Characterization of the *Arabidopsis* MADS-box transcription factor, *AGL104*

A thesis presented in partial fulfillment of the requirements for the degree of

Master of Science in Plant Biology

at Massey University, Palmerston North, New Zealand

Arti S. Reddy

2007

ABSTRACT

AGL104 is an *Arabidopsis* MADS-box transcription factor belonging to the MIKC* clade. The exclusive expression of MIKC* genes in the gametophyte generation of both mosses and angiosperms has fueled questions regarding the function of these genes in both these taxa and the notion that the developmental program of the gametophyte generation in both these taxa may be fundamentally similar even though the structures themselves differ greatly in their phenotype. Since transcription factors control development and changes in the developmental control genes is thought to be a major source of evolutionary changes in morphology, characterization of MIKC* genes is expected to provide clues to the evolutionary changes in land plant body form. In angiosperms, *AGL104* is reported to be expressed late and exclusively in the male (pollen) and female (embryo sac) gametophyte. Since late pollen development, such as pollen germination and pollen tube elongation, is thought to occur independently of transcription, the exclusive and high level of expression of a transcription factor is thus intriguing.

We report the expression of *AGL104* in developing anthers, mature pollen, pollen tubes and the egg apparatus of the embryo sac. Our study is the first report of *AGL104* expression in the pollen tubes. Our data showing spatial expression of *AGL104* in the different developmental stages of pollen, with weak expression in the uninucleate microspore that increases and culminates in the mature pollen, is also novel since spatial expression of this gene during pollen development had not been previously reported.

Functional characterization through gain-of-function and loss-of-function analyses shows that *AGL104* promotes pollen germination and an increased pollen tube length when measured 4 hours after pollination. The implication of this data is that, despite popular notions, active gene regulation is taking place during pollen germination and tube elongation. Further functional analysis in the pollen and the embryo sac is required to establish the precise role of *AGL104* in the angiosperms. This information will then lay the groundwork for future comparisons of MIKC* activity in the basal and higher plants and determine if changes in MIKC* gene function were responsible for evolutionary changes in land plant body form.

ACKNOWLEDGEMENTS

First and foremost I thank my supervisor, Dr Barbara Ambrose. Barbara, without your help this thesis would not be what it is today. Thank you for encouraging me through every step of my degree, both in the laboratory and out of it. My heartfelt gratitude for being my mentor and a friend when I just needed to talk. Thank you for all your advice, for your contagious enthusiasm and for helping me keep things in perspective. I am taking away a lot more than the knowledge and skills you have imparted to me. There are words of yours I will always remember. I am grateful for the amount of patience you had with me. I know I have pushed you to the limits. Thank you for being understanding and most of all for believing in me. You are the greatest supervisor I could ever have hoped for. Whenever I will look back on my experience as a Master student, your influence will be the first thing I will remember. **You are the best !!!**

I am also very grateful to my lab mates, Kalika, Emilio, Ryohei and Xiuwen. Emilio, my plants would not have survived had it not been for you. Kalika, thank you for being my "big brother" and teaching me things I would never have learnt otherwise. Ryohei, thank you very much for all the little favors you have done that has made a huge difference in the project. You are the best "resource-finder" I know. Xiuwen, I wish I had more time to learn all the knowledge and skills you brought to the lab. Without you guys, my experience would not have been as enjoyable.

My special thanks to all my friends who have helped me in various ways during the writing of this thesis. Bernie, without you I could not have made it through those long, stressful nights, especially towards the end of my stay in NZ. You are a true friend. A special thank you to Lulu for staying up with us through the nights. You are the cutest and most wonderful baby. Suzy, thank you for all the help you provided. I will miss you guys so much. I am also very grateful to Gregory Gallaway, Sachi, Sumit, Kim May, Jasmine, Kaufoou, Manisha and my brother Amrit for their contribution in making this thesis possible.

I also thank NZAID for providing me with the scholarship that made this degree possible. A special thank you to Professor Peter Lockhart, who had the patience and provided me with resources to enable me to complete this thesis. I hope I will be able to give back someday.

