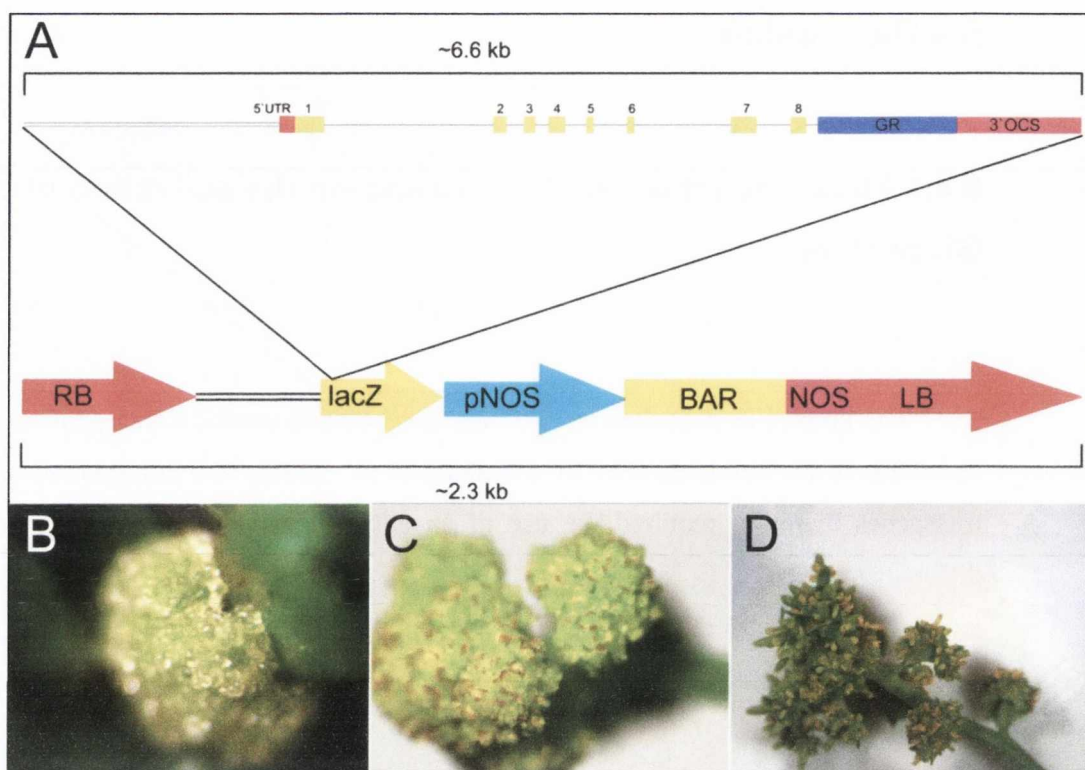


Figure 5.7 – Phenotypes of *pCAL::CAL-GR ap1 cal* plants. (A) A schematic of the T-DNA insertion carrying the *pCAL::CAL-GR* construct, which was designed to drive expression of *CAL* in a wild-type pattern. The binary vector pML-BART was used in conjunction with *Agrobacterium tumefaciens* to transform *ap1 cal* plants. For full details, see Section 2.13.15. (B) An inflorescence-like meristem before treatment with a dexamethasone-containing solution. (C) An inflorescence-like meristem 7 days after treatment with a dexamethasone-containing solution. (D) An inflorescence-like meristem 13 days after treatment with a dexamethasone-containing solution. Images shown in panels B – D were kindly provided by Dr. Jiao Jin (Riechmann Group, Catalan Institution for Research and Advanced Studies, Barcelona). 3'OCS, octopine synthase terminator sequence; RB, right border of T-DNA; LB, left border of T-DNA. lacZ, a modified version of the β -galactosidase gene from *E. coli*; pNOS, nopaline synthase promoter sequence; BAR, BASTA® resistance gene; NOS, nopaline synthase terminator sequence; GR, glucocorticoid receptor.

untreated *35S-FIS* plants suggests that there is a reduced level of nuclear AP1-GR activity in untreated *pAP1-FIS* plants. Early flowering has been observed in plants ectopically over-expressing *AP1* (Mandel and Yanofsky, 1995). However, over-expressing *AP1* does not alter floral morphology (Mandel and Yanofsky, 1995). Therefore, the cause of decreased petal size and reduced pollen production in *35S-FIS* plants relative to *pAP1-FIS* plants is currently unclear. It can be speculated, though, that the altered floral organ morphology is a consequence of the ectopic expression of the AP1-GR fusion protein.

The results of *in situ* hybridizations (Figure 5.5) showed that the expression pattern of *API-GR* in *pAPI::API-GR* plants resembled that of endogenous *API* in wild-type plants (Gustafson-Brown et al., 1994). *API* mRNA usually accumulates throughout early flower buds until stage 3 of flower development when *AG* expression commences (Gustafson-Brown et al., 1994). *AG* has been shown to repress the expression of *API* in the inner whorls of developing flowers (Gustafson-Brown et al., 1994). Therefore, *API* mRNA is normally restricted to petals and sepals after stage 3 of flower development (Gustafson-Brown et al., 1994). The fact the *pAPI::API-GR* transgene can recapitulate the expression pattern of endogenous *API* mRNA suggests that the genomic fragment of *API* used for generating the construct contains all the regulatory regions required for proper *API* expression.

In sum, the *pAPI-FIS* led to an improvement in the morphology of petals and the fertility of stamens, relative to *35S-FIS* plants, when treated with dexamethasone. In addition, the transitioning time of the inflorescence-like meristem to flowering is delayed in the *pAPI-FIS* when compared to the *35S-FIS*, suggesting a reduced leakiness of the system. Finally, the expression pattern of the *API-GR* mRNA in *pAPI-FIS* matches that of endogenous *API* mRNA in wild-type plants. Therefore, *pAPI-FIS* plants appear to behave more similarly to wild-type plants than *35S-FIS* plants.

5.4.2 Floral induction systems based on the activation of a CAL-GR fusion protein

In order to identify genes that are regulated by the CAL transcription factor, two transgenic lines were generated. First, the *CAL* coding region was translationally fused to a portion of the *GR* receptor and placed under the control of the CaMV 35S promoter. This construct was used to transform *ap1 cal* plants. The primary transgenic lines responded to the dexamethasone treatment with the formation of synchronized flowers, which contained sepals and petals. The floral organ identity restoration observed in these plants was unexpected, since formation of these organs was not observed when the *CAL* coding region was expressed under the control of a portion of the *API* promoter in *ap1-1* plants (Alvarez-Buylla et al., 2006). However, rescue of the *ap1-1* mutant phenotype was observed when

the same promoter region was used to drive expression of the *API* coding sequence. Furthermore, it has been reported that overexpression of the *CAL* cDNA from the CaMV 35S promoter was not sufficient to restore sepal and petal identity in *ap1 cal* mutants (Savidge, 1996). Contrary to previously published results (Alvarez-Buylla et al., 2006; Savidge, 1996), the phenotypes presented here suggest that *CAL* has the capacity to compensate for *API* functions during flower development. A possible explanation for the discrepancy between the literature and the results presented here is that the rescue of sepal and petal development by *CAL* may require extremely high expression levels, which may not have been reached in the previous studies.

The second transgenic line generated expressed the *CAL-GR* fusion protein under the control of the putative *CAL* promoter. When transgenic lines harbouring this transgene were treated with a dexamethasone-containing solution, plants appeared to flower synchronously, and no sepals or petals developed within the resulting flowers. This phenotype is in agreement with the fact that *CAL* is not expressed in sepals or petals and does not appear to be involved in the formation of these organ types (Kempin et al., 1995). Given this information and the differences between the *35S::CAL-GR ap1 cal* and the *pCAL::CAL-GR ap1 cal* lines, it appears that the different expression patterns of *API* and *CAL* are the underlying causes of phenotypic differences observed in *ap1* and *cal* mutant plants. Therefore, the *pCAL::CAL-GR ap1 cal* line is most suited for use in the experiments we will conduct to identify genes regulated by the *CAL* transcription factor.

5.5 Future Work

In order to identify genes that are regulated specifically by *CAL*, microarray experiments will be performed using RNA extracted from dexamethasone-treated inflorescence tissue of *pCAL::CAL-GR ap1 cal* plants at different time points after hormone-treatment. This will lead to the identification of genes specifically regulated by *CAL*. The differentially expressed genes identified in these experiments can be compared with the genes differentially expressed upon *API-GR* activation (Wellmer et al., 2006). In doing so, the molecular basis of the phenotypic redundancy observed between *API* and *CAL* can be understood. Furthermore, the *pCAL::CAL-GR* line will be used for in ChIP assays for the genome-wide localization of *CAL* binding sites. A comparison of the microarray

datasets and binding data should then lead to detailed insights into the regulatory hierarchy downstream of this key floral regulator.

Chapter 6 – Domain-specific perturbation of *AG* and *KNU* function through amiRNAs

6.1 Introduction

AG is expressed in the third and fourth whorls of *Arabidopsis* flowers, where it mediates the specification of stamen and carpel identity (Yanofsky et al., 1990). Another function of *AG* is to control determinacy of the floral meristem (Drews et al., 1991; Yanofsky et al., 1990). In *ag* mutant flowers, the expression of the transcription factor *WUSCHEL* (*WUS*) persists for a longer developmental period than in wild-type flowers (Lenhard et al., 2001). *WUS* is required for meristem maintenance (Laux et al., 1996), and down-regulation of *WUS* expression coincides with the termination of the floral meristem (Lenhard et al., 2001). It is thought that *AG* is involved in repressing the expression of *WUS*, however, it is currently unclear whether this repression is direct or indirect (Lohmann et al., 2001). Given that *WUS* is expressed in only a few cells in the L3 layer of the floral meristem, and that *AG* is expressed in a much broader domain (Figure 6.1), it is possible that *AG* activity is only required in the region where *WUS* is expressed in order to repress its expression and terminate meristematic activity. Alternatively, *AG* expression could be required throughout the centre of the floral meristem to down-regulate *WUS*.

Although it is currently unclear exactly how *AG* terminates the floral meristem, the spatial domains in which *AG* activity is required for its different functions have been studied extensively. However, the precise regions where *AG* must be expressed in order to terminate growth of the floral meristem has only been partially examined. Using the *Cre/loxP* mosaic analysis method, Sieburth et al. (1998) restored *AG* function to specific cell layers in *ag-3* mutants. They showed that *AG* is essential in the L2 clonal cell layer in order to specify carpel and stamen identity, and is required in both the L2 and L3 clonal layers in order to terminate floral meristem growth (Sieburth et al., 1998). Moreover, they showed that epidermal cell identity could be communicated from the L2 to the L1 clonal cell layer (Sieburth et al., 1998). More recently, Urbanus et al. (2010) expressed an *AG*-GFP fusion protein in the L1 cell layer in a strong *ag* mutant. They found that the

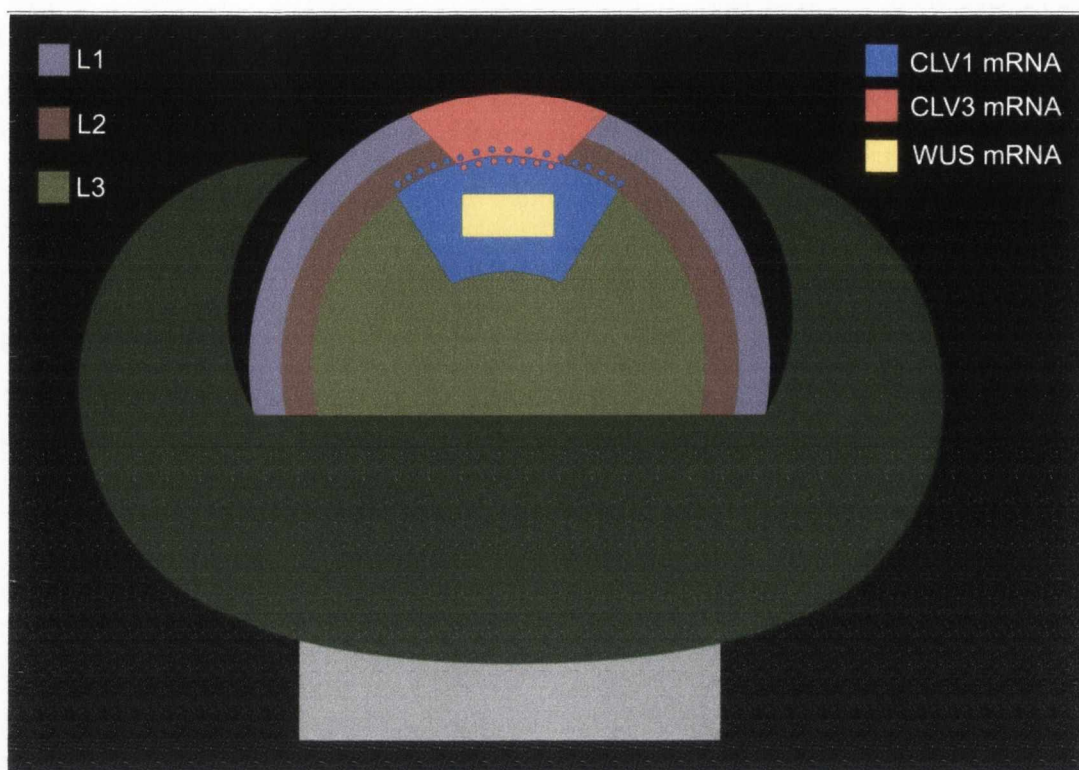


Figure 6.1 A schematic representation of the expression patterns of *CLV1*, *CLV3* and *WUS* within a stage 3 flower. The *CLV1* mRNA (blue) is present from stage 1 to the end of stage 5 of flower development in a group of cells of the L3 cell layer. It is also present in a few L2 clonal cell layers (Long and Barton, 1998). The expression domain of *CLV1* overlaps with whorl 4 and with the inner part of whorl 3 (Clark et al., 1997). The *CLV3* mRNA (red) is present in the L1 and L2 clonal cell layers in addition to a few underlying L3 cells (Fletcher et al., 1999). This domain is included within that of the whorl 4 of developing flowers from stage 2 until the end of stage 5. The *WUS* mRNA (yellow) is expressed in a small group of cells of the L3 clonal cell layer, which overlaps with whorl 4. The mRNA is present from stage 1 of flower development until the end of stage 5 (Lohmann et al., 2001).

expression of AG-GFP in this cell layer was sufficient to restore stamen and carpel identity, but not determinacy of the floral meristem. In addition, Urbanus et al. (2010) showed that the AG-GFP protein moved from the L1 to the L2 cell layer (but not to the L3 cell layer). The presence of AG-GFP in both L1 and L2 could explain in part the rescue of stamen and carpel development in these AG-GFP lines.

KNU has also been implicated in terminating the floral meristem; therefore, it is also of interest to identify the regions in which *KNU* activity is required (see Section 4.1 for a description of the role of *KNU* in floral meristem termination). It is unclear in which regions of the floral meristem *KNU* is required to terminate floral meristem proliferation. To more

accurately define the domains in which AG and KNU activities are required to terminate the floral meristem, I aimed at expressing the *AG4-miRNA* and the *KNU-miRNA* (see Section 3.2.1.1 for a description of the *AG4-miRNA*, see Section 4.2.1 for a description of the *KNU-miRNA*) under the control of the promoters of the meristem regulatory genes *CLAVATA1* (*CLV1*), *CLAVATA3* (*CLV3*) and *WUS* to disrupt AG and KNU functions in specific domains of the floral meristem. For this purpose, I used a previously described two-component expression system (Moore et al., 1998). The promoter sequences of *CLV1*, *CLV3* and *WUS* were transcriptionally fused to the coding region of the chimeric transcription factor LhG4 in previous studies (Gross-Hardt et al., 2002; Lenhard and Laux, 2003; Moore et al., 1998; Schoof et al., 2000). The LhG4 transcription factor can induce the expression of any sequence downstream of six copies of an optimized *lac* operator sequence (denoted as *6xOP* or *pOP* hereafter) (Moore et al., 1998). In order to disrupt expression of AG and KNU in the *CLV1*, *CLV3* and *WUS* domains, the *AG4-miRNA* and the *KNU-miRNA* were placed under the control of a *pOP* promoter. Previously available transgenic plants (Gross-Hardt et al., 2002; Lenhard and Laux, 2003; Schoof et al., 2000) containing the different Promoter::LhG4 constructs (these lines are referred to as “driver lines” hereafter) were transformed with T-DNA encoding the *pOP::AG4-miRNA* and *pOP::KNU-miRNA* transgenes.

6.2 Results

6.2.1 Expression of the *AG4-miRNA* in the *CLV1* expression domain

The *pCLV1::LhG4* driver line was transformed with a vector encoding the *6xOP::AG4-miRNA*. The resulting transgenic plants should express the *AG4-miRNA* in a domain in which the *CLV1* promoter is active (*i.e.* in the same domain as *CLV1* itself), which should result in a decrease of AG mRNA in this region (Figure 6.2A). Twenty-one primary transformants were obtained and are denoted as *pCLV1>>AG4* hereafter.

Of these 21 independent transformants, five plants produced wild-type flowers and six plants developed mildly bulged gynoecia (not shown). Seven additional independent

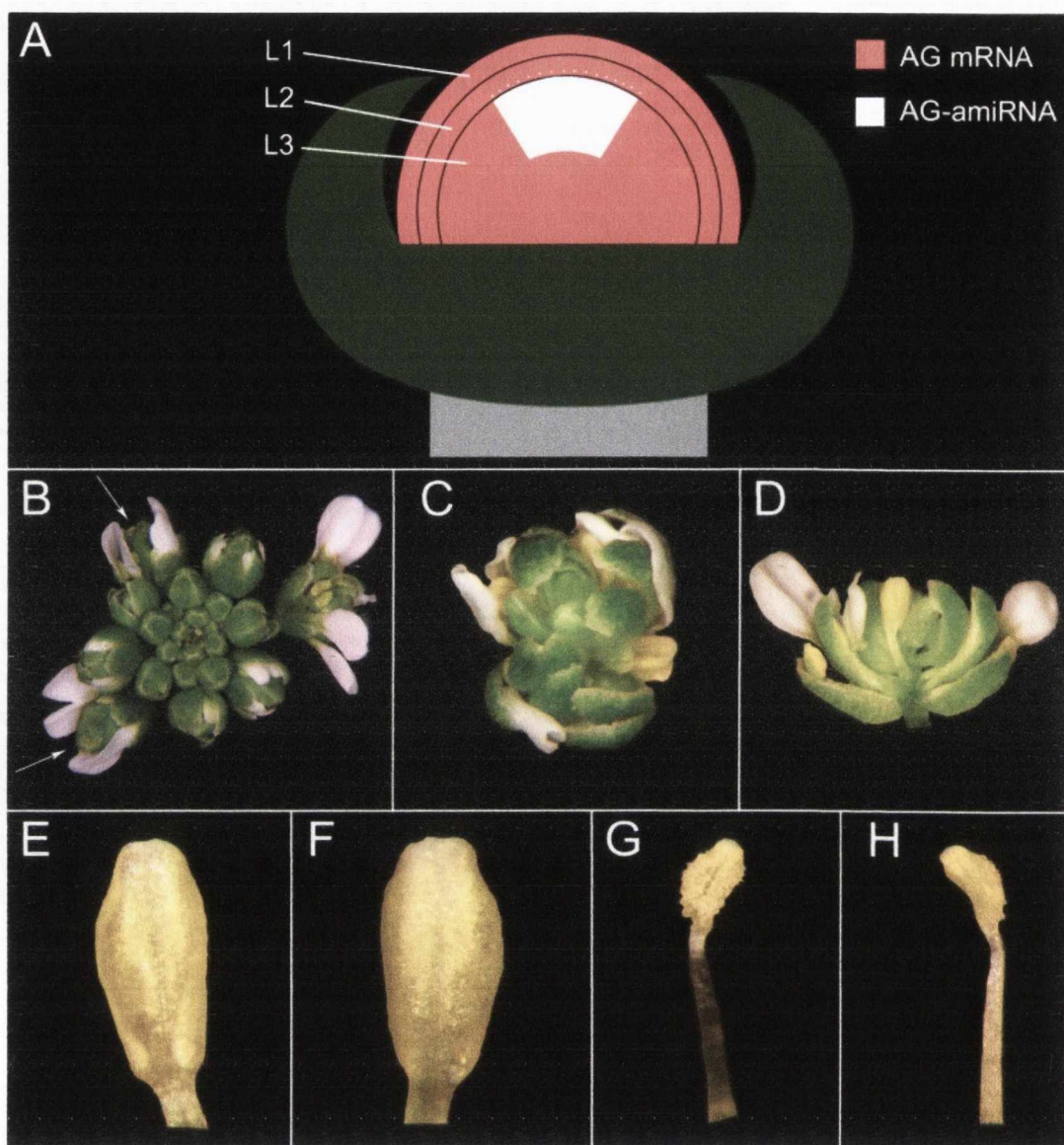


Figure 6.2 Expression of the *AG4-miRNA* in the *CLV1* mRNA domain. (A) A schematic of the region of the floral meristem where the *AG* mRNA should be absent/present and where the *AG4-miRNA* is expressed. (B-H) Floral phenotypes presented in *pCLV1>>AG4-miRNA* plants. (B) An inflorescence containing rounded flowers due to indeterminacy and early arising flowers that retain carpel identity (arrows). (C-D) Flowers that have lost determinacy and carpel identity. (C) Top view. (D) Side view with outer organs removed to facilitate viewing of the floral architecture. (E-F) A stamen with a partial petal identity from a flower at an early stage of development. (E) Adaxial view. (F) Abaxial view. (G-H) A stamen containing an anther that was partially converted to a petal from a flower at a late stage of development. (G) Adaxial view. (H) Abaxial view.

plants had lost floral meristem determinacy (*i.e.* they showed reiterations of sepals, petals and stamens) and lacked carpel identity, except in the first flowers formed (Figure 6.2B). In addition, the stamens of these plants were partially converted to petals (Figure 6.2E – H). The remaining three plants produced flowers that lacked any carpel identity, were indeterminate and produced stamens with partial petal identity (Figure 6.2C, D). These three transgenic lines were different because carpel tissue was not present on early-arising flowers whereas carpel tissue was present on the early-arising flower of the other transformants. The fourth whorl organs of these flowers had a sepal-like morphology and subsequent whorls repeated in a petal-stamen-sepal pattern. Organs in the inner whorls were often chimeric, containing both stamen and petal features (similar to Figure 6.3E, F but not shown for pCLV1>>AG4 plants).

More specifically, third whorl organs of these pCLV1>>AG4 flowers did not contain any locules at early stages of flower development (Figure 6.2E, F) and were lighter in colour than wild-type stamens at a similar stage. Moreover, these organs had an overall shape that was more similar to petals than to stamens. Despite these differences at early stage of flower development, third whorl organs developed into stamens that were similar to wild-type stamens, although the anthers remained partially petal-like (Figure 6.2G, H).

In summary, the phenotypes presented here show that expression of the *AG4-miRNA* from the *CLV1* promoter is sufficient to abolish determinacy and carpel identity. Moreover, it resulted in a partial conversion of stamens into petals.

6.2.2 Expression of the *AG4-miRNA* in the *CLV3* expression domain

Plants containing the *pCLV3::LhG4* transgene were transformed with the a T-DNA containing the *6xOP::AG4-miRNA* transgene using the floral dip method (Zhang et al., 2006). The resulting plants should result in a disruption of *AG* activity in the *CLV3* expression domain (Figure 6.3A), *i.e.* within the stem cells of the floral meristem. Twenty-one primary transformants were obtained and are denoted as pCLV3>>AG4 hereafter.

Of the 21 independent transformants obtained, four plants developed flowers with bulged gynoecia (not shown), nine plants displayed indeterminacy and partially disrupted

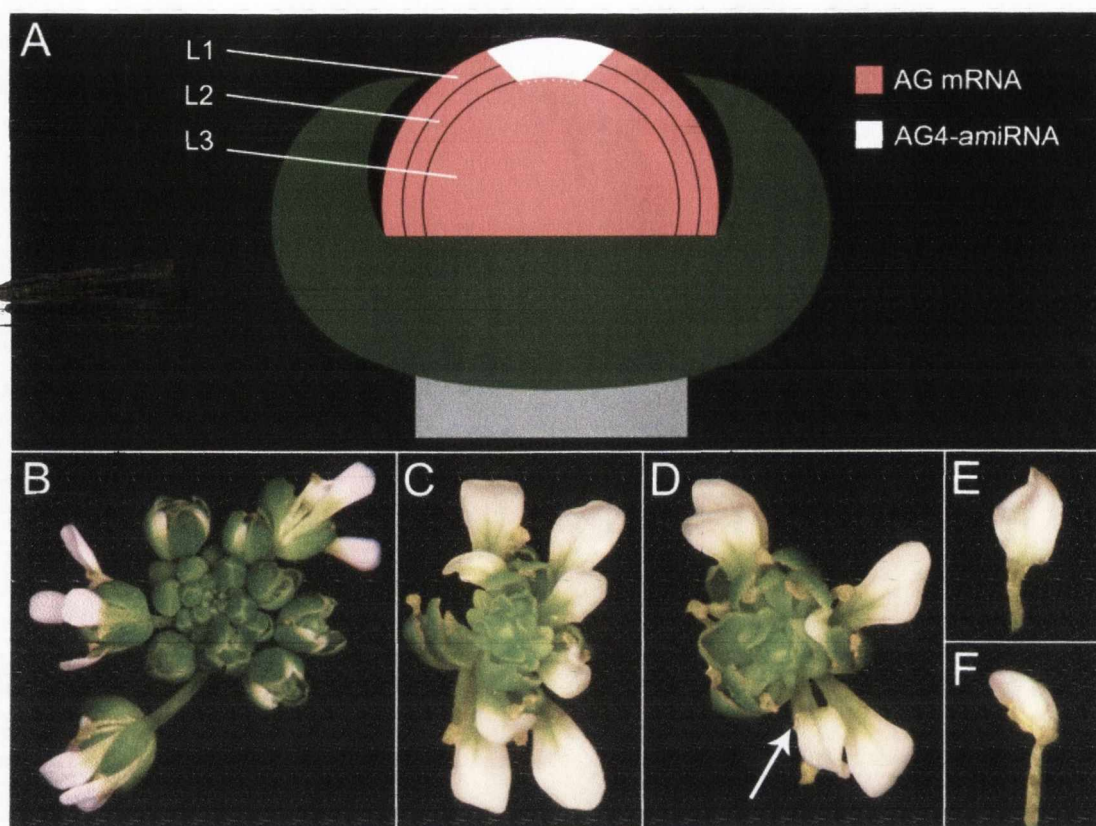


Figure 6.3 Expression of the *AG4-amiRNA* in the *CLV3* mRNA domain. (A) A schematic of the region of the FM where the *AG4-amiRNA* is expressed. (B-F) Floral phenotypes presented in pCLV3>>AG4-amiRNA plants. (B) An inflorescence containing rounded flowers due to indeterminacy and a loss of carpel identity. (C) A flower that has lost carpel identity and determinacy with reiterations of petals, stamens and sepals. (D) A flower that has lost carpel identity and determinacy, presenting petals with mild stamen identity (arrow) or morphologically wild-type petals. (E-F) A sample of a chimeric fifth whorl organ with stamen and petal identities. (E) A chimeric organ with a filamentous base and a petaloid tip. (F) A chimeric organ with a filamentous base and anther/petal tissue at the tip.

carpel identity (not shown), and the flowers of the remaining eight first generation transformants had lost all carpel identity and were indeterminate (Figure 6.3B – D). The fourth whorl organs of these flowers had a sepal-like morphology and subsequent whorls repeated in a petal-stamen-sepal pattern (Figure 6.3C), however, organs in the reiterated whorls were often chimeric, containing stamen and petal identities (Figure 6.3E, F). Stamens were not affected in these plants, primarily because the *CLV3* promoter should not be active in the third whorl of flowers.

In sum, these transgenic plants show that expression of the AG4-amiRNA from the *CLV3* promoter can abolish meristem determinacy and carpel identity, similarly to the

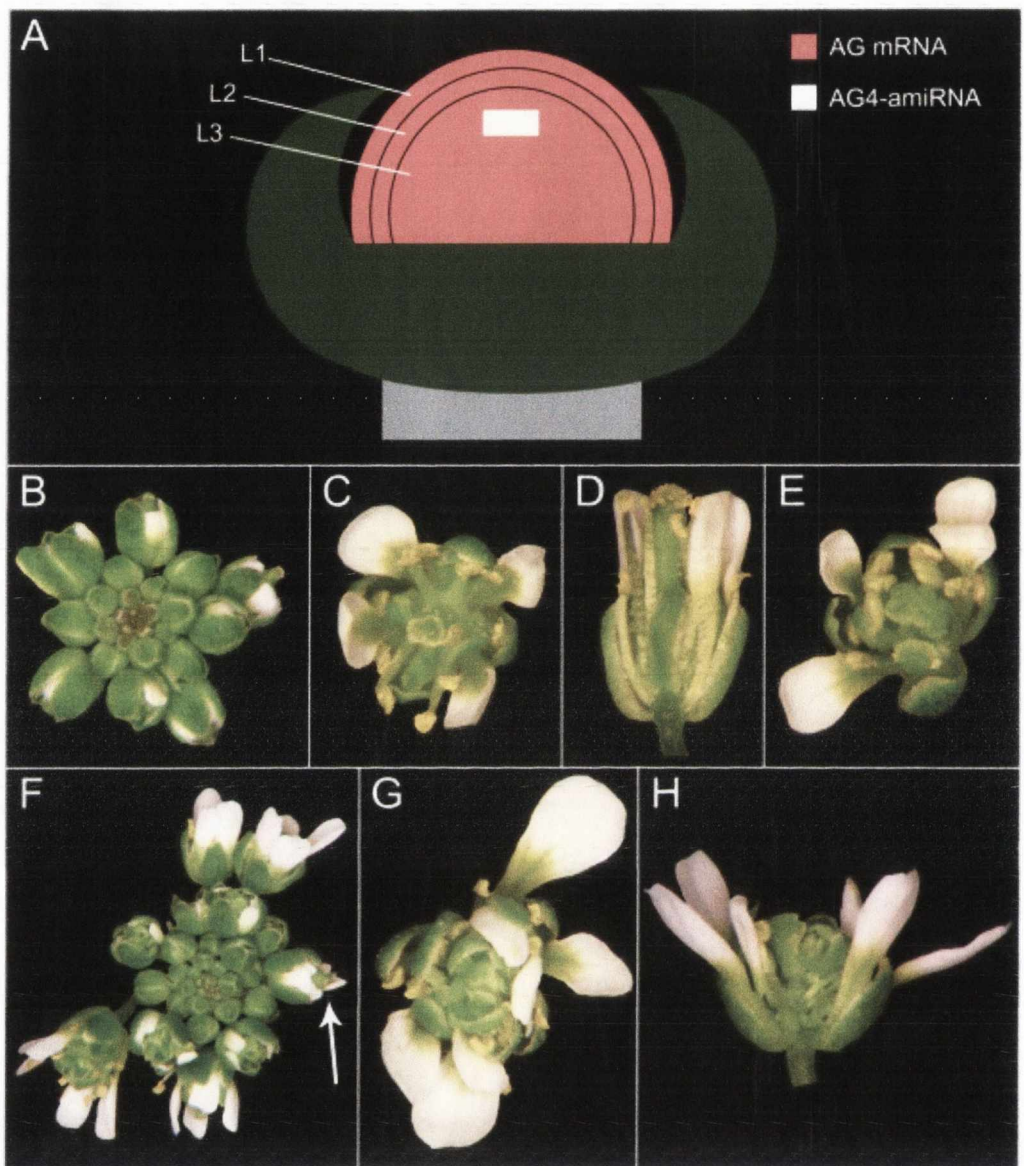


Figure 6.4 Expression of the *AG4-miRNA* in the *WUS* mRNA domain. (A) A schematic of the region of the FM where the *AG4-miRNA* is expressed. (B-E) Floral phenotypes presented in pWUS>>AG4-miRNA plants. (B) An inflorescence containing mildly rounded flowers due to the presence of additional whorls of organs. (C-D) Representative flowers from pWUS>>AG4-miRNA plants. (C) Top view. Note the presence of an additional whorl of sepals, as well as an additional whorl of stamens. (D) Side view of the flower shown in (C). Petals and sepals were removed. (E) Although carpels were sometimes not fused, carpel identity was always retained. (F) Second generation transformants, whose early-arising flowers retain carpel identity (arrow), but whose later-arising flowers lack a central gynoecium. (G and H) Flower of a second generation pWUS>>AG4 line showing a loss of determinacy with reiteration of sepals, petals and stamens and lacking carpels. (G) Top view. (H) Side view (some sepals and petals have been removed).

effects obtained when expressing the same amiRNA in the *CLV1* expression domain. Moreover, the identity of subsequent whorls can be affected.