Abbreviations

µ	microns
µg	microgram(s)
µl	microlitre(s)
35S	35S promoter from the cauliflower mosaic virus, CaMV
A.tumefaciens	*Agrobacterium tumefaciens*
Amp	Ampicillin
ATP	Adenosine triphosphate
bp	base pairs
C.elegans	*Caenorhabditis elegans*, a nematode
cDNA	complementary deoxyribonucleic acid
d	day(s)
DNA	deoxyribonucleic acid
dNTP	deoxy-nucleotide-triphosphate
E.coli	*Escherichia coli*; a bacterium
EDTA	ethylene diamine tetra acetate
g	gram(s)
GUS	ß-glucuronidase
hap	hours after pollination
HCl	Hydrochloric acid
hr(s)	hour(s)
Kan	Kanamycin
kb	kilobase(s)
KOH	Potassium hydroxide
l	litre(s)
LB	Luria-Bertani; used as bacterial growth media
M	molar; moles per litre
mg	milligram(s)
milliQ	water purified using the Milli-Q Ultrapure System
min	minute(s)

ml	millilitre(s)
mM	millimolar
MOPS	3-(N-morpholino) propanesulfonic acid
MS	Murashige Skoog, media
NaCl	Sodium chloride
NaOH	Sodium hydroxide
ng	nanogram(s)
o/n	over night
°C	degrees Celsius
ORF	open reading frame
PCR	polymerase chain reaction
psi	unit for pressure
Rif	Rifampicin
RNA	ribonucleic acid
rpm	revolutions per minute
RT	reverse transcription/transcriptase
rt	room temperature
SC	sperm cell
sec	second(s)
Spec	Spectinomycin
TBE	Tris borate EDTA
T-DNA	transposon DNA
TE	Tris-EDTA buffer
Tris	tris(hydroxymethyl) aminomethane
U	enzyme units
UV	ultra violet
V	volts
VC	vegetative cell
VN	vegetative nuclei
w/v	weight/volume
wt	wildtype

X unit for buffer concentration

- Gene names (and loci) are italicised eg. *AGL104*
- Proteins are capitalised eg. AGL104

TABLE OF CONTENTS

Chapter 2: Materials and Method

Chapter 4: Gain-of-function analysis

Chapter 5: Loss-of-function analysis

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Background

Comparative morphological and traditional phylogenetic analyses have revealed the major patterns of land plant evolution (Kenrick and Crane, 1997). However, the mechanisms which generated the vast diversity of land plant body plans have remained elusive. The natural history of life is usually considered a succession of adult phenotypes, a theory called phylogeny or evolution (Theissen and Saedler, 1995). However, in multicellular organisms, their natural history is a succession of complete life cycles. To enable scientists to better understand the plethora of different plant and animal forms, a new approach called the "evo-devo" theory was proposed. Evolutionary developmental (thus, evo-devo) genetics makes the assumption that there is a close interrelationship between developmental and evolutionary processes (Gilbert et al., 1996). This assumption seems logical since even the most complex organisms are generated anew in each generation from a single fertilized egg cell, the zygote, through a complex process called development or ontogeny, which comprises pattern formation, differentiation, morphogenesis and growth. Given that development can put serious constraints on evolution, which could work either as a negative force by preventing advantageous alterations, or as a positive force by allowing preferred changes, development and evolution can indeed be regarded as interrelated processes (Theissen and Saedler, 1995). Thus, in the case of multicellular organisms, evolution of form is always the evolution of developmental processes (Theissen and Saedler, 1995).

1.2 Transcription factors and evolution

Development to a large extent is under genetic control and changes in the developmental control genes is thought to be a major aspect of evolutionary changes in morphology

(Gilbert et al., 1996; Theissen and Saedler, 1995). The key developmental control genes are members of multigene families which encode transcription factors. Transcription factors are proteins that bind to the regulatory regions in the genome and help control gene expression. They comprise a substantial fraction of all eukaryotic genomes and are usually grouped into different gene families according to the type of DNA-binding domains that they encode. In *Arabidopsis thaliana* (commonly called *Arabidopsis*), more than 5% of the genes encode transcription factors (Ratcliffe et al., 2000). Transcription factors form intricate networks, both through protein-protein interactions (among themselves and with proteins of other families) and at the transcriptional level.