6.2.3 Expression of the *AG4-amiRNA* in the *WUS* expression domain

The *AG4-amiRNA* was inserted downstream of the *6xOP* promoter and placed into the binary vector *pART27*. Plants containing the *pWUS::LhG4* transgene were transformed with this vector using the floral dip method. The resulting plants should result in a disruption of *AG* activity in the *WUS* domain (Figure 6.4A).

Sixteen primary transformants were obtained and are denoted as *pWUS>>AG4* hereafter. Similar to *pCLV1>>AG4* and *pCLV3>>AG4* lines, the flowers of *pWUS>>AG4* plants displayed floral indeterminacy. However, the defects were weaker, in that a single repetition of sepal-like organs and stamens (Figure 6.4C, D) was observed. Furthermore, in contrast with *pCLV1>>AG4* and *pCLV3>>AG4* lines, carpel identity was not completely lost. In fact, flowers of *pWUS>>AG4* T1 plants developed a central gynoecium, which had 2 fused carpels and was fertile (Figure 6.4B – D). No petals or carpel-like organs were observed in the reiterated whorls of these primary transformants. These plants sometimes produced flowers with carpels that were not completely fused. However, these carpels still formed stigmatic tissue and ovules (Figure 6.4E).

The progeny of all first generation transformants was then analysed. Twelve second-generation transformants developed flowers that were identical to those of the parental lines (not shown). Three other second-generation transformants developed flowers that sometimes lacked a central gynoecium. (Figure 6.4F). Finally, one second-generation plant did not form any flower that contained an organ with carpel identity (Figure 6.4F, G). The reiteration of whorls was also different, in that it presented petals (*i.e.* a reiteration of petals-stamens-sepals was observed).

To summarise, the expression of the *AG4-amiRNA* from the *WUS* promoter resulted in plants that produced flowers that were partially indeterminate and that sometimes lacked carpel identity in the second generation (see Section 6.3.2 for Discussion).

A

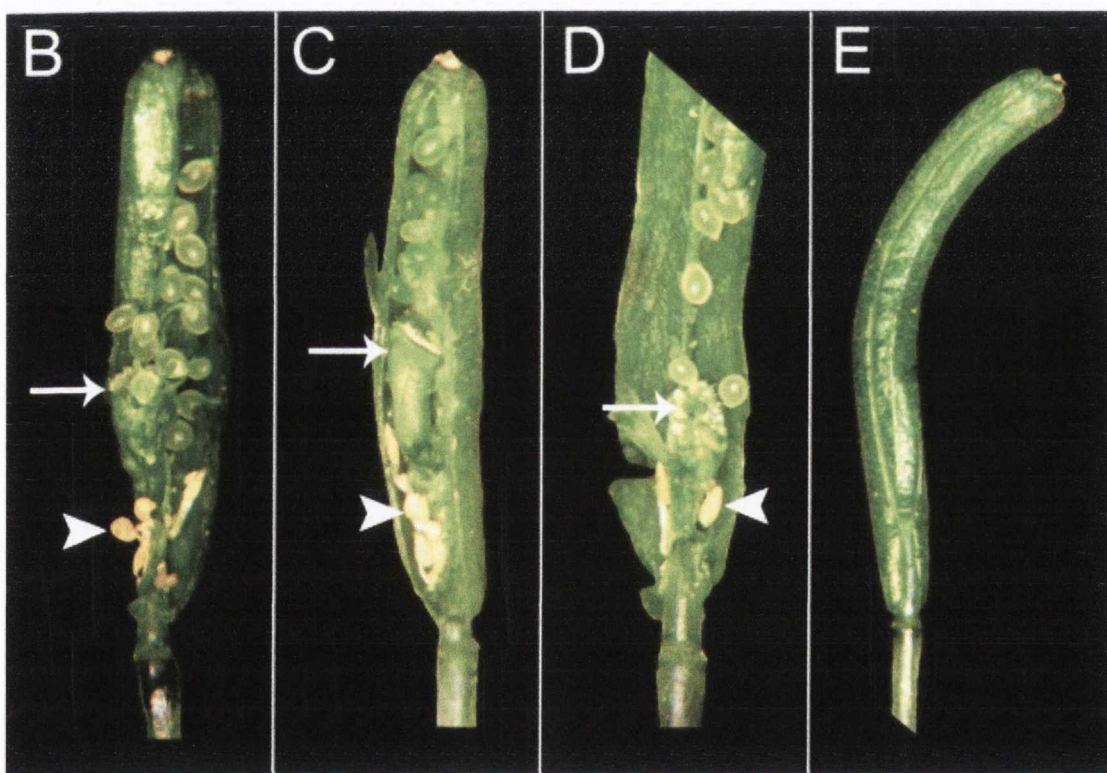
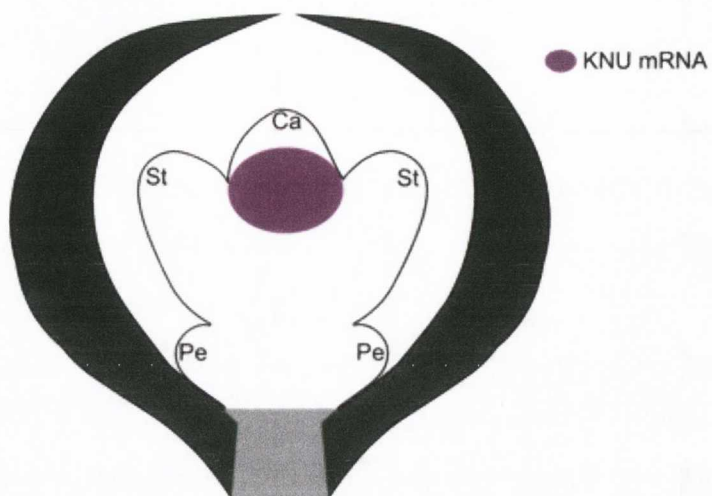


Figure 6.5 Expression of the *KNU-amiRNA* in the *CLV1*, *CLV3* and *WUS* expression domains. (A) A schematic of a stage 6 flower when *KNU* mRNA is first observed at the base of the developing carpel. Ca, carpel primordia; St, stamen primordia; Pe, petal primordia. (B to D) Ectopic reproductive organs within the developing siliques of *pCLV1>>KNU-amiRNA* plants (B), *pCLV3>>KNU-amiRNA* plants (C), and *pWUS>>KNU-amiRNA* plants (D). Arrows indicate ectopic gynoecia, whereas arrowheads indicate ectopic stamens. (E) A tricarpelloid silique from a *pWUS>>KNU-amiRNA* plant.

6.2.4 Expression of the *KNU-amiRNA* in the *CLV1*, *CLV3* and *WUS* expression domains

Plants expressing the LhG4 transcription factor in the domains of *CLV1*, *CLV3* and *WUS* were transformed with a plasmid carrying the *6xOP::KNU-amiRNA* transgene. Primary transformants were identified based on their resistance to kanamycin and by genotyping. These plants are termed pCLV1>>KNU-amiRNA, pCLV3>>KNU-amiRNA and pWUS>>KNU-amiRNA, respectively.

Twenty-one pCLV1>>KNU-amiRNA primary transformants were obtained. Although mild *knu*-like phenotypes, such as short siliques, were sometimes observed (not shown), the siliques of these plants did not form internal ectopic organs. The progeny of these plants was analysed. Twelve of these second-generation transformants appeared phenotypically wild type. Six other T₂ populations presented mild *knu*-like phenotypes, such as short siliques. As observed in the T₁ generation, these siliques did not contain ectopic organs (not shown). Finally, the offspring of six additional independent transformants developed ectopic reproductive organs within the siliques (Figure 6.5B). The phenotypes presented were similar to those observed when the *KNU-amiRNA* was over-expressed from the *35S* promoter (see Chapter 4). However, the frequency of bulged siliques was much higher in these pCLV1>>KNU-amiRNA lines compared to *35S::KNU-amiRNA* lines (Figure 6.6, pCLV1>>KNU-amiRNA, red bar). Inflorescences from plants that presented *knu*-like phenotypes were collected and pooled, RNA was extracted and cDNA was synthesized. Quantitative real-time PCR analysis revealed that *KNU* mRNA levels were reduced in these plants to approximately 28% of wild-type levels (Figure 6.6). This showed that *KNU* activity can be efficiently disrupted by expressing the *KNU-amiRNA* in the background of a pCLV1 driver line.

Thirty-four independent pCLV3>>KNU-amiRNA transformants were isolated. Of these transformants, 21 did not present with any phenotype, one contained short siliques and twelve produced siliques with ectopic reproductive organs within them (Figure 6.5C). These phenotypes were similar to the ones described for the ectopic over-expression of the *KNU-amiRNA* (Chapter 4). One homozygous line that exhibited a strong *knu* phenotype was

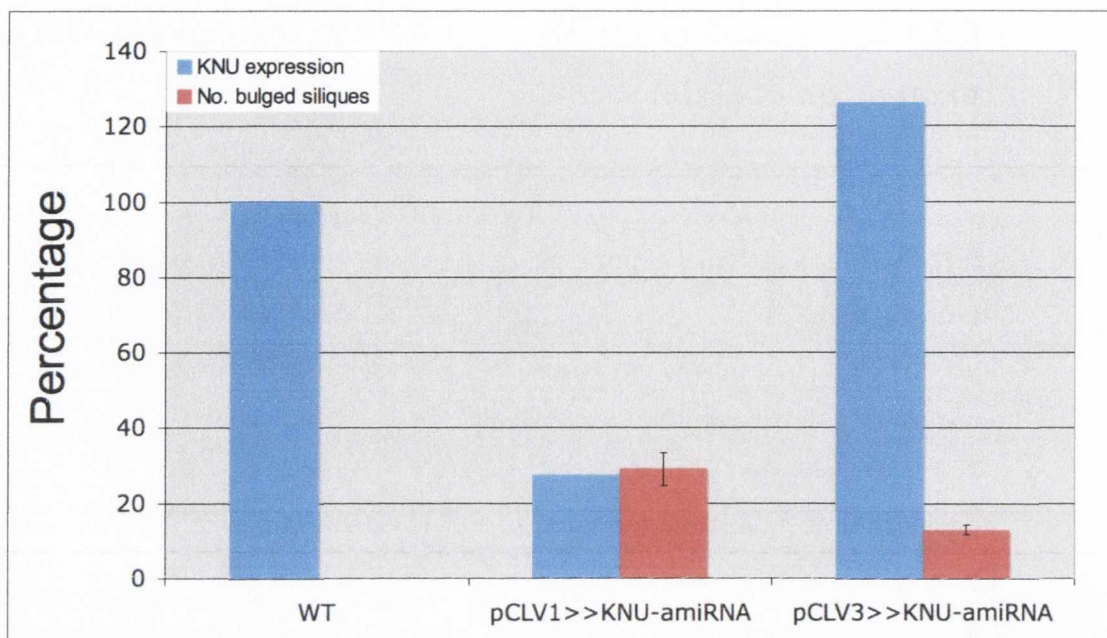


Figure 6.6 Expression of the *KNU* mRNA and percentage of bulged siliques in pCLV1>>KNU-amiRNA and pCLV3>>KNU-amiRNA plants. The qRT-PCR expression values (blue bars) are represented as a fraction of the wild-type *KNU* mRNA levels. No error bars are present, as these values are an average of technical duplicates. The difference between each technical duplicate was less than 0.5 Cp values. Number of bulged siliques is relative to the total number of siliques counted (red bars): pCLV1>>KNU-amiRNA, 7 plants were counted with at least 16 siliques per plant; pCLV3>>KNU-amiRNA, 17 plants counted with at least 19 siliques per plant. Error bars are s.e.m. The primer pairs DM-94/DM-95 and DM-97/DM-98 were used to amplify the *KNU* and *REF2* (At4g34270) cDNAs, respectively.

isolated and the number of knuckle-like siliques was counted (Figure 6.6, pCLV3>>KNU-amiRNA, red bar). Approximately 13% of the plants derived from this line contained bulged siliques. In order to test if the phenotype correlated with decreased levels of *KNU* mRNA, qRT-PCR was carried out and showed that the levels of *KNU* mRNA in this pCLV3>>KNU-amiRNA line were not significantly reduced in comparison to the wild type (Figure 6.6, pCLV3>>KNU-amiRNA, blue bar). The lack of a correlation between phenotypic effects and *KNU* transcript levels in this line could be a consequence of the comparatively small expression domain of *CLV3*. In other words, *KNU* expression might not appear to be down-regulated in pCLV3>>KNU-amiRNA plants because of a spatially limited knockdown effect.

Twenty-eight independent pWUS>>KNU-amiRNA transformants were isolated. Of these lines, 24 appeared phenotypically wild type, and only four lines displayed a *knu*-like phenotype (Figure 6.5C). These phenotypes were similar to those observed in the

35S::*KNU-amiRNA* lines, although the frequency of bulged siliques was lower in the pWUS>>*KNU-amiRNA* plants (two per plant on average). Some of the siliques produced by these four plants were tri-carpeloid (Figure 6.5D), which was not observed in pCLV1>>*KNU-amiRNA* or pCLV3>>*KNU-amiRNA* plants. In summary, the phenotypes observed suggest that expressing the *KNU-amiRNA* from the *WUS* promoter could disrupt *KNU* activity, although the extent of this disruption (*i.e.* the number of siliques affected and expression levels of *KNU* mRNA) is currently unknown.

6.3 Discussion

Artificial miRNAs that have been shown to target the *AG* and *KNU* mRNAs efficiently (described in Chapters 3 and 4) were cloned under the control of a *6xOP* promoter and introduced into driver lines that express the *LhG4* transcription from the *CLV1*, the *CLV3* or the *WUS* promoters (denoted as *Promoter Of Interest (POI)* when referring to all three promoters at once). I analysed the phenotypes of the different transgenic plants obtained to gain further insights into the role of *AG* and *KNU* in the control of meristem determinacy. The transgenic plants isolated produced a subset of the *AG* and *KNU* mutant phenotypes, however, in order to interpret these phenotypes correctly, additional control experiments need to be performed. These control experiments, as well as the results obtained, are discussed below.

6.3.1 General considerations on the expression of amiRNAs in specific domains

In order to correctly interpret the phenotypes presented above and validate our approach, several control experiments must be performed. For example, the expression pattern of *LhG4* must be determined in each of the driver lines. For each driver line generated, the promoter region used has been shown to be sufficient to rescue the respective mutant, when fused to the coding sequence of the corresponding gene. However, in the driver lines, the coding regions of *CLV1*, *CLV3* and *WUS* have been replaced by that of *LhG4*, which could result in the removal of important regulatory elements. Thus, the

expression pattern of the *LhG4* transcription factor needs to be validated in the different driver lines used.

The genomic region encompassing ~5.6 kb upstream and ~700 bp downstream of the *CLV1* coding sequence was shown to be sufficient to partially rescue the *clv1-1* phenotype (Clark et al., 1997). In order to generate the *pCLV1::LhG4* driver line, the same genomic region was used, except that the *CLV1* coding sequence was replaced by that of *LhG4* (Schoof et al., 2000). The expression pattern of the *LhG4* mRNA in that line has not been analysed in detail. It therefore needs to be confirmed that the LhG4 transcription factor is indeed expressed in the same domain as *CLV1*.

It has been shown that a genomic fragment containing the *CLV3* coding sequence, as well as ~3.9 kb upstream of the start codon, is sufficient for a rescue of the *clv3-1* mutant (Lenhard and Laux, 2003). To generate a *pCLV3::LhG4* driver line, the same ~3.9 kb promoter region was placed upstream of the *LhG4* coding sequence (Lenhard and Laux, 2003). The expression pattern of *LhG4* in this driver line has not been determined directly. However, when a *pOP::CLV3* construct was introduced into this *pCLV3::LhG4* driver line, it was sufficient to rescue the *clv3-1* phenotype. Furthermore, the domain of GUS activity in a *pCLV3::LhG4/pOP::GUS* line was found to be similar to that of the endogenous *CLV3* gene (Lenhard and Laux, 2003).