1.3 *Arabidopsis* as a model plant

The past two decades has seen the emergence of *Arabidopsis* as one of the most widely used model organisms for the study of plant biology. As a member of the mustard family, *Arabidopsis* is closely related to many food plants such as canola, cabbage, cauliflower, broccoli, turnip, rutabaga, kale, brussel sprouts, kohlrabi, and radish. This species is small in stature, easy to grow, easy to manipulate genetically, has a short generation time, and a highly compact genome of about 130Mb with little interspersed repetitive DNA (Somerville and Somerville, 1999). Due to these factors, researchers sequenced the genome of *Arabidopsis* as a representative dicot model to enable further research in higher plants.

With the *Arabidopsis* genome now being fully sequenced, identification of and understanding the functions of its many genes, including transcription factors, is a step towards understanding the evolution and workings of higher plants as a whole. Recent years have seen an accumulation of a large amount of genetic and molecular data from a variety of plants (*Arabidopsis*, *Petunia*, *Antirrhinum*, Rice and Maize), thus providing resources for comparative evolutionary developmental biology.

1.4 Genome duplication, functional diversification and retention in plants

The existence of large gene families (including transcription factors), whose members have distinct but related functions, raised the possibility that gene duplication followed by functional diversification of the duplication products were an essential factor in the evolution of morphology (Svensson et al., 2000). This was confirmed through genomic studies that revealed that eukaryotes harbored large families of duplicate genes that had persisted through a long period of time (Lynch et al., 2001). Gene duplication was thus thought to contribute to the establishment of new gene functions (Long and Langley, 1993) with positive selection playing a role in preserving some gene copies, especially those that were advantageous (Messier and Stewart, 1997; Moore and Purugganan, 2003; Wagner et al., 1999). Since gene duplication was believed to provide a substrate for evolution, understanding the contribution of single-gene and whole-genome duplications and the fate of duplicates became fundamental to clarifying the link between phenotypic evolution and gene family diversification and evolution (Martinez-Castilla and Alvarez-Buylla, 2003).

Two models that aimed to explain the fate of duplicate genes were proposed:

- The classic "Ohno model" proposed that one of the duplicates was either lost (pseudogenization) or gained a new function (neofunctionalization) (Moore and Purugganan, 2005). However, this model failed to explain the predominance of duplicates with similar functions in the genome.

- A second model, known as "duplication-degeneration-complementation (DDC) model" or "subfunctionalization model" (Moore and Purugganan, 2005), suggested that one copy either decayed to a pseudogene (pseudogenization), acquired a new function (neofunctionalization), or both genes partitioned the function of their ancestral progenitor (subfunctionalization) (Force et al., 1999; Hughes, 1994; Lynch and Force, 2000).

Since deleterious mutations were more probable than advantageous mutations and the duplicates were likely to be functionally redundant upon duplication, models such as the ones described above presumed that most duplicate copies were lost and only a few neofunctional or subfunctional loci were maintained by selection (Force et al., 1999; Hughes, 1994; Lynch and Force, 2000). However, studies on expression divergence between duplicate genes that suggested that subfunctionalization was both frequent and rapid in nature (Blanc and Wolfe, 2004; Gu et al., 2002; Makova and Li, 2003).

In plants duplicate loci compose a large fraction of their genome, partly because of the frequent occurrence of genomic segmental duplications and polyploidization. In *Arabidopsis*, up to 90% of the loci are duplicated and it is estimated that 70-80% of the angiosperm species have undergone polyploidization at some point in their evolutionary history (Blanc et al., 2003; Bowers et al., 2003). The degree of expansion of transcription factor families in *Arabidopsis* is substantially higher than their expansion in other eukaryotes or expansion of other non-transcription factor families (Shiu et al., 2005). The observation that transcription factor duplicates in plants had a different rate of retention compared to other eukaryotes and persisted for longer periods in plant than in animal genomes implied that they play a more significant developmental role in plants than in their animal counterparts (Moore and Purugganan, 2005). Thus, the importance of understanding the phylogeny/evolution of the transcription factor genes was emphasized with the expectation that it would contribute significantly towards understanding the evolution of plant and animal form (Doebley and Lukens, 1998; Purugganan, 1998).

1.5 MADS-box transcription factors in plants

1.5.1 Phylogeny and origin

The MADS-box family of transcription factors has been one of the most extensively studied families of genes in plants. In *Arabidopsis*, as much as 107 MADS-box genes have been identified (De Bodt et al., 2003a; Kofuji et al., 2003; Nam et al., 2004; Parenicova et al., 2003). These genes bear a particular conserved DNA-binding domain

called the "MADS-box" named after the founding member proteins: MINICHROMOSOMAL MAINTENANCE 1 (MCM1) from yeast (Passmore et al., 1988), AGAMOUS (AG) from *Arabidopsis* (Yanofsky et al., 1990), DEFICIENS (DEF) from *Antirrhinum* (Schwarzsommer et al., 1992; Sommer et al., 1990), and SERUM RESPONSE FACTOR (SRF) from humans (Norman et al., 1988). The MADS domain is then followed by a C-terminal extension of varying length and structure.

Phylogenetic analysis of numerous *Arabidopsis* MADS-box transcription factor genes led to the proposal of two evolutionary lineages, represented by type I and type II genes (Alvarez-Buylla et al., 2000b; De Bodt et al., 2003a; Martinez-Castilla and Alvarez-Buylla, 2003; Verelst et al., 2007). The type I and type II lineages are thought to have arisen through duplication of an ancestral gene. Phylogenetic analyses recovered the type II genes as a strongly supported monophyletic lineage (Figure 1.1) (De Bodt et al., 2003b; Martinez-Castilla and Alvarez-Buylla, 2003; Parenicova et al., 2003). The monophyly of type I genes, on the other hand, was not well established as analyses by different groups gave different results (Becker and Theissen, 2003; Kofuji et al., 2003; Martinez-Castilla and Alvarez-Buylla, 2003; Parenicova et al., 2003). The two lineages differ in the amino acid sequence of their MADS-box and the domain structure of the protein such that the plant type I genes consist of one or two exons while type II genes always have more than five exons (Martinez-Castilla and Alvarez-Buylla, 2003). In addition, the type II genes possess a characteristic K-domain in their C terminus which is missing in the plant type I genes (De Bodt et al., 2003a; Martinez-Castilla and Alvarez-Buylla, 2003; Verelst et al., 2007).

Analysis of the distribution and origin of the *Arabidopsis* type I and type II genes showed that type I genes were concentrated on chromosome I and V, whereas type II genes were uniformly distributed among the five chromosomes (Martinez-Castilla and Alvarez-Buylla, 2003). Furthermore, the origin of the type I genes could be traced to intrachromosomal duplications whereas half of type II genes seemed to have originated from interchromosomal duplications (Martinez-Castilla and Alvarez-Buylla, 2003). Lynch and Conery (2000) discovered that recent duplications occurred more frequently