The expression pattern of *LhG4* in the *pWUS::LhG4* driver line has not been fully characterised. However, it has been described that introduction of a *pOP::WUS-GUS* coding sequence into this *pWUS::LhG4* driver line was sufficient to rescue the *wus-1* mutant (Gross-Hardt et al., 2002). Although the expression pattern of the *pOP::WUS-GUS* reporter in ovules was identical to that of the endogenous *WUS* mRNA, the expression pattern of the WUS-GUS reporter or of the LhG4 transcription factor in the floral meristem was not verified in these lines.

In summary, although there is evidence that the LhG4 transcription factor is expressed in the expected domains, all 3 driver lines used in this study need to be characterized in more detail. In addition to a correct spatial localization of the LhG4 transcription factor, its expression at different stages of flower development is also important. Indeed, the rate of decay of the *LhG4* mRNA or of the LhG4 protein might be very slow, which could lead to the presence of the LhG4 transcription factor long after stage 5 of flower development, which is when expression of the *CLV1*, *CLV3* and *WUS* mRNAs ends (see Section 6.4.3 for further discussion).

In addition to verifying the expression pattern of LhG4, the cell autonomy of the amiRNAs used must be established. Non-cell autonomous activity of certain endogenous miRNAs has been observed (Allen et al., 2005; Carlsbecker et al., 2010) that transmit a non-cell autonomous signal by targeting *trans*-acting short interfering RNAs (tasiRNAs) which in turn trigger double-stranded RNA formation and siRNA production (see Section 1.5.2.1 and (Ossowski et al., 2008)). Therefore, the localization of the amiRNAs, and the effect on the expression pattern of target transcripts, should be analysed in the mis-expression lines. The expression pattern of the amiRNAs can be investigated by *in situ* hybridisation, using a published method (Valoczi et al., 2006). Furthermore, the effect of the amiRNAs on the mRNAs of the target genes can be analysed using *in situ* hybridisation or reporter proteins. Using the *pAG::AG-GFP* transgenic line (Chapter 3) in combination with the *pOP::AG4-amiRNA* and *pPOI::LhG4* transgenes, it should be possible, through fluorescence confocal microscopy, to determine the meristem regions in which the amiRNA is functional and suppresses AG-GFP accumulation..

These control experiments are essential to correctly interpret the phenotypes observed above and could be referred to as a proof of concept for further experiments of a similar design.

6.3.2 Expression of the *AG4-amiRNA* in the *CLV1*, *CLV3* and *WUS* mRNA domains

The characterization of lines expressing the *AG4-amiRNA* in the *CLV1*, *CLV3* and *WUS* expression domains suggests that the down-regulation of AG activity in these 3 domains results in the abolishment of carpel identity and meristem determinacy. Additionally, the expression of the *AG4-amiRNA* in the *CLV1* domain resulted in some stamen to petal transformations, which were not observed in the pCLV3 and pWUS driver lines. These phenotypes are discussed below.

The *CLV3* promoter is active in the centre of floral meristems in the L1 and L2 layers, as well as in the upper portion of the L3 layer until stage 5 of flower development (Lenhard and Laux, 2003; Mayer et al., 1998). Sieburth *et al.* (1998) showed that expression of *AG* in the L2 and L1 layers was essential to restore carpel and stamen development in an *ag-3* mutant, whereas expression of *AG* in both the L2 and L3 was required to restore

determinacy. Based on these results, I expected that down-regulation of AG activity in the *CLV3* expression domain would lead to carpel development defects, but not necessarily to a disruption of stamen development, because the expression of the *AG4-miRNA* in the *CLV3* domain should not affect the levels of *AG* mRNA in the presumptive third whorl of the flower. Furthermore, because the *CLV3* promoter is active in the L2 layer and in the uppermost region of the L3 layer, indeterminate phenotypes were expected. The phenotype observed in pCLV3>>AG4-miRNA lines was consistent with both suppositions (*i.e.* stamen development was not affected, whereas carpels did not develop and determinacy was affected).

The *CLV1* promoter is active in the centre of floral meristems in the L3 layer and in a few cells of the L2 cell layer (Long and Barton, 1998). *CLV1* mRNA is present until stage 5 and underlies the third and fourth whorls (Figure 6.1) (Clark et al., 1997). Based on the expression of *CLV1* and the results presented by Sieburth *et al.* (1998), one might expect pCLV1>>AG4-miRNA plants to form an indeterminate flower with carpel identity, and potentially with stamen development defects. Furthermore, Urbanus *et al.* (2010) found that carpel and stamen identity could be restored in an *ag* mutant by expressing an *AG-GFP* transgene from the L1. If the presence of the AG protein in the L1 is sufficient to restore carpel and stamen identity, then pCLV1>>AG4-miRNA flowers should produce carpels, considering that AG protein movement from the L1 to the L2 should not be disrupted in the pCLV1>>AG4-miRNA lines (only *AG* mRNA accumulation should be affected). The phenotypes observed in pCLV1>>AG4-miRNA plants were only partially consistent with these data. Carpel identity was abolished in pCLV1>>AG4-miRNA lines, which is inconsistent with the L1 expression of AG being sufficient to define carpel identity. It is possible that the *AG4-miRNA* is expressed in a broader domain than described due to activity of the LhG4 transcription factor. In addition, non-cell autonomous effects of the *AG4-miRNA* could contribute to the phenotype observed.

In addition to indeterminacy and the lack of carpels, stamen development was affected in the pCLV1>>AG4-miRNA lines. Stamens had partial petal identity at early stages of flower development, but ultimately developed into stamens with long filaments and viable pollen. However, the anthers of these stamens still had partial petal identity (Figure 6.2H). The reversion of petaloid-stamens to stamens during flower development is consistent with the findings of the ethanol-inducible *AG4-miRNA* experiments, which suggest that stamens are specified over a large developmental window (Chapter 3, Figure

3.5). Ito *et al.* (2007) found that anther and filament formation was specified between stages 7 and 8 (Ito *et al.*, 2007), however, anthers were partially transformed into petals in the pCLV1>>AG4-miRNA lines. This suggests that AG-dependent anther development begins before stage 6 of flower development.

The *WUS* promoter used should drive expression of *LhG4* in the centre of floral meristems in the inner region of the L3 domain (Figure 6.1); therefore, the levels of *AG* mRNA are expected to decrease in this small domain of the floral meristem only, and should not be affected in the L2 and L1 layers (Figure 6.4A). Because previously published data indicated that expression of *AG* is required in both the L2 and L3 layers in order to restore determinacy in an *ag-3* mutant, I expected that flowers of pWUS>>AG4-miRNA lines would show determinacy defects. Furthermore, Sieburth *et al.* have shown that the formation of carpels and stamens require AG activity in the L1 and L2 layers, but not in the L3 layer. Similarly, Urbanus *et al.* (2010) describe a restoration of reproductive organ identity, but not determinacy, when *AG* was expressed from an epidermis specific promoter (Urbanus *et al.*, 2010). Thus, I expected that expression of the AG4-miRNA in the pWUS driver lines would not result in the abolishment of stamens or carpels. My results did not always support these assumptions, and varied depending on the generation that was being analysed (*i.e.* T1 or T2 generations). The flower defects observed in plants from the T1 generation were closer to the expected phenotypes, in that flowers of these plants formed carpels and stamens and had a weak indeterminacy phenotype (*i.e.* flowers only had an extra whorl of sepals and an additional whorl of stamens). In contrast, the offspring of some of the T1 lines (the T2 generation) showed stronger defects, as carpels could be absent and additional extra whorls could be formed. It is possible that the stronger phenotypes observed in some lines of the T2 generation are due to higher levels of the *AG4-miRNA*, as transgene expression can vary depending on the genomic position of a given T-DNA. Additionally, the number of T-DNAs in a first generation transformant can vary. Control experiments are being carried out to explain this stronger phenotype in the T2 generation. Independently of this problem, it remains unclear why reducing *AG* mRNA levels in the *WUS* expression domain could result in such strong *ag*-like phenotypes. There are several possible hypotheses, which need to be explored in future experiments. For example, it is possible that the pWUS::*LhG4* driver line expresses *LhG4* in a broader domain than expected. As mentioned above, the expression of the WUS-GUS reporter fusion in a pWUS::*LhG4*/pOp::*WUS-GUS* line has only been tested in ovules and not in early-stage

floral meristem. Moreover, potential non-cell autonomous effects of the *AG4-miRNA* could account for this phenotype.

In summary, the lack of carpel identity in pWUS>>AG4-miRNA and pCLV1>>AG4-miRNA plants is not consistent with the findings of Urbanus *et al.* (2010). Furthermore, the lack of carpel identity in pWUS>>AG4-miRNA plants is not in agreement with the findings of Sieburth *et al.* (1998) or Urbanus *et al.* (2010). However, these phenotypes are consistent with the idea of WUS mediating the activation of *AG* in the centre of the flower (Lohmann *et al.*, 2001). The WUS protein has been proposed to activate the transcription of *AG*, along with the floral meristem identity gene, *LEAFY* (*LFY*). Therefore, in the pWUS>>AG4-miRNA lines, the *WUS* promoter would be active before the expression of *AG* and the *AG4-miRNA* would already be present. In this scenario, *AG* mRNA would be degraded as it is transcribed. This, in theory, could suppress carpel identity formation.

Overall, the results presented here suggest that *AG* activity is required throughout the floral meristem to generate a determinate floral meristem. This is in keeping with the idea that floral meristem determinacy is very sensitive to *AG* perturbation, as was observed in the ethanol-inducible experiments (see Section 3.2.2.2). Specification of carpel identity also appears to be sensitive to *AG* perturbation, although less so than floral meristem determinacy. This is in agreement with the *AG* dosage-dependence of each activity of *AG* proposed by Sieburth *et al.* (1995). There are some points that require further investigation, most notably the phenotypes observed in the pWUS>>AG4-miRNA lines. Experiments outlining further investigation are presented in Section 6.5.

6.3.3 Expression of the *KNU-miRNA* in the *CLV1*, *CLV3* and *WUS* mRNA domains

All driver lines expressing the *KNU-miRNA* developed *knu*-like phenotypes. The mRNAs of *CLV1*, *CLV3* and *WUS* have been shown to be present until stage 5 of flower development when expression ceases. Because expression of *KNU* has been reported to commence at stage 6 of flower development (Payne *et al.*, 2004), no phenotypes should be observed in the different driver lines expressing the *KNU-miRNA*. There are several possible explanations for this: 1) The rate of decay of the *LhG4* mRNA or LhG4 protein

could be slow enough so as to be present at stage 6 of flower development, or beyond; 2) The rate of decay of the amiRNA could be slow enough so that it could be present at stage 6 of flower development, or beyond; 3) The activity of the *WUS*, *CLV1* and *CLV3* promoters continues beyond stage 5; or 4) *KNU* is expressed earlier than stage 6 of flower development. The control experiments described in section 6.3.1 should enable us to test the first 3 possibilities. As for the fourth hypothesis, the *KNU* expression pattern was described by Payne *et al.* (2004) using translational fusions between a putative *KNU* promoter, the *KNU* coding sequence and a *GUS* reporter, as well as transcriptional fusions using the same promoter sequences and *GUS*. Additionally, *in situ* hybridisations were performed. Based on these 3 approaches, the authors concluded that *KNU* mRNA was not present before stage 6 of flower development. However, a weak expression at earlier stages of flower development could have been missed, which needs to be investigated.

Despite the above-mentioned caveats, the experiments presented here suggest that expression of a *KNU* amiRNA in the 3 different driver lines used is sufficient to cause *knu*-like phenotypes. Surprisingly, the *knu*-like phenotype of pCLV1>>*KNU*-amiRNA lines was more frequent than that observed in plants expressing the same *KNU* amiRNA from the 35S promoter, suggesting that expression of *KNU* in the *CLV1* expression domain could be important to *KNU* function (Figure 6.6). Although this experiment has only been performed once and needs to be repeated, it appears that the levels of *KNU* mRNA were similar in pCLV1>>*KNU*-amiRNA plants (28%) and in the 35S::KNU-amiRNA plants (25%) (with the exception of line #6). Altogether, these results suggest that the *KNU* amiRNA expressed from the 35S promoter is not efficiently targeting the *KNU* mRNA in domains that are essential for its activity. Activity of the 35S CaMV promoter activity has been described to vary between different tissues (Skinner *et al.*, 2004) and could be a source of the low frequency of bulged siliques in 35S::KNU-amiRNA lines compared to the *knu-1* allele and the pCLV1>>*KNU*-amiRNA lines.

Tri-carpelloid siliques were only observed in pWUS>>*KNU*-amiRNA plants, which could suggest that *KNU* activity may have to be disrupted in a specific subset of cells (*i.e.* the *WUS* expression domain) to generate this phenotypic effect,. It is unclear why pCLV1>>*KNU*-amiRNA lines developed ectopic organs within siliques but not tri-carpelloid siliques, since the expression domain of *CLV1* encompasses that of *WUS*. This point will have to be investigated further.

To summarise, the phenotypes presented here show that KNU activity can be disrupted by expressing the *KNU-amiRNA* in different domains of the flower. The disparity between the temporal expression patterns of *CLV1*, *CLV3*, *WUS* and *KNU* needs to be investigated further, as described above.

6.4 Future Work

As mentioned in section 6.3.1, the expression pattern of *LhG4* must be confirmed in every driver line. This will be done using *in situ* hybridisation. In addition, the cell autonomous effect of the amiRNAs must be investigated. This again can be addressed through *in situ* hybridisation, but also through confocal microscopy of lines expressing an AG-GFP fusion protein in the *POI>>AG4-amiRNA* background. The latter experiments can also be used to confirm decreased levels of the target mRNA/protein in each of the lines generated. The success of these experiments is essential for interpretation of the phenotypes observed and could be cited as a proof of concept for experiments of similar design.

Additionally, the *pOP::AG4-amiRNA* and *pOP::KNU-amiRNA* transgenes from each of the strongest lines will be isolated and backcrossed into the original driver lines and the phenotypes will be compared. This will allow the assessment of the contribution of the T-DNA positions.

Finally, line #5 of the *pCLV1>>KNU-amiRNA* plants will be isolated in a homozygous state in an *L-er* background and crossed with the FIS. This line will be compared with line #28 of the *35S::KNU-amiRNA* lines for its ability to identify genes responsive to KNU function. Additionally, the *KNU-amiRNA* will be placed under the control of the *KNU* promoter and expressed in plants. The phenotypes will be compared with the *35S::KNU-amiRNA* and *pCLV1>>KNU-amiRNA* lines.

Chapter 7 – General Discussion and Future Perspectives

In this study, I set out to characterise the function of *AGAMOUS* (*AG*), a C-class floral homeotic gene, as a master regulator of the gene regulatory network underlying flower development (Bowman et al., 1989). *AG* is a MADS-box domain transcription factor that functions to specify the organ identity of male and female reproductive organs in *Arabidopsis* (Yanofsky et al., 1990). Despite recent advances in our understanding of *AG* function, it has remained largely unclear how *AG* acts to control the specification and development of the reproductive organs. In particular, functions of *AG* at early stages of flower development, when organs are specified, have remained elusive. The floral induction system (Wellmer et al., 2006) used (and improved) in this study represents an important progression in the analysis of gene regulatory networks underlying flower development. This system allows the collection of large amounts of synchronous flower buds at very early stages of flower development, which, prior to the generation of the floral induction system, was extremely difficult (Wellmer et al., 2004). Although improvements of the floral induction system developed by Wellmer et al. (2006) have been made in this study, its versatility could be improved further, as discussed below.