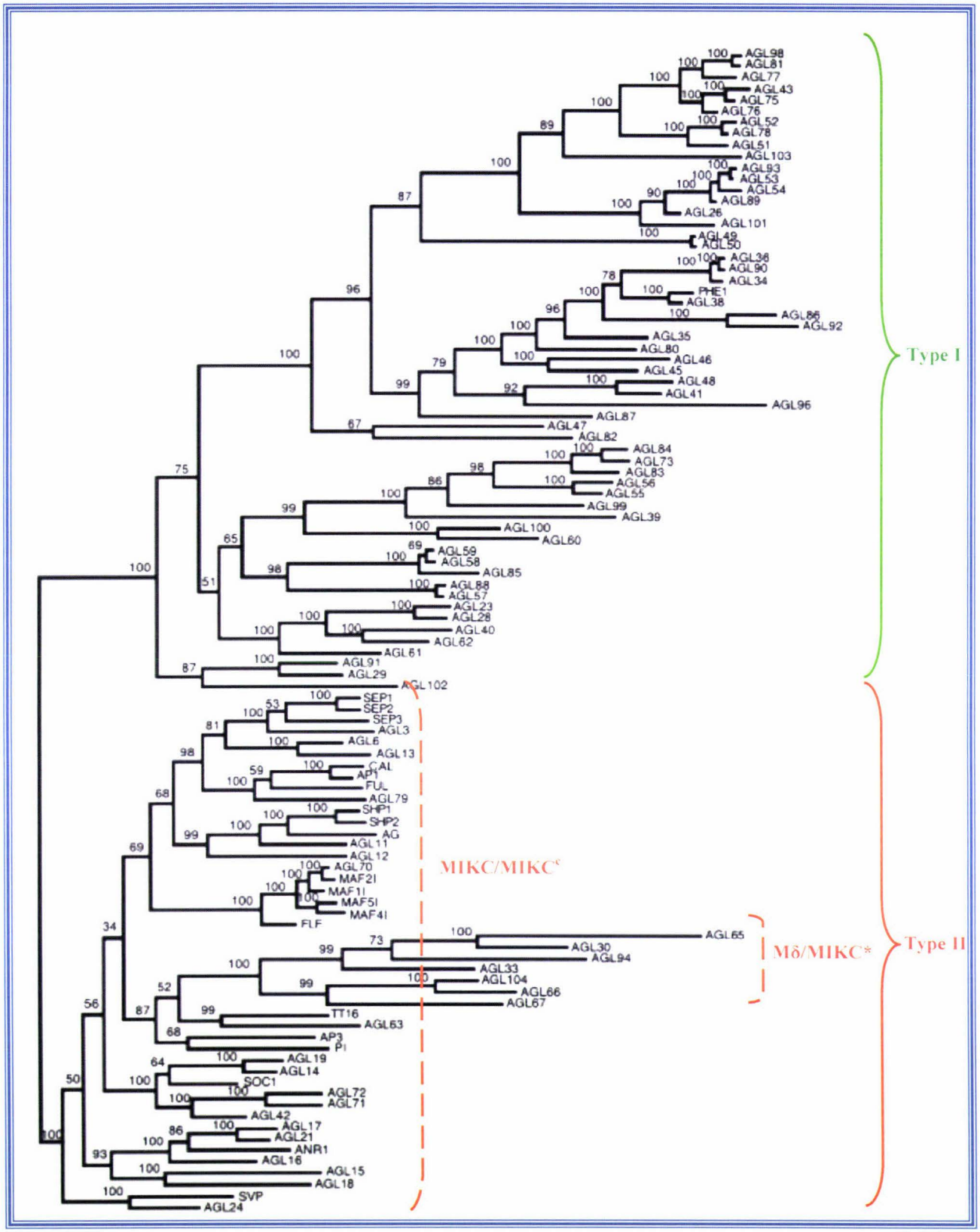

Figure 1.1: Phylogenetic tree of the *Arabidopsis* MADS-box genes (Modified from Martinez-Castilla and Alvarez-Buylla, (2003))

The MADS-box genes are divided into type I and type II lineages. The type II lineage has a strongly supported monophyletic origin whereas the monophyly of type I genes is not well established. Mδ and MIKC represent nomenclature used by Parenicova et al. (2003) whereas MIKC^c and MIKC* represent nomenclature used by Henschel et al. (2002).

within (intra) rather than between (inter) two chromosomes, which implied that duplications leading to type I genes would have occurred more recently than duplications leading to type II genes and thus the amount of difference among the sequences of type I genes was expected to be less than the amount of difference among the sequences of type II genes. Close scrutiny revealed that contrary to the above expectations, type I sequences were in fact more divergent among themselves in comparison to type II sequences (Lynch and Conery, 2000). Analyses by Nam et al. (2004) showed that type I genes had experienced a higher rate of birth-and-death evolution than type II genes in angiosperms. The higher birth rate of type I genes was believed to have been caused by a higher rate of segmental duplications and the higher death rate appeared to be partly due to a weaker purifying selection in type I genes, implying that the death of type I genes were less harmful to the organism (Nam et al., 2004). This explained the persistence of a larger number of duplicates from the type II lineage in the *Arabidopsis* genome (Martinez-Castilla and Alvarez-Buylla, 2003). It was proposed that after duplication the type II genes became functionally differentiated in a relatively short time and therefore were maintained as functional genes (Nam et al., 2004).

1.5.2 Structure of the MADS-box proteins

The structure of the different MADS-box proteins is extremely important to their function. Comparison of the MADS domain shows that in animals and fungi, the type I genes are represented by members of the *SRF* gene family whereas type II genes are represented by members of the *MYOCYTE ENHANCER FACTOR 2* (*MEF2*) gene family (Figure 1.2). In plants, the MADS domain of the type I genes also have a higher homology to the animal and fungi *SRF*-like genes and that of the type II genes are related to the *MEF2*-like genes (Alvarez-Buylla et al., 2000b). However, the structure of the MADS-box proteins starts to differ when the C-terminal domains between animals, fungi and plants are compared.

In animal and fungi type I genes, the C-terminal domains are well defined and are called SAM (acronym for *SRF*, *AG* and *MCM1*) (Shore and Sharrocks, 1995) whereas the C- terminal of their type II genes are called MEF2 (Mueller and Nordheim, 1991; Shore and Sharrocks, 1995).

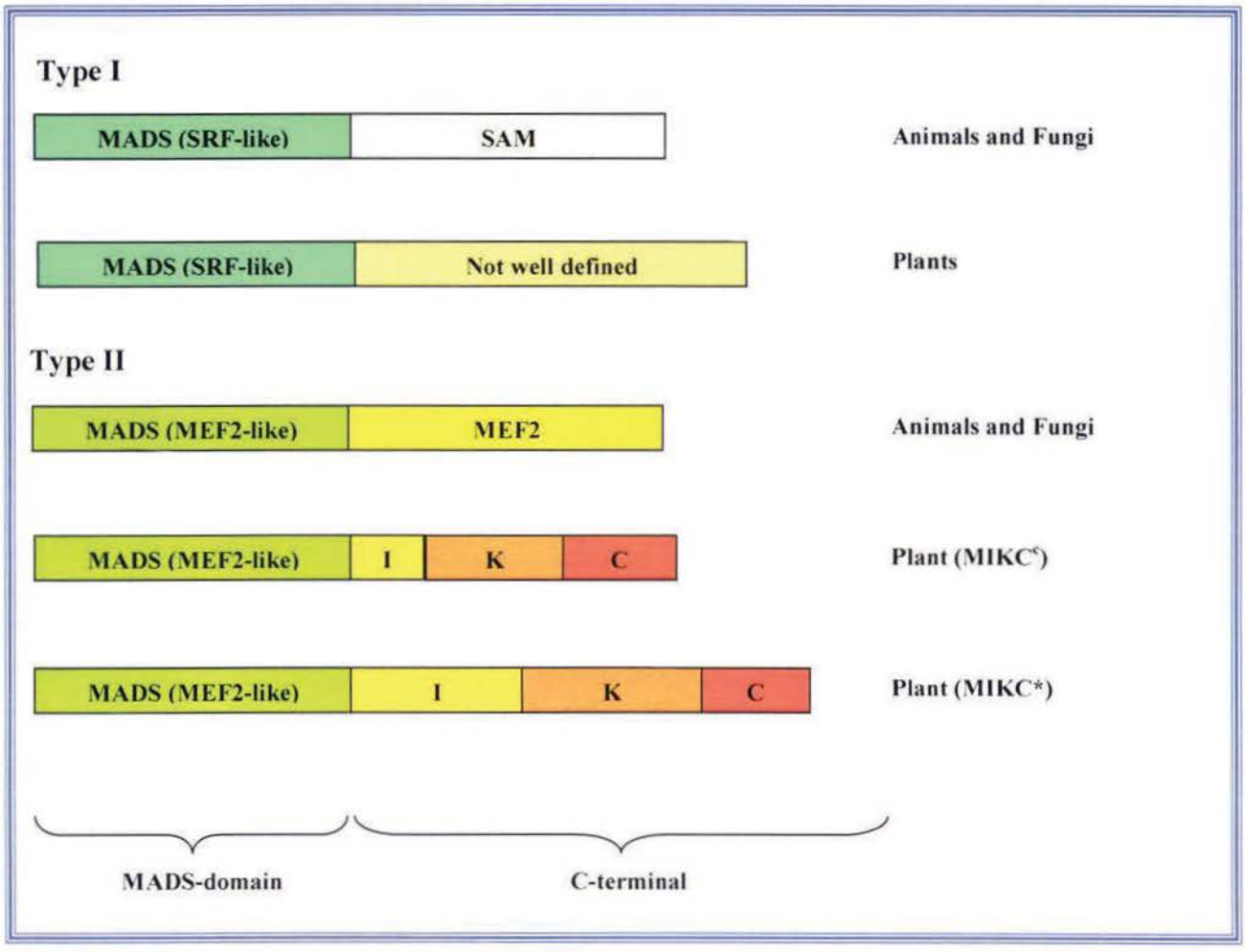

Figure 1.2: Domains of the fungal, animal and plant type I and type II MADS-box proteins.