In this study, I generated some valuable data generated by perturbing *AG* activity followed by microarray expression profiling, and by identifying binding sites of *AG* using ChIP-Seq technology. These data have greatly facilitated our understanding of the gene regulatory networks underlying flower development. However, they would not have been sufficient on their own, as another important advance I used was based on the prediction and generation of artificial microRNAs (Schwab et al., 2006). Although publicly available mutant collections based on the integration of a T-DNA or a transposon in different regions of the genome have been generated, it is not possible to specifically and inducibly disrupt the function of a given gene using such insertion lines (Alonso and Ecker, 2006; Alonso et al., 2003). Furthermore, the use of RNAi can have significant drawbacks, in that it can lead to a large number of unknown off-target genes. In contrast, because the sequence targeting the mRNA of the gene of interest is much shorter than for RNAi, possible off-target effects are reduced. However, as illustrated by our attempts to generate an efficient *AG*-specific amiRNA, several candidate amiRNAs need to be screened in order to identify one that perturbs sufficiently the function of the gene of interest. Finally, a functional amiRNA can

be placed under the control of an inducible promoter to identify the genes that are directly responsive to a disruption of the target gene's function. Alternatively, specific amiRNAs can be expressed in different tissues to infer domain specific functions of a given gene.

Using an ethanol-inducible perturbation system to drive expression of amiRNAs highlighted another difficulty, namely the limited selection of highly responsive inducible systems (Moore et al., 2006). Of the systems available, the glucocorticoid receptor (GR)-based steroid inducible system appears to affect the development of plants the least and, when combined with the *LhG4/pOP* system, can induce the expression of genes under the control of the *pOP* promoter to high levels (Moore et al., 2006). However, the current forms of the floral induction system, including its new variants, are based on the fusion of a GR domain with AP1, which precludes the use of this induction system to express amiRNAs. As exemplified by the large number of genes that are misexpressed when using an ethanol-inducible system in our experiments, the identification of a suitable alternative inducible promoter is currently a limiting step in performing high-throughput inducible perturbation analyses.

Over the last decade, an ever-increasing number of genes that are directly regulated by AG have been identified, most of them encoding other transcription factors (Sablowski, 2009). This led to the conjecture that AG activity leads to the foundation of a complex regulatory network involving numerous transcription factors and signalling pathways to form reproductive organs and tissues.

To test this hypothesis, and to further understand how reproductive organs are formed, I used expression profiling by microarray analysis and ChIP-Seq to identify the set of genes that are directly or indirectly activated and repressed by AG. The data obtained largely confirmed the idea that AG acts as a master regulator of flower development primarily by directly activating other transcription factors. The transcription factors regulated by AG then specify distinct tissues and organ identities in specific sections of the reproductive organs. Our data further suggests that AG itself may continue being active in these tissues and could co-regulate other downstream targets by forming protein complexes with some of the transcription factors it activated at an earlier stage of flower development. This idea can be illustrated by the regulation of the B3 transcription factor *VERDANDI* (*VDD*), which has been shown to be involved in the specification of female reproductive development (Matias-Hernandez et al., 2010). Expression of this transcription factor has been shown to be regulated by the MADS-box domain transcription factors

SHATTERPROOF 1 (SHP1), *SHP2*, *SEEDSTICK (STK)* and *SEP3 (SEPALLATA 3)*. In particular, binding of *SEP3* and *STK* to a putative regulatory region of the *VDD* promoter has been demonstrated (Matias-Hernandez et al., 2010). Here, we have found that *AG* binds to the same region of the *VDD* promoter, and that *VDD* expression is reduced in an *ag* mutant background. Although no direct evidence is available yet, these results suggest that *AG* may be part of a protein complex with *SEP3* and *STK* that is involved in the regulation of *VDD*. Indeed, it was previously proposed that *SHP1*, *SHP2*, *STK* and *AG* form a transcriptional complex that regulates ovule development (Favaro et al., 2003). Therefore, *AG* function could be altered depending on the availability of other transcription factors in a particular tissue or cell type.

Our data also suggests that *AG* might be part of protein complexes involving transcription factors of the basic Leucine Zipper Domain (bZIP) or basic Helix-Loop-Helix (bHLH) class. It has been proposed that these 2 classes of transcription factors bind to a conserved motif, termed the G-box (Siberil et al., 2001). Analysis of the neighbouring genomic regions bound by *AG* showed an enrichment in the G-box motif. This observation was also made by Kaufmann et al. (2010) when analyzing the vicinity of *SEP3* binding sites (Kaufmann et al., 2009). Altogether, this data suggests that *AG* and *SEP3*, which function together in transcriptional complexes, may bind together with other non-MADS box transcription factors to activate and repress genes involved in flower development.

In summary, by identifying genes that are directly regulated by *AG*, we have furthered our understanding of the *AG*-dependent gene regulatory networks underlying the specification and the development of *Arabidopsis* reproductive organs. A similar strategy can now be adopted for other genes involved in flower development, such as for example genes that act downstream of the master regulators. Such studies would lay the foundations for the generation of a detailed gene regulatory network underlying the different processes required to form a flower.

Chapter 8 - References

- Aida, M., Ishida, T., Fukaki, H., Fujisawa, H. and Tasaka, M. (1997). Genes involved in organ separation in *Arabidopsis*: an analysis of the cup-shaped cotyledon mutant. *Plant Cell* **9**, 841-57.
- Allen, E., Xie, Z., Gustafson, A. M. and Carrington, J. C. (2005). microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell* **121**, 207-21.
- Alonso, J. M. and Ecker, J. R. (2006). Moving forward in reverse: genetic technologies to enable genome-wide phenomic screens in *Arabidopsis*. *Nat Rev Genet* **7**, 524-36.
- Alonso, J. M., Stepanova, A. N., Leisse, T. J., Kim, C. J., Chen, H., Shinn, P., Stevenson, D. K., Zimmerman, J., Barajas, P., Cheuk, R. et al. (2003). Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* **301**, 653-7.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403-10.
- Alvarez, J. and Smyth, D. R. (1999). CRABS CLAW and SPATULA, two *Arabidopsis* genes that control carpel development in parallel with AGAMOUS. *Development* **126**, 2377-86.
- Alvarez, J. P., Goldshmidt, A., Efroni, I., Bowman, J. L. and Eshed, Y. (2009). The NGATHA distal organ development genes are essential for style specification in *Arabidopsis*. *Plant Cell* **21**, 1373-93.
- Alvarez-Buylla, E. R., Garcia-Ponce, B. and Garay-Arroyo, A. (2006). Unique and redundant functional domains of APETALA1 and CAULIFLOWER, two recently duplicated *Arabidopsis thaliana* floral MADS-box genes. *J Exp Bot* **57**, 3099-107.
- Alvarez-Buylla, E. R., Pelaz, S., Liljegren, S. J., Gold, S. E., Burgeff, C., Ditta, G. S., Ribas de Pouplana, L., Martinez-Castilla, L. and Yanofsky, M. F. (2000). An ancestral MADS-box gene duplication occurred before the divergence of plants and animals. *Proc Natl Acad Sci U S A* **97**, 5328-33.
- Ausin, I., Alonso-Blanco, C. and Martinez-Zapater, J. M. (2005). Environmental regulation of flowering. *Int J Dev Biol* **49**, 689-705.
- Bailey, T. L. (2011). DREME: motif discovery in transcription factor ChIP-seq data. *Bioinformatics* **27**, 1653-9.
- Benjamini, Y. and Hochberg, Y. (1995). Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B-Methodological* **57**, 289-300.
- Birnbaum, K., Shasha, D. E., Wang, J. Y., Jung, J. W., Lambert, G. M., Galbraith, D. W. and Benfey, P. N. (2003). A gene expression map of the *Arabidopsis* root. *Science* **302**, 1956-60.
- Bomblies, K., Lempe, J., Eppe, P., Warthmann, N., Lanz, C., Dangl, J. L. and Weigel, D. (2007). Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants. *PLoS Biol* **5**, e236.
- Bourc'his, D. and Bestor, T. H. (2002). Helicase homologues maintain cytosine methylation in plants and mammals. *Bioessays* **24**, 297-9.
- Bowman, J. L., Alvarez, J., Weigel, D., Meyerowitz, E. M. and Smyth, D. R. (1993). Control of flower development in *Arabidopsis thaliana* by APETALA1 and interacting genes. *Development* **119**, 721-743.
- Bowman, J. L., Drews, G. N. and Meyerowitz, E. M. (1991a). Expression of the *Arabidopsis* floral homeotic gene AGAMOUS is restricted to specific cell types late in flower development. *Plant Cell* **3**, 749-58.

- Bowman, J. L., Sakai, H., Jack, T., Weigel, D., Mayer, U. and Meyerowitz, E. M.** (1992). SUPERMAN, a regulator of floral homeotic genes in Arabidopsis. *Development* **114**, 599-615.
- Bowman, J. L. and Smyth, D. R.** (1999). CRABS CLAW, a gene that regulates carpel and nectary development in Arabidopsis, encodes a novel protein with zinc finger and helix-loop-helix domains. *Development* **126**, 2387-96.
- Bowman, J. L., Smyth, D. R. and Meyerowitz, E. M.** (1989). Genes directing flower development in Arabidopsis. *Plant Cell* **1**, 37-52.
- Bowman, J. L., Smyth, D. R. and Meyerowitz, E. M.** (1991b). Genetic interactions among floral homeotic genes of Arabidopsis. *Development* **112**, 1-20.
- Breuil-Broyer, S., Morel, P., de Almeida-Engler, J., Coustham, V., Negrutiu, I. and Trehin, C.** (2004). High-resolution boundary analysis during Arabidopsis thaliana flower development. *Plant J* **38**, 182-92.
- Cairns, J., Spyrou, C., Stark, R., Smith, M. L., Lynch, A. G. and Tavare, S.** (2011). BayesPeak--an R package for analysing ChIP-seq data. *Bioinformatics* **27**, 713-4.
- Carlsbecker, A., Lee, J. Y., Roberts, C. J., Dettmer, J., Lehesranta, S., Zhou, J., Lindgren, O., Moreno-Risueno, M. A., Vaten, A., Thitamadee, S. et al.** (2010). Cell signalling by microRNA165/6 directs gene dose-dependent root cell fate. *Nature* **465**, 316-21.
- Castillejo, C., Romera-Branchat, M. and Pelaz, S.** (2005). A new role of the Arabidopsis SEPALLATA3 gene revealed by its constitutive expression. *Plant J* **43**, 586-96.
- Cecchetti, V., Altamura, M. M., Falasca, G., Costantino, P. and Cardarelli, M.** (2008). Auxin regulates Arabidopsis anther dehiscence, pollen maturation, and filament elongation. *Plant Cell* **20**, 1760-74.
- Chen, K. and Rajewsky, N.** (2007). The evolution of gene regulation by transcription factors and microRNAs. *Nat Rev Genet* **8**, 93-103.
- Chen, X.** (2004). A microRNA as a translational repressor of APETALA2 in Arabidopsis flower development. *Science* **303**, 2022-5.
- Chen, X.** (2005). MicroRNA biogenesis and function in plants. *FEBS Lett* **579**, 5923-31.
- Cheng, H., Qin, L., Lee, S., Fu, X., Richards, D. E., Cao, D., Luo, D., Harberd, N. P. and Peng, J.** (2004). Gibberellin regulates Arabidopsis floral development via suppression of DELLA protein function. *Development* **131**, 1055-64.
- Cheng, H., Song, S., Xiao, L., Soo, H. M., Cheng, Z., Xie, D. and Peng, J.** (2009). Gibberellin acts through jasmonate to control the expression of MYB21, MYB24, and MYB57 to promote stamen filament growth in Arabidopsis. *PLoS Genet* **5**, e1000440.
- Cheng, Y., Dai, X. and Zhao, Y.** (2006). Auxin biosynthesis by the YUCCA flavin monooxygenases controls the formation of floral organs and vascular tissues in Arabidopsis. *Genes Dev* **20**, 1790-9.
- Clark, S. E., Williams, R. W. and Meyerowitz, E. M.** (1997). The CLAVATA1 gene encodes a putative receptor kinase that controls shoot and floral meristem size in Arabidopsis. *Cell* **89**, 575-85.
- Coen, E. S. and Meyerowitz, E. M.** (1991). The war of the whorls: genetic interactions controlling flower development. *Nature* **353**, 31-7.
- Cole, M., Chandler, J., Weijers, D., Jacobs, B., Comelli, P. and Werr, W.** (2009). DORNROSCHEN is a direct target of the auxin response factor MONOPTEROS in the Arabidopsis embryo. *Development* **136**, 1643-51.
- Colombo, M., Brambilla, V., Marcheselli, R., Caporali, E., Kater, M. M. and Colombo, L.** (2010). A new role for the SHATTERPROOF genes during Arabidopsis gynoecium development. *Dev Biol* **337**, 294-302.

- Craft, J., Samalova, M., Baroux, C., Townley, H., Martinez, A., Jepson, I., Tsiantis, M. and Moore, I.** (2005). New pOp/LhG4 vectors for stringent glucocorticoid-dependent transgene expression in Arabidopsis. *Plant J* **41**, 899-918.
- Czechowski, T., Stitt, M., Altmann, T., Udvardi, M. K. and Scheible, W. R.** (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant Physiol* **139**, 5-17.
- De Veylder, L., Beeckman, T. and Inze, D.** (2007). The ins and outs of the plant cell cycle. *Nat Rev Mol Cell Biol* **8**, 655-65.
- Deveaux, Y., Peaucelle, A., Roberts, G. R., Coen, E., Simon, R., Mizukami, Y., Traas, J., Murray, J. A., Doonan, J. H. and Laufs, P.** (2003). The ethanol switch: a tool for tissue-specific gene induction during plant development. *Plant J* **36**, 918-30.
- Dillon, S. C., Zhang, X., Trievel, R. C. and Cheng, X.** (2005). The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biol* **6**, 227.
- Dinneny, J. R., Weigel, D. and Yanofsky, M. F.** (2006). NUBBIN and JAGGED define stamen and carpel shape in Arabidopsis. *Development* **133**, 1645-55.
- Ditta, G., Pinyopich, A., Robles, P., Pelaz, S. and Yanofsky, M. F.** (2004). The SEP4 gene of Arabidopsis thaliana functions in floral organ and meristem identity. *Curr Biol* **14**, 1935-40.
- Drews, G. N., Bowman, J. L. and Meyerowitz, E. M.** (1991). Negative regulation of the Arabidopsis homeotic gene AGAMOUS by the APETALA2 product. *Cell* **65**, 991-1002.
- Durinck, S., Bullard, J., Spellman, P. T. and Dudoit, S.** (2009). GenomeGraphs: integrated genomic data visualization with R. *BMC Bioinformatics* **10**, 2.
- Edwards, K., Johnstone, C. and Thompson, C.** (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Res* **19**, 1349.
- Ehrenreich, I. M. and Purugganan, M. D.** (2008). Sequence variation of MicroRNAs and their binding sites in Arabidopsis. *Plant Physiol* **146**, 1974-82.
- Eshed, Y., Baum, S. F. and Bowman, J. L.** (1999). Distinct mechanisms promote polarity establishment in carpels of Arabidopsis. *Cell* **99**, 199-209.
- Eshed, Y., Baum, S. F., Perca, J. V. and Bowman, J. L.** (2001). Establishment of polarity in lateral organs of plants. *Curr Biol* **11**, 1251-60.
- Fahlgren, N., Howell, M. D., Kasschau, K. D., Chapman, E. J., Sullivan, C. M., Cumbie, J. S., Givan, S. A., Law, T. F., Grant, S. R., Dangl, J. L. et al.** (2007). High-throughput sequencing of Arabidopsis microRNAs: evidence for frequent birth and death of MIRNA genes. *PLoS One* **2**, e219.
- Favaro, R., Pinyopich, A., Battaglia, R., Kooiker, M., Borghi, L., Ditta, G., Yanofsky, M. F., Kater, M. M. and Colombo, L.** (2003). MADS-box protein complexes control carpel and ovule development in Arabidopsis. *Plant Cell* **15**, 2603-11.
- Feng, S., Jacobsen, S. E. and Reik, W.** (2010). Epigenetic reprogramming in plant and animal development. *Science* **330**, 622-7.
- Ferrandiz, C., Gu, Q., Martienssen, R. and Yanofsky, M. F.** (2000). Redundant regulation of meristem identity and plant architecture by FRUITFULL, APETALA1 and CAULIFLOWER. *Development* **127**, 725-34.
- Ferrandiz, C., Pelaz, S. and Yanofsky, M. F.** (1999). Control of carpel and fruit development in Arabidopsis. *Annu Rev Biochem* **68**, 321-54.
- Fletcher, J. C., Brand, U., Running, M. P., Simon, R. and Meyerowitz, E. M.** (1999). Signaling of cell fate decisions by CLAVATA3 in Arabidopsis shoot meristems. *Science* **283**, 1911-4.
- Glaser, P., Kunst, F., Arnaud, M., Coudart, M. P., Gonzales, W., Hullo, M. F., Ionescu, M., Lubochinsky, B., Marcelino, L., Moszer, I. et al.** (1993). Bacillus subtilis genome

project: cloning and sequencing of the 97 kb region from 325 degrees to 333 degrees. *Mol Microbiol* **10**, 371-84.