The MADS domain of the animal, fungi and plant type I genes have a structure similar to SRF-like genes whereas the MADS-domain of type II genes in these three Kingdoms has a structure similar to MEF2-like genes. The C-terminal of the animal and fungi type I is called SAM whereas the C-terminal of their type II genes is called MEF2. In plants, the C-terminal of type I proteins are of variable length and does not have a defined structure, whereas the C-terminal of plant type II proteins have a characteristic "I", "K" and "C" regions. The "I" and "K" regions of plant MIKC* proteins are longer than the "I" and "K" regions of plant MIKC^c proteins.

SRF- SERUM RESPONSE FACTOR; **MEF2-** MYOCYTE ENHANCER FACTOR 2; **SAM-** acronymn for SRF, AG (AGAMOUS) and MCM1(MINI CHROMOSOMAL MAINTENANCE 1); **I-** intermediate domain; **K-** keratin-like domain; **C-** C-terminal domain.

In plants however, the C-terminal of the type I genes are not well defined and is of variable length (De Bodt et al., 2003a). In contrast, the C-terminal domain of plant type II genes is divided into characteristic "I", "K" and "C" domains (Riechmann et al., 1996; Riechmann and Meyerowitz, 1997) (Figure 1.2). The plant type II genes are thus also known as MIKC-type genes. The "M" of the MIKC represents the MADS-box composed of a well conserved sequence of approximately 60 amino acids, which folds into an antiparallel coiled coil of 2 α-helices. This "M" domain is involved in binding of the transcription factor to specific DNA sequences known as the CArG boxes; in dimerization; in accessory factor binding; and in the nuclear localization of the transcription factor (Immink et al., 2002; Ng and Yanofsky, 2001). The "M" domain is followed by an intermediate or "I" domain, which is a short stretch of 30 amino acids that constitutes a key regulatory determinant for the selective formation of DNA-binding dimers (Egea-Cortines et al., 1999). The "K"-box (keratin-like domain) spans 70 amino acids and has regularly spaced hydrophobic residues resulting in an amphipathic α-helix structure that is important for protein dimerization. The "K" domain is exclusive to plant MADS-box type II genes (Alvarez-Buylla et al., 2000b; Theissen et al., 2000). The final domain is the "C"-terminal sequence of variable length that mediates protein-protein interactions by acting as a transactivation domain; assists in the formation of multimeric transcription factor complexes; and contributes to functional specificity (Lamb and Irish, 2003).

1.5.2.1 Further classification of type II MADS-box genes: MIKC and MIKC^c*

Initial phylogenetic analysis by Parenicova et al. (2003) resolved the type II genes into two well-supported sublineages or clades (gene subfamilies), which they named Mδ and MIKC (Figure 1.1). Upon the discovery of two different types of MADS-box genes in mosses, Henschel et al. (2002) renamed these clades MIKC* and MIKC^c (for classic MIKC), respectively. The difference between these two clades lies in the intron/exon structure of their "I" domain. The "I" domains of the MIKC^c type proteins are considerably shorter than those of the MIKC* proteins. Whereas the "I" domain of MIKC^c proteins are generally encoded by one exon, that of MIKC* is encoded by four or five exons (Henschel et al., 2002). The K-domain of the MIKC* also deviates with respect to both length and sequence from the MIKC^c proteins such that there are three indels where the K-domain of the MIKC* proteins are longer than that of the MIKC^c proteins and some of the

hydrophobic amino acids are shifted by one or two positions in the MIKC* proteins (Henschel et al., 2002). It has been suggested that the "I" region of present-day MIKC^c genes evolved by exon deletion from a MIKC*-like ancestor (Kofuji et al., 2003).

In *Arabidopsis* there are approximately 40 MIKC^c type and 6 MIKC* type genes (Becker and Theissen, 2003; Kofuji et al., 2003; Martinez-Castilla and Alvarez-Buylla, 2003; Nam et al., 2004; Parenicova et al., 2003). The MIKC* type genes tend to form a monophyletic group, which is further distinguished into two lineages: one comprising of *AGL66* (*Agamous-like 66*), *AGL104* and *AGL67*, and the other comprising of *AGL30*, *AGL65* and *AGL94* (Kofuji et al., 2003; Verelst et al., 2007). The MIKC^c type genes, on the other hand, are resolved into twelve clades with a diverse range of functions and expression patterns (Becker and Theissen, 2003; Henschel et al., 2002; Kofuji et al., 2003).