Gomez-Mena, C., de Folter, S., Costa, M. M., Angenent, G. C. and Sablowski, R. (2005). Transcriptional program controlled by the floral homeotic gene AGAMOUS during early organogenesis. *Development* **132**, 429-38.

Gregis, V., Sessa, A., Colombo, L. and Kater, M. M. (2006). AGL24, SHORT VEGETATIVE PHASE, and APETALA1 redundantly control AGAMOUS during early stages of flower development in Arabidopsis. *Plant Cell* **18**, 1373-82.

Gregory, B. D., Yazaki, J. and Ecker, J. R. (2008). Utilizing tiling microarrays for whole-genome analysis in plants. *Plant J* **53**, 636-44.

Gremski, K., Ditta, G. and Yanofsky, M. F. (2007). The HECATE genes regulate female reproductive tract development in Arabidopsis thaliana. *Development* **134**, 3593-601.

Gross-Hardt, R., Lenhard, M. and Laux, T. (2002). WUSCHEL signaling functions in interregional communication during Arabidopsis ovule development. *Genes Dev* **16**, 1129-38.

Guo, H., Li, L., Ye, H., Yu, X., Algreen, A. and Yin, Y. (2009). Three related receptor-like kinases are required for optimal cell elongation in Arabidopsis thaliana. *Proc Natl Acad Sci USA* **106**, 7648-53.

Guo, H. S., Fei, J. F., Xie, Q. and Chua, N. H. (2003). A chemical-regulated inducible RNAi system in plants. *Plant J* **34**, 383-92.

Gustafson-Brown, C., Savidge, B. and Yanofsky, M. F. (1994). Regulation of the arabidopsis floral homeotic gene APETALA1. *Cell* **76**, 131-43.

Haring, M., Offermann, S., Danker, T., Horst, I., Peterhansel, C. and Stam, M. (2007). Chromatin immunoprecipitation: optimization, quantitative analysis and data normalization. *Plant Methods* **3**, 11.

Havecker, E. R., Wallbridge, L. M., Hardcastle, T. J., Bush, M. S., Kelly, K. A., Dunn, R. M., Schwach, F., Doonan, J. H. and Baulcombe, D. C. (2010). The Arabidopsis RNA-directed DNA methylation argonautes functionally diverge based on their expression and interaction with target loci. *Plant Cell* **22**, 321-34.

Heisler, M. G., Atkinson, A., Bylstra, Y. H., Walsh, R. and Smyth, D. R. (2001). SPATULA, a gene that controls development of carpel margin tissues in Arabidopsis, encodes a bHLH protein. *Development* **128**, 1089-98.

Hiratsu, K., Matsui, K., Koyama, T. and Ohme-Takagi, M. (2003). Dominant repression of target genes by chimeric repressors that include the EAR motif, a repression domain, in Arabidopsis. *Plant J* **34**, 733-9.

Hiratsu, K., Ohta, M., Matsui, K. and Ohme-Takagi, M. (2002). The SUPERMAN protein is an active repressor whose carboxy-terminal repression domain is required for the development of normal flowers. *FEBS Lett* **514**, 351-4.

Hochberg, Y. and Benjamini, Y. (1990). More powerful procedures for multiple significance testing. *Stat Med* **9**, 811-8.

Honma, T. and Goto, K. (2001). Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. *Nature* **409**, 525-9.

Ikeda, M., Mitsuda, N. and Ohme-Takagi, M. (2009). Arabidopsis WUSCHEL is a bifunctional transcription factor that acts as a repressor in stem cell regulation and as an activator in floral patterning. *Plant Cell* **21**, 3493-505.

Inoue, H., Nojima, H. and Okayama, H. (1990). High efficiency transformation of Escherichia coli with plasmids. *Gene* **96**, 23-8.

Irish, V. F. and Sussex, I. M. (1990). Function of the apetala-1 gene during Arabidopsis floral development. *Plant Cell* **2**, 741-53.

- Ishida, T., Aida, M., Takada, S. and Tasaka, M.** (2000). Involvement of CUP-SHAPED COTYLEDON genes in gynoecium and ovule development in *Arabidopsis thaliana*. *Plant Cell Physiol* **41**, 60-7.
- Ito, T., Ng, K. H., Lim, T. S., Yu, H. and Meyerowitz, E. M.** (2007). The homeotic protein AGAMOUS controls late stamen development by regulating a jasmonate biosynthetic gene in *Arabidopsis*. *Plant Cell* **19**, 3516-29.
- Ito, T., Wellmer, F., Yu, H., Das, P., Ito, N., Alves-Ferreira, M., Riechmann, J. L. and Meyerowitz, E. M.** (2004). The homeotic protein AGAMOUS controls microsporogenesis by regulation of SPOROCTELESS. *Nature* **430**, 356-60.
- Jack, T.** (2004). Molecular and genetic mechanisms of floral control. *Plant Cell* **16 Suppl**, S1-17.
- Jackson, J. P., Lindroth, A. M., Cao, X. and Jacobsen, S. E.** (2002). Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* **416**, 556-60.
- Jiao, Y. and Meyerowitz, E. M.** (2010). Cell-type specific analysis of translating RNAs in developing flowers reveals new levels of control. *Mol Syst Biol* **6**, 419.
- Jofuku, K. D., den Boer, B. G., Van Montagu, M. and Okamoto, J. K.** (1994). Control of *Arabidopsis* flower and seed development by the homeotic gene APETALA2. *Plant Cell* **6**, 1211-25.
- Johnson, D. S., Mortazavi, A., Myers, R. M. and Wold, B.** (2007). Genome-wide mapping of in vivo protein-DNA interactions. *Science* **316**, 1497-502.
- Kaufmann, K., Muino, J. M., Jauregui, R., Airolidi, C. A., Smaczniak, C., Krajewski, P. and Angenent, G. C.** (2009). Target genes of the MADS transcription factor SEPALLATA3: integration of developmental and hormonal pathways in the *Arabidopsis* flower. *PLoS Biol* **7**, e1000090.
- Kaufmann, K., Muino, J. M., Osteras, M., Farinelli, L., Krajewski, P. and Angenent, G. C.** (2010a). Chromatin immunoprecipitation (ChIP) of plant transcription factors followed by sequencing (ChIP-SEQ) or hybridization to whole genome arrays (ChIP-CHIP). *Nat Protoc* **5**, 457-72.
- Kaufmann, K., Wellmer, F., Muino, J. M., Ferrier, T., Wuest, S. E., Kumar, V., Serrano-Mislata, A., Madueno, F., Krajewski, P., Meyerowitz, E. M. et al.** (2010b). Orchestration of floral initiation by APETALA1. *Science* **328**, 85-9.
- Kelley, D. R., Skinner, D. J. and Gasser, C. S.** (2009). Roles of polarity determinants in ovule development. *Plant J* **57**, 1054-64.
- Kempin, S. A., Savidge, B. and Yanofsky, M. F.** (1995). Molecular basis of the cauliflower phenotype in *Arabidopsis*. *Science* **267**, 522-5.
- Kerk, N. M., Ceserani, T., Tausta, S. L., Sussex, I. M. and Nelson, T. M.** (2003). Laser capture microdissection of cells from plant tissues. *Plant Physiol* **132**, 27-35.
- Knizewski, L., Ginalska, K. and Jerzmanowski, A.** (2008). Snf2 proteins in plants: gene silencing and beyond. *Trends Plant Sci* **13**, 557-65.
- Krizek, B. A. and Fletcher, J. C.** (2005). Molecular mechanisms of flower development: an armchair guide. *Nat Rev Genet* **6**, 688-98.
- Krizek, B. A., Lewis, M. W. and Fletcher, J. C.** (2006). RABBIT EARS is a second-whorl repressor of AGAMOUS that maintains spatial boundaries in *Arabidopsis* flowers. *Plant J* **45**, 369-83.
- Krizek, B. A. and Meyerowitz, E. M.** (1996). The *Arabidopsis* homeotic genes APETALA3 and PISTILLATA are sufficient to provide the B class organ identity function. *Development* **122**, 11-22.

- Kuusk, S., Sohlberg, J. J., Long, J. A., Fridborg, I. and Sundberg, E.** (2002). STY1 and STY2 promote the formation of apical tissues during Arabidopsis gynoecium development. *Development* **129**, 4707-17.
- Laufs, P., Peaucelle, A., Morin, H. and Traas, J.** (2004). MicroRNA regulation of the CUC genes is required for boundary size control in Arabidopsis meristems. *Development* **131**, 4311-22.
- Laux, T., Mayer, K. F., Berger, J. and Jurgens, G.** (1996). The WUSCHEL gene is required for shoot and floral meristem integrity in Arabidopsis. *Development* **122**, 87-96.
- Law, J. A. and Jacobsen, S. E.** (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet* **11**, 204-20.
- Lee, T. I., Johnstone, S. E. and Young, R. A.** (2006). Chromatin immunoprecipitation and microarray-based analysis of protein location. *Nat Protoc* **1**, 729-48.
- Lenhard, M., Bohnert, A., Jurgens, G. and Laux, T.** (2001). Termination of stem cell maintenance in Arabidopsis floral meristems by interactions between WUSCHEL and AGAMOUS. *Cell* **105**, 805-14.
- Lenhard, M. and Laux, T.** (2003). Stem cell homeostasis in the Arabidopsis shoot meristem is regulated by intercellular movement of CLAVATA3 and its sequestration by CLAVATA1. *Development* **130**, 3163-73.
- Li, R., Yu, C., Li, Y., Lam, T. W., Yiu, S. M., Kristiansen, K. and Wang, J.** (2009a). SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* **25**, 1966-7.
- Li, R. Q., Yu, C., Li, Y. R., Lam, T. W., Yiu, S. M., Kristiansen, K. and Wang, J.** (2009b). SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* **25**, 1966-1967.
- Liu, C., Thong, Z. and Yu, H.** (2009). Coming into bloom: the specification of floral meristems. *Development* **136**, 3379-91.
- Liu, Y. G., Mitsukawa, N., Oosumi, T. and Whittier, R. F.** (1995). Efficient isolation and mapping of Arabidopsis thaliana T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant J* **8**, 457-63.
- Lloyd, A. M., Schena, M., Walbot, V. and Davis, R. W.** (1994). Epidermal cell fate determination in Arabidopsis: patterns defined by a steroid-inducible regulator. *Science* **266**, 436-9.
- Lohmann, J. U., Hong, R. L., Hobe, M., Busch, M. A., Parcy, F., Simon, R. and Weigel, D.** (2001). A molecular link between stem cell regulation and floral patterning in Arabidopsis. *Cell* **105**, 793-803.
- Lohmann, J. U. and Weigel, D.** (2002). Building beauty: the genetic control of floral patterning. *Dev Cell* **2**, 135-42.
- Long, J. A. and Barton, M. K.** (1998). The development of apical embryonic pattern in Arabidopsis. *Development* **125**, 3027-35.
- Mallory, A. C., Elmayan, T. and Vaucheret, H.** (2008). MicroRNA maturation and action--the expanding roles of ARGONAUTES. *Curr Opin Plant Biol* **11**, 560-6.
- Mandaokar, A., Thines, B., Shin, B., Lange, B. M., Choi, G., Koo, Y. J., Yoo, Y. J., Choi, Y. D., Choi, G. and Browse, J.** (2006). Transcriptional regulators of stamen development in Arabidopsis identified by transcriptional profiling. *Plant J* **46**, 984-1008.
- Mandel, M. A., Gustafson-Brown, C., Savidge, B. and Yanofsky, M. F.** (1992). Molecular characterization of the Arabidopsis floral homeotic gene APETALA1. *Nature* **360**, 273-7.
- Mandel, M. A. and Yanofsky, M. F.** (1995). A gene triggering flower formation in Arabidopsis. *Nature* **377**, 522-4.

- Mangelsdorf, D. J., Thummel, C., Beato, M., Herrlich, P., Schutz, G., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P. et al.** (1995). The nuclear receptor superfamily: the second decade. *Cell* **83**, 835-9.
- Massie, C. E. and Mills, I. G.** (2008). ChIPping away at gene regulation. *EMBO Rep* **9**, 337-43.
- Materna, S. C. and Oliveri, P.** (2008). A protocol for unraveling gene regulatory networks. *Nat Protoc* **3**, 1876-87.
- Matias-Hernandez, L., Battaglia, R., Galbiati, F., Rubes, M., Eichenberger, C., Grossniklaus, U., Kater, M. M. and Colombo, L.** (2010). VERDANDI is a direct target of the MADS domain ovule identity complex and affects embryo sac differentiation in Arabidopsis. *Plant Cell* **22**, 1702-15.
- Mayer, K. F., Schoof, H., Haecker, A., Lenhard, M., Jurgens, G. and Laux, T.** (1998). Role of WUSCHEL in regulating stem cell fate in the Arabidopsis shoot meristem. *Cell* **95**, 805-15.
- McBride, K. E. and Summerfelt, K. R.** (1990). Improved binary vectors for Agrobacterium-mediated plant transformation. *Plant Mol Biol* **14**, 269-76.
- McCormick, S.** (1993). Male Gametophyte Development. *Plant Cell* **5**, 1265-1275.
- Messenguy, F. and Dubois, E.** (2003). Role of MADS box proteins and their cofactors in combinatorial control of gene expression and cell development. *Gene* **316**, 1-21.
- Mi, S., Cai, T., Hu, Y., Chen, Y., Hodges, E., Ni, F., Wu, L., Li, S., Zhou, H., Long, C. et al.** (2008). Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5' terminal nucleotide. *Cell* **133**, 116-27.
- Mitsuda, N., Hiratsu, K., Todaka, D., Nakashima, K., Yamaguchi-Shinozaki, K. and Ohme-Takagi, M.** (2006). Efficient production of male and female sterile plants by expression of a chimeric repressor in Arabidopsis and rice. *Plant Biotechnol J* **4**, 325-32.
- Mizukami, Y. and Ma, H.** (1992). Ectopic expression of the floral homeotic gene AGAMOUS in transgenic Arabidopsis plants alters floral organ identity. *Cell* **71**, 119-31.
- Mizukami, Y. and Ma, H.** (1997). Determination of Arabidopsis floral meristem identity by AGAMOUS. *Plant Cell* **9**, 393-408.
- Mlotshwa, S., Yang, Z., Kim, Y. and Chen, X.** (2006). Floral patterning defects induced by Arabidopsis APETALA2 and microRNA172 expression in Nicotiana benthamiana. *Plant Mol Biol* **61**, 781-93.
- Moore, I., Galweiler, L., Grosskopf, D., Schell, J. and Palme, K.** (1998). A transcription activation system for regulated gene expression in transgenic plants. *Proc Natl Acad Sci U S A* **95**, 376-81.
- Moore, I., Samalova, M. and Kurup, S.** (2006). Transactivated and chemically inducible gene expression in plants. *Plant J* **45**, 651-83.
- Morohashi, K. and Grotewold, E.** (2009). A systems approach reveals regulatory circuitry for Arabidopsis trichome initiation by the GL3 and GL1 selectors. *PLoS Genet* **5**, e1000396.
- Muino, J. M., Hoogstraat, M., van Ham, R. C. and van Dijk, A. D.** (2011a). PRI-CAT: a web-tool for the analysis, storage and visualization of plant ChIP-seq experiments. *Nucleic Acids Res* **39**, W524-7.
- Muino, J. M., Kaufmann, K., van Ham, R. C., Angenent, G. C. and Krajewski, P.** (2011b). ChIP-seq Analysis in R (CSAR): An R package for the statistical detection of protein-bound genomic regions. *Plant Methods* **7**, 11.
- Murmu, J., Bush, M. J., DeLong, C., Li, S., Xu, M., Khan, M., Malcolmson, C., Fobert, P. R., Zachgo, S. and Hepworth, S. R.** (2010). Arabidopsis basic leucine-zipper transcription factors TGA9 and TGA10 interact with floral glutaredoxins ROXY1 and ROXY2 and are redundantly required for anther development. *Plant Physiol* **154**, 1492-504.

- Neff, M. M., Neff, J. D., Chory, J. and Pepper, A. E.** (1998). dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: experimental applications in *Arabidopsis thaliana* genetics. *Plant J* **14**, 387-92.
- Niu, Q. W., Lin, S. S., Reyes, J. L., Chen, K. C., Wu, H. W., Yeh, S. D. and Chua, N. H.** (2006). Expression of artificial microRNAs in transgenic *Arabidopsis thaliana* confers virus resistance. *Nat Biotechnol* **24**, 1420-8.
- Ohta, M., Matsui, K., Hiratsu, K., Shinshi, H. and Ohme-Takagi, M.** (2001). Repression domains of class II ERF transcriptional repressors share an essential motif for active repression. *Plant Cell* **13**, 1959-68.
- Olmedo-Monfil, V., Duran-Figueroa, N., Arteaga-Vazquez, M., Demesa-Arevalo, E., Autran, D., Grimanelli, D., Slotkin, R. K., Martienssen, R. A. and Vielle-Calzada, J. P.** (2010). Control of female gamete formation by a small RNA pathway in *Arabidopsis*. *Nature* **464**, 628-32.
- Ossowski, S., Schwab, R. and Weigel, D.** (2008). Gene silencing in plants using artificial microRNAs and other small RNAs. *Plant J* **53**, 674-90.
- Palatnik, J. F., Allen, E., Wu, X., Schommer, C., Schwab, R., Carrington, J. C. and Weigel, D.** (2003). Control of leaf morphogenesis by microRNAs. *Nature* **425**, 257-63.
- Parcy, F.** (2005). Flowering: a time for integration. *Int J Dev Biol* **49**, 585-93.
- Parenicova, L., de Folter, S., Kieffer, M., Horner, D. S., Favalli, C., Busscher, J., Cook, H. E., Ingram, R. M., Kater, M. M., Davies, B. et al.** (2003). Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in *Arabidopsis*: new openings to the MADS world. *Plant Cell* **15**, 1538-51.
- Park, P. J.** (2009). ChIP-seq: advantages and challenges of a maturing technology. *Nat Rev Genet* **10**, 669-80.
- Park, W., Zhai, J. and Lee, J. Y.** (2009). Highly efficient gene silencing using perfect complementary artificial miRNA targeting AP1 or heteromeric artificial miRNA targeting AP1 and CAL genes. *Plant Cell Rep* **28**, 469-80.
- Payne, T., Johnson, S. D. and Koltunow, A. M.** (2004). KNUCKLES (KNU) encodes a C2H2 zinc-finger protein that regulates development of basal pattern elements of the *Arabidopsis* gynoecium. *Development* **131**, 3737-49.
- Pekker, I., Alvarez, J. P. and Eshed, Y.** (2005). Auxin response factors mediate *Arabidopsis* organ asymmetry via modulation of KANADI activity. *Plant Cell* **17**, 2899-910.
- Pelaz, S., Tapia-Lopez, R., Alvarez-Buylla, E. R. and Yanofsky, M. F.** (2001). Conversion of leaves into petals in *Arabidopsis*. *Curr Biol* **11**, 182-4.
- Pillot, M., Baroux, C., Vazquez, M. A., Autran, D., Leblanc, O., Vielle-Calzada, J. P., Grossniklaus, U. and Grimanelli, D.** (2010). Embryo and endosperm inherit distinct chromatin and transcriptional states from the female gametes in *Arabidopsis*. *Plant Cell* **22**, 307-20.
- Pinyopich, A., Ditta, G. S., Savidge, B., Liljegren, S. J., Baumann, E., Wisman, E. and Yanofsky, M. F.** (2003). Assessing the redundancy of MADS-box genes during carpel and ovule development. *Nature* **424**, 85-8.
- Prunet, N., Morel, P., Thierry, A. M., Eshed, Y., Bowman, J. L., Negrutiu, I. and Trehin, C.** (2008). REBELOTE, SQUINT, and ULTRAPETALA1 function redundantly in the temporal regulation of floral meristem termination in *Arabidopsis thaliana*. *Plant Cell* **20**, 901-19.
- Raj, A. and van Oudenaarden, A.** (2008). Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* **135**, 216-26.

- Ramachandran, V. and Chen, X.** (2008). Small RNA metabolism in Arabidopsis. *Trends Plant Sci* **13**, 368-74.
- Rangwala, S. H. and Richards, E. J.** (2004). The value-added genome: building and maintaining genomic cytosine methylation landscapes. *Curr Opin Genet Dev* **14**, 686-91.
- Riechmann, J. L., Krizek, B. A. and Meyerowitz, E. M.** (1996). Dimerization specificity of Arabidopsis MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA, and AGAMOUS. *Proc Natl Acad Sci U S A* **93**, 4793-8.
- Riechmann, J. L. and Meyerowitz, E. M.** (1997). MADS domain proteins in plant development. *Biol Chem* **378**, 1079-101.
- Rieu, I., Eriksson, S., Powers, S. J., Gong, F., Griffiths, J., Woolley, L., Benlloch, R., Nilsson, O., Thomas, S. G., Hedden, P. et al.** (2008). Genetic analysis reveals that C19-GA 2-oxidation is a major gibberellin inactivation pathway in Arabidopsis. *Plant Cell* **20**, 2420-36.
- Ritchie, M. E., Silver, J., Oshlack, A., Holmes, M., Diyagama, D., Holloway, A. and Smyth, G. K.** (2007). A comparison of background correction methods for two-colour microarrays. *Bioinformatics* **23**, 2700-2707.
- Robertson, G., Hirst, M., Bainbridge, M., Bilenky, M., Zhao, Y., Zeng, T., Euskirchen, G., Bernier, B., Varhol, R., Delaney, A. et al.** (2007). Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat Methods* **4**, 651-7.
- Romanel, E. A., Schrago, C. G., Counago, R. M., Russo, C. A. and Alves-Ferreira, M.** (2009). Evolution of the B3 DNA binding superfamily: new insights into REM family gene diversification. *PLoS One* **4**, e5791.
- Roslan, H. A., Salter, M. G., Wood, C. D., White, M. R., Croft, K. P., Robson, F., Coupland, G., Doonan, J., Laufs, P., Tomsett, A. B. et al.** (2001). Characterization of the ethanol-inducible alc gene-expression system in Arabidopsis thaliana. *Plant J* **28**, 225-35.
- Sablowski, R.** (2009). Genes and functions controlled by floral organ identity genes. *Semin Cell Dev Biol*.
- Sablowski, R. W. and Meyerowitz, E. M.** (1998). A homolog of NO APICAL MERISTEM is an immediate target of the floral homeotic genes APETALA3/PISTILLATA. *Cell* **92**, 93-103.
- Sakai, H., Krizek, B. A., Jacobsen, S. E. and Meyerowitz, E. M.** (2000). Regulation of SUP expression identifies multiple regulators involved in arabidopsis floral meristem development. *Plant Cell* **12**, 1607-18.
- Sakai, H., Medrano, L. J. and Meyerowitz, E. M.** (1995). Role of SUPERMAN in maintaining Arabidopsis floral whorl boundaries. *Nature* **378**, 199-203.
- Savidge, B.** (1996). Floral meristem specification and floral organ development in Arabidopsis., vol. PhD (ed.: University of California at San Diego, La Jolla, CA).
- Savidge, B., Rounsley, S. D. and Yanofsky, M. F.** (1995). Temporal relationship between the transcription of two Arabidopsis MADS box genes and the floral organ identity genes. *Plant Cell* **7**, 721-33.
- Schiefthaler, U., Balasubramanian, S., Sieber, P., Chevalier, D., Wisman, E. and Schneitz, K.** (1999). Molecular analysis of NOZZLE, a gene involved in pattern formation and early sporogenesis during sex organ development in Arabidopsis thaliana. *Proc Natl Acad Sci U S A* **96**, 11664-9.
- Schoof, H., Lenhard, M., Haecker, A., Mayer, K. F., Jurgens, G. and Laux, T.** (2000). The stem cell population of Arabidopsis shoot meristems is maintained by a regulatory loop between the CLAVATA and WUSCHEL genes. *Cell* **100**, 635-44.

- Schultz, E. A. and Haughn, G. W.** (1991). LEAFY, a Homeotic Gene That Regulates Inflorescence Development in Arabidopsis. *Plant Cell* **3**, 771-781.
- Schwab, R., Ossowski, S., Riester, M., Warthmann, N. and Weigel, D.** (2006). Highly specific gene silencing by artificial microRNAs in Arabidopsis. *Plant Cell* **18**, 1121-33.
- Schwab, R., Palatnik, J. F., Riester, M., Schommer, C., Schmid, M. and Weigel, D.** (2005). Specific effects of microRNAs on the plant transcriptome. *Dev Cell* **8**, 517-27.
- Schwarz-Sommer, Z., Huijser, P., Nacken, W., Saedler, H. and Sommer, H.** (1990). Genetic Control of Flower Development by Homeotic Genes in *Antirrhinum majus*. *Science* **250**, 931-936.
- Sessions, A., Nemhauser, J. L., McColl, A., Roe, J. L., Feldmann, K. A. and Zambryski, P. C.** (1997). ETTIN patterns the Arabidopsis floral meristem and reproductive organs. *Development* **124**, 4481-91.
- Shiraishi, H., Okada, K. and Shimura, Y.** (1993). Nucleotide sequences recognized by the AGAMOUS MADS domain of Arabidopsis thaliana in vitro. *Plant J* **4**, 385-98.
- Shpak, E. D., Berthiaume, C. T., Hill, E. J. and Torii, K. U.** (2004). Synergistic interaction of three ERECTA-family receptor-like kinases controls Arabidopsis organ growth and flower development by promoting cell proliferation. *Development* **131**, 1491-501.
- Shpak, E. D., Lakeman, M. B. and Torii, K. U.** (2003). Dominant-negative receptor uncovers redundancy in the Arabidopsis ERECTA Leucine-rich repeat receptor-like kinase signaling pathway that regulates organ shape. *Plant Cell* **15**, 1095-110.
- Siberil, Y., Doireau, P. and Gantet, P.** (2001). Plant bZIP G-box binding factors. Modular structure and activation mechanisms. *Eur J Biochem* **268**, 5655-66.
- Sieburth, L. E., Drews, G. N. and Meyerowitz, E. M.** (1998). Non-autonomy of AGAMOUS function in flower development: use of a Cre/loxP method for mosaic analysis in Arabidopsis. *Development* **125**, 4303-12.
- Sieburth, L. E., Running, M. P. and Meyerowitz, E. M.** (1995). Genetic separation of third and fourth whorl functions of AGAMOUS. *Plant Cell* **7**, 1249-58.
- Silver, J. D., Ritchie, M. E. and Smyth, G. K.** (2009). Microarray background correction: maximum likelihood estimation for the normal-exponential convolution. *Biostatistics* **10**, 352-363.
- Silverstone, A. L., Jung, H. S., Dill, A., Kawaide, H., Kamiya, Y. and Sun, T. P.** (2001). Repressing a repressor: gibberellin-induced rapid reduction of the RGA protein in Arabidopsis. *Plant Cell* **13**, 1555-66.
- Skinner, D. J., Hill, T. A. and Gasser, C. S.** (2004). Regulation of ovule development. *Plant Cell* **16 Suppl**, S32-45.
- Slotkin, R. K., Vaughn, M., Borges, F., Tanurdzic, M., Becker, J. D., Feijo, J. A. and Martienssen, R. A.** (2009). Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. *Cell* **136**, 461-72.
- Smyth, D. R., Bowman, J. L. and Meyerowitz, E. M.** (1990). Early flower development in Arabidopsis. *Plant Cell* **2**, 755-67.
- Smyth, G. K.** (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* **3**, Article3.
- Smyth, G. K.** (2005). Individual channel analysis of two-colour microarrays. In: *55th Session of the International Statistics Institute, 5-12 April 2005, Sydney Convention & Exhibition Centre, Sydney, Australia (CD)*, 116.
- Smyth, G. K. and Speed, T.** (2003). Normalization of cDNA microarray data. *Methods* **31**, 265-273.

- Sohlberg, J. J., Myrenas, M., Kuusk, S., Lagercrantz, U., Kowalczyk, M., Sandberg, G. and Sundberg, E.** (2006). STY1 regulates auxin homeostasis and affects apical-basal patterning of the Arabidopsis gynoecium. *Plant J* **47**, 112-23.
- Spyrou, C., Stark, R., Lynch, A. G. and Tavare, S.** (2009). BayesPeak: Bayesian analysis of ChIP-seq data. *BMC Bioinformatics* **10**, 299.
- Sridhar, V. V., Surendrarao, A. and Liu, Z.** (2006). APETALA1 and SEPALLATA3 interact with SEUSS to mediate transcription repression during flower development. *Development* **133**, 3159-66.
- Stahle, M. I., Kuehlich, J., Staron, L., von Arnim, A. G. and Golz, J. F.** (2009). YABBYs and the Transcriptional Corepressors LEUNIG and LEUNIG_HOMOLOG Maintain Leaf Polarity and Meristem Activity in Arabidopsis. *Plant Cell*, 3105-18.
- Stintzi, A. and Browse, J.** (2000). The Arabidopsis male-sterile mutant, opr3, lacks the 12-oxophytodienoic acid reductase required for jasmonate synthesis. *Proc Natl Acad Sci U S A* **97**, 10625-30.
- Sun, B., Xu, Y., Ng, K. H. and Ito, T.** (2009). A timing mechanism for stem cell maintenance and differentiation in the Arabidopsis floral meristem. *Genes Dev* **23**, 1791-804.
- Sundaresan, V., Springer, P., Volpe, T., Haward, S., Jones, J. D., Dean, C., Ma, H. and Martienssen, R.** (1995). Patterns of gene action in plant development revealed by enhancer trap and gene trap transposable elements. *Genes Dev* **9**, 1797-810.
- Sundstrom, J. F., Nakayama, N., Glimelius, K. and Irish, V. F.** (2006). Direct regulation of the floral homeotic APETALA1 gene by APETALA3 and PISTILLATA in Arabidopsis. *Plant J* **46**, 593-600.
- Takeda, S., Matsumoto, N. and Okada, K.** (2004). RABBIT EARS, encoding a SUPERMAN-like zinc finger protein, regulates petal development in Arabidopsis thaliana. *Development* **131**, 425-34.
- Thorstensen, T., Grini, P. E. and Aalen, R. B.** (2011). SET domain proteins in plant development. *Biochim Biophys Acta* **1809**, 407-20.
- Triezenberg, S. J., Kingsbury, R. C. and McKnight, S. L.** (1988). Functional dissection of VP16, the trans-activator of herpes simplex virus immediate early gene expression. *Genes Dev* **2**, 718-29.
- Trigueros, M., Navarrete-Gomez, M., Sato, S., Christensen, S. K., Pelaz, S., Weigel, D., Yanofsky, M. F. and Ferrandiz, C.** (2009). The NGATHA genes direct style development in the Arabidopsis gynoecium. *Plant Cell* **21**, 1394-409.
- Urbanus, S. L., Martinelli, A. P., Dinh, Q. D., Aizza, L. C., Dornelas, M. C., Angenent, G. C. and Immink, R. G.** (2010). Intercellular transport of epidermis-expressed MADS domain transcription factors and their effect on plant morphology and floral transition. *Plant J* **63**, 60-72.
- Valoczi, A., Varallyay, E., Kauppinen, S., Burgyan, J. and Havelda, Z.** (2006). Spatio-temporal accumulation of microRNAs is highly coordinated in developing plant tissues. *Plant J* **47**, 140-51.
- Vaucheret, H.** (2008). Plant ARGONAUTES. *Trends Plant Sci* **13**, 350-8.
- Villanueva, J. M., Broadhvest, J., Hauser, B. A., Meister, R. J., Schneitz, K. and Gasser, C. S.** (1999). INNER NO OUTER regulates abaxial- adaxial patterning in Arabidopsis ovules. *Genes Dev* **13**, 3160-9.
- Voinnet, O.** (2008). Use, tolerance and avoidance of amplified RNA silencing by plants. *Trends Plant Sci* **13**, 317-28.

- Wang, H., Tang, W., Zhu, C. and Perry, S. E.** (2002). A chromatin immunoprecipitation (ChIP) approach to isolate genes regulated by AGL15, a MADS domain protein that preferentially accumulates in embryos. *Plant J* **32**, 831-43.
- Warnes, G., Bolker, B. and Lumley, T.** gplots: Various R programming tools for plotting data.
- Weigel, D., Alvarez, J., Smyth, D. R., Yanofsky, M. F. and Meyerowitz, E. M.** (1992). LEAFY controls floral meristem identity in Arabidopsis. *Cell* **69**, 843-59.
- Wellmer, F., Alves-Ferreira, M., Dubois, A., Riechmann, J. L. and Meyerowitz, E. M.** (2006). Genome-wide analysis of gene expression during early Arabidopsis flower development. *PLoS Genet* **2**, e117.
- Wellmer, F. and Riechmann, J. L.** (2005). Gene network analysis in plant development by genomic technologies. *Int J Dev Biol* **49**, 745-59.
- Wellmer, F., Riechmann, J. L., Alves-Ferreira, M. and Meyerowitz, E. M.** (2004). Genome-wide analysis of spatial gene expression in Arabidopsis flowers. *Plant Cell* **16**, 1314-26.
- Wesley, S. V., Helliwell, C. A., Smith, N. A., Wang, M. B., Rouse, D. T., Liu, Q., Gooding, P. S., Singh, S. P., Abbott, D., Stoutjesdijk, P. A. et al.** (2001). Construct design for efficient, effective and high-throughput gene silencing in plants. *Plant J* **27**, 581-90.
- Wielopolska, A., Townley, H., Moore, I., Waterhouse, P. and Helliwell, C.** (2005). A high-throughput inducible RNAi vector for plants. *Plant Biotechnol J* **3**, 583-90.
- Wollmann, H., Mica, E., Todesco, M., Long, J. A. and Weigel, D.** (2010). On reconciling the interactions between APETALA2, miR172 and AGAMOUS with the ABC model of flower development. *Development* **137**, 3633-42.
- Wu, D., Lim, E., Vaillant, F., Asselin-Labat, M. L., Visvader, J. E. and Smyth, G. K.** (2010). ROAST: rotation gene set tests for complex microarray experiments. *Bioinformatics* **26**, 2176-82.
- Wuest, S. E.** (2011). A Functional Genomics Approach to Study Plant Reproductive Development. In *Smurfit Institute of Genetics*, vol. PhD (ed. Dublin: Trinity College Dublin).
- Yang, W. C., Ye, D., Xu, J. and Sundaresan, V.** (1999). The SPOROCTELESS gene of Arabidopsis is required for initiation of sporogenesis and encodes a novel nuclear protein. *Genes Dev* **13**, 2108-17.
- Yanofsky, M. F., Ma, H., Bowman, J. L., Drews, G. N., Feldmann, K. A. and Meyerowitz, E. M.** (1990). The protein encoded by the Arabidopsis homeotic gene agamous resembles transcription factors. *Nature* **346**, 35-9.
- Yu, H., Ito, T., Zhao, Y., Peng, J., Kumar, P. and Meyerowitz, E. M.** (2004). Floral homeotic genes are targets of gibberellin signaling in flower development. *Proc Natl Acad Sci USA* **101**, 7827-32.
- Zhang, X., Henriques, R., Lin, S. S., Niu, Q. W. and Chua, N. H.** (2006). Agrobacterium-mediated transformation of Arabidopsis thaliana using the floral dip method. *Nat Protoc* **1**, 641-6.
- Zhu, L. J., Gazin, C., Lawson, N. D., Pages, H., Lin, S. M., Lapointe, D. S. and Green, M. R.** (2010). ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinformatics* **11**, 237.
- Znaidi, S., Barker, K. S., Weber, S., Alarco, A. M., Liu, T. T., Boucher, G., Rogers, P. D. and Raymond, M.** (2009). Identification of the Candida albicans Cap1p regulon. *Eukaryot Cell* **8**, 806-20.

Appendix 1

Primer	Sequence 5' - 3'
DM-1	gaTTTTTATTCACTCCGGGCGGTtctctctttgtattcc
DM-2	gaACGGCCCCGAGTGAATAAAAAAtcaaagagaatcaatga
DM-3	gaACAGCCCCGAGTGTATAAAATtcacaggtcgtgatatg
DM-4	gaATTTTATACACTCCGGGCTGTtctacatatattcct
DM-5	gaTTTACACTAACTGGAGAGCGCtctctctttgtattcc
DM-6	gaGCGCTCTCCAGTTAGTGTAATcaaagagaatcaatga
DM-7	gaGCACTCTCCAGTTTGTGTAATtcacaggtcgtgatatg
DM-8	gaATTACACAACTGGAGAGTGCTtctacatatattcct
DM-9	gaATAGCCGGGAGTGTATAAAATtcacaggtcgtgatatg
DM-10	gaATTTTATACACTCCCGGCTATtctacatatattcct
DM-11	gaTTTTTATTCACTCCCGGCGATtctctctttgtattcc
DM-12	gaATCGCCGGGAGTGAATAAAAAAtcaaagagaatcaatga
DM-13	gaTTTATCGTTATGCAAATCAGCtctctctttgtattcc
DM-14	gaGCTGATTTGCATAACGATAAAAtcaaagagaatcaatga
DM-15	gaGCCGATTTGCATATCGATAATtcacaggtcgtgatatg
DM-16	gaATTATCGATATGCAAATCGGCtctacatatattcct
DM-17	gaTTGTTGATAATACTGTGCGTTtctctctttgtattcc
DM-18	gaAACGCACAGTATTATCAACAAtcaaagagaatcaatga
DM-19	gaAAAGCACAGTATTTTCAACATtcacaggtcgtgatatg
DM-20	gaATGTTGAAAATACTGTGCTTTtctacatatattcct
DM-21	gaTTCGTAGTTAGATCCTGCCGGTctctctttgtattcc
DM-22	gaCCGGCAGGATCTAACTACGAAtcaaagagaatcaatga
DM-23	gaCCAGCAGGATCTATCTACGATtcacaggtcgtgatatg
DM-24	gaATCGTAGATAGATCCTGCTGGtctacatatattcct
DM-25	gaTTAGCTCCGATTGGTAGGCGGtctctctttgtattcc
DM-26	gaCCGCCTACCAATCCGAGCTAAAtcaaagagaatcaatga
DM-27	gaCCACCTACCAATCCGAGCTATtcacaggtcgtgatatg
DM-28	gaATAGCTCGGATTGGTAGGTGGtctacatatattcct
DM-29	CTGCAAGGCGATTAAGTTGGGTAAC
DM-30	GCGGATAACAATTTTACACAGGAAACAG
DM-31	GGTAAGGATCTGAGCTACACATGCTC
DM-32	GAAAAAGAAGACGTTCCAACCAC
DM-33	CCGCATAGCTGAACATCTATATAAC
DM-34	GGA CAA TTC TAA CAC CGG ATC
DM-35	GATATATTAATATATGTTGATAAATCACTTA
DM-36	ATAGAATTACCTTCTTGGATCGG
DM-37	GAATCGATAGCTGCCACTAACTGGAGAGCGGTTTGG
DM-38	GTGAAACAAATTTTCTGCGAGAATGTCACT
DM-39	CCGGTGTTAGAATTGTCCGATATTGC
DM-40	TATCGGACAATTCTAACACCGGATCG
DM-41	CAGAATTCGCCACTAACTGGAGAGCGGTTTGG
DM-42	AATGATTGATCGGAGC
DM-43	AAGCTCGAGAAGTATTACCCGAATCCGATC
DM-44	TAGAATTCGCGGTTTGGTCTTGGCGA
DM-45	CGTCTAGAGAGAAGTATTACCCGAATCCGATC

DM-46 TAGGATCCGCGGTTTGGTCTTGG
 DM-47 TACTCGAGATCACGGCGTACCAATCG
 DM-48 GAATCCCGGGCACTAACTGGAGAGCGGTTT
 DM-75 GCTCACTGAGTTCGAGT
 DM-76 GTTGCTTGTTGGAGCTAAG
 DM-77 atctcgagcgtggtag
 DM-78 ACATTCTAGATGCGGCGAAGCAGC
 DM-79 CTCTAGAGATCAAAAATGGGAAGGGGTAGGGTTC
 DM-80 CGGATCCTGCGGCGAAGCAGCCAAG
 DM-81 gaTAAAAGGGTACACATACGGGTtctctctttgtattcc
 DM-82 gaACCCGTATGTGTACCCTTTTAtcaaagagaatcaatga
 DM-83 gaACACGTATGTGTAGCCTTTTTtcacaggtcgtgatg
 DM-84 gaAAAAAGGCTACACATACGTGTtctacatatattcct
 DM-85 CCCGGGCCATGGGTATGAGTAAAGGAGAAGAAGAACTTTTCACT
 DM-86 GTCCCGCAATTATACATTTAATACGCG
 DM-87 CTCGTCATCACTCTCGCTAGG
 DM-88 CTCCTACCCGGTATCTCCTTCAAG
 DM-89 TACTGCAGATCACGGCGTACCAATCG
 DM-90 GAATCTCGAGACACTAACTGGAGAGCGGTTT
 DM-91 CTCTAGAGATCAAAAATGGGAAGGGGTAGGGTTC
 DM-92 CGGATCCTGCGGCGAAGCAGCCAAG
 DM-93 tgtatgcaagctagttagg
 DM-94 TGGCTTGCTGAATCCCTTTG
 DM-95 GCAGCTTCTTCTCTGTCTGC
 DM-96 AGCTCGAACATGAGATGGAAGAG
 DM-97 GCGACATCTATGATAGAGCGCCAC
 DM-98 GGCAGAACCGGTCAAACCTAAAAG
 DM-99 AGTACATTAAAAACGTCCGCAATGTG
 DM-100 TCCGACAAGATCCTTTGAGCAC
 DM-101 ATAAATATGGATTGCACGACGAGTC
 DM-102 ACCAGATTCTTCGTGCAAAGATAGCTG
 DM-103 AAGCTGCTCGTAGTTAGATCCTCCTG
 DM-104 AAACACACGCTCGGACGCATATTAC
 DM-105 GGCGACTCGGTATTTGGATGAATGAG
 DM-106 ATGGATCCGAGAAATGGGAAGGGGTAGGGTT
 DM-107 TAGGATCCAGCGGCGTAACAGCCAAGGTAATTG
 DM-108 ATACGCGTCGAGATGTGTCTGC
 DM-109 ATCTCGAGTAGCGGCGTAACAGCCAAGGTAATTG
 DM-110 TCCGACAAGATCCTTTGAGCAC
 DM-111 ATAAATATGGATTGCACGACGAGTC
 DM-112 GCAGCTTCTTCTCTGTCTGC
 DM-113 AGCTCGAACATGAGATGGAAGAG
 DM-114 TACTGCAGCTCAAAAATGGCGGAACCAACCAC
 DM-115 TTGGTACCTTATAAACGGAGAGAAAGGTCTAGATC
 DM-116 aagcaactgcagcttaaaccttccc
 DM-117 gccgagaattcgTAAACGGAGAGAAAGGTCTA
 DM-118 ATCtgagtggaataacaaagagacagaagc
 DM-119 TTGAATTCCCTAAACGGAGAGAAAGGTCTAGATC

DM-120	CATACGTATATCCATTTCGCATTATCCGGTAAC
DM-121	GAATGGATATACGTATGGATACGGATGAAACG
DM-122	ACTTCTGGAAACGAACCGCTTCAG
DM-123	CGGATGAAACGGATCGTAGCCATC
DM-124	GGAGCTCCCAGAAGCTCGAAAAACAAAGAAAAAAATC
DM-125	AAGCG^GTTGTGGAGAACATGATACG
DM-126	TGGAGAGCTTGATTTGCGAAATACCG
DM-127	ATCTGCG^AAAGGGTATCCAGTTGAC
DM-128	TGGAAGCCTCTGACTGATGGAGC

Appendix 2

CTAB buffer

100 mM Tris HCl (pH8.0)

20 mM EDTA

1.4 M NaCl

2% (w/v) CTAB (cetyltrimethyl ammonium bromide – nonionic detergent)

facultative 1% PVP 40,000 (polyvinyl pyrrolidone)

M1 buffer

10 mM NaPi pH7.0

100 mM NaCl

H₂O

Added fresh to M1 stock just before use

1 M hexylene glycol

10 mM b-ME

1x CPIC

M3 buffer

10 mM NaPi pH7.0

100 mM NaCl

Added fresh to M3 stock just before use

10 mM b-ME

1x Complete Protease Inhibitor cocktail

ChIP dilution buffer

16.7 mM Tris-HCl pH8.1

167 mM NaCl

1.2 mM EDTA

1.1% Triton X-100

0.01% SDS

1x CPIC was added fresh just before use

LiCl/Detergent Wash

10 mM Tris-HCl pH8.0

0.5% deoxycholic acid (w/v)

0.5% NP-40

1 mM EDTA

250 mM LiCl

Elution buffer 1

100 mM NaHCO₃

1% SDS

Elution Buffer 2

0.1 M glycine
0.5 M NaCl
0.05% Tween-20
pH 2.8.

MSTET buffer

5% Triton X-100
50mM EDTA pH 8.0
50mM Tris pH 8.0
5% Sucrose

Stacking gel

4x Stacking Buffer
40% Acrylamide:Bis (29:1)
Temed
APS 10%

4x Separating Buffer

1.5M Tris-Base
20% 0.4%
Adjusted to pH 8.8 with HCl

20x Transfer buffer (500 mL)

200 mM Tris base
2.0 M glycine

1x Transfer buffer (2L)

20x transfer buffer
Methanol
H₂O

100x agarose gel-loading buffer (10 mL)
Bromophenol blue
Xylene cyanol FF
Orange G
Glycerol
H₂O
Dilute in 50% glycerol to make 10x solution

MC Buffer

10mM NaPi pH7.0
50mM NaCl

0.1M Sucrose

M2 buffer

10 mM NaPi pH7.0

100 mM NaCl

10 mM MgCl₂

0.5% Triton X100

H₂O

Added fresh to M2 stock just before use

1 M hexylene glycol

10 mM b-ME

1x CPIC

Lysis buffer

50 mM Tris-HCl pH 8.1

10 mM EDTA

1% SDS

1x CPIC was added fresh just before use

Mixed Micelle Wash Buffer

20 mM Tris-HCl pH8.0

150 mM NaCl

5 mM EDTA

5 % sucrose (w/v)

1% Triton X100

0.2% SDS

Buffer 500

50 mM Hepes pH7.5

0.1% deoxycholic acid (w/v)

1 mM EDTA

1% Triton X100

IP Buffer

0.5M HEPES (pH 7.5)

4M NaCl

100mM ZnSO₄

1M MgCl₂

25% Triton-X (Sigma)

20% SDS (Sigma)

TES buffer

50mM Tris pH 8.0

25% Sucrose

2mM EDTA pH 8.0

Separating Gel

4x Separating Buffer
40% Acrylamide:Bis (29:1)
Temed
APS 10%

4x Stacking Buffer

0.5M Tris-Base
0.4% SDS
Adjusted to pH 6.8 with HCl

2x SDS loading buffer

25 ml 4× stacking buffer
20 ml glycerol
4% SDS
2 ml 2-ME
1 mg bromphenol blue
Add H₂O to 100 ml and mix

Ethidium bromide stock solution

Stock: 10 mg/ml in water; store at RT or 4°C and protect from light.
Working concentration in gel or buffer: 0.5 µg/ml

50x TAE buffer

50× stock solution:
242 g Tris base
57.1 ml glacial acetic acid
37.2 g Na₂EDTA · 2H₂O
H₂O to 1 liter
Autoclave
Working solution, pH ~8.5:
40 mM Tris · acetate
2 mM Na₂EDTA · 2H₂O

Appendix 3

Vector	Description
vDM-1	35S-pBJ36
vDM-2	<i>pAlcA</i> -pBJ36
vDM-3	pML-BART
vDM-4	35S:: <i>AlcR</i> -pML-BART
vDM-5	pHANNIBAL
vDM-6	35S- <i>SRDX</i> -pBJ36
vDM-7	35S- <i>WUSB</i> -pBJ36
vDM-8	pBJ36-6xHA
vDM-9	pBJ36- <i>mGFP5</i>
vDM-10	pART27
vDM-11	pBJ36-GR
vDM-12	pGREEN-GR
vDM-13	6xOP-pBJ36