1.5.3 Initial identification of MADS-box genes in flowers and its implications

MADS-box transcription factors have been an object of scientific curiosity primarily due to the fact that the initial identification of these genes in plants implicated its role in flower development. A flower is a structure with distinct pattern formation. A typical dicot flower consists of four distinct organs arranged in four whorls. The outermost whorl (whorl 1) consists of green, leaf-like sepals; the second whorl (whorl 2) is composed of usually brightly colored petals; the third whorl (whorl 3) consists of the male reproductive organ, the stamen; and the fourth, innermost whorl (whorl 4) consists of the female reproductive organ, the carpel. Although the phenotype of these four organs differs at maturity, each organ starts as a bulge of cells. Flower formation, therefore, can be thought of as a series of distinct developmental steps: floral induction, formation of the flower primordia, and the production of floral organs.

Investigation into the molecular and genetic processes that controlled flower development led to the identification of several homeotic loci belonging to the MADS-box family of transcription factors (Bradley et al., 1993; Goto and Meyerowitz, 1994; Huijser et al., 1992; Jack et al., 1992; Mandel et al., 1992; Sommer et al., 1990; Trobner et al., 1992; Weigel et al., 1992; Yanofsky et al., 1990). Flower development was thus identified as a significant and useful framework for

comparative studies of gene family function and evolution. Previous comparative studies of developmental processes had shown that despite major morphological differences, different species utilized similar genetic processes to direct development, which implied that new physical features evolved due to small changes in the function of key regulatory genes (Irish and Litt, 2005).

1.5.4 Function of MADS-box genes in flower development

The MADS-box genes that govern floral development belong to the $MIKC^c$ type II group (with the exception of *AP2*, see below) (Bradley et al., 1993; Goto and Meyerowitz, 1994; Huijser et al., 1992; Jack et al., 1992; Mandel et al., 1992; Sommer et al., 1990; Trobner et al., 1992; Weigel et al., 1992; Yanofsky et al., 1990). Characterization of these floral homeotic loci indicated that these genes could be classified into two different groups: meristem identity genes comprising of *APETALA1* (*AP1*), *LEAFY* (*LFY*) (not a MADS-box gene) and *CAULIFLOWER* (*CAL*) and floral organ identity genes comprising of *AP1*, *APETALA2* (*AP2*- not a MADS-box gene), *APETALA3* (*AP3*), *PISTILLATA* (*PI*) and *AGAMOUS* (*AG*). Further research into the floral organ identity genes showed that the five genes (*AP1*, *AP2*, *AP3*, *PI* and *AG*) fell into three separate classes and interaction between these classes formed the basis of distinct organ formation in *Arabidopsis* flowers (Bowman et al., 1989; Bowman et al., 1991). This led to the proposal of the "ABC model" of floral organ development with the A class genes in *Arabidopsis* being *AP1* and *AP2*; the B class genes being *AP3* and *PI*; and *AG* as the only C class gene.

1.5.4.1 The ABC and quartet model of flower development

The ABC model states that the A, B and C classes of genes act in a combinatorial manner to specify floral organ identity (Coen and Meyerowitz, 1991). It proposes that sepal identity is specified by genes of A class alone, petals identity by A+B classes, stamen identity by B+C classes and carpel identity by genes of the C class alone. *In situ* hybridization and ectopic expression experiments provided strong evidence supporting the ABC model and showed that each whorl of the flower primordia had a unique combination of floral organ identity gene activities (Coen and Meyerowitz, 1991; Jack et al., 1994; Krizek and Meyerowitz, 1996). The

ABC model was also shown to be applicable in monocots, specifically in economically important grasses such as maize and rice (Ambrose et al., 2000; Nagasawa et al., 2003; Whipple et al., 2004).

Recently, the E class, composed of the MADS-box genes *SEPALLATA1* (*SEP1*), *SEP2* and *SEP3* (Pelaz et al., 2000) and *SEP4* (Ditta et al., 2004), has been added to the ABC model. These genes are known to combine with genes of the A, B and C classes to specify the identity of the floral organs. For example, petals are specified by A+B+E classes of genes (Honma and Goto, 2001; Pelaz et al., 2000), stamens are specified by B+C+E genes (Honma and Goto, 2001; Pelaz et al., 2000), carpels by C+E genes (Pelaz et al., 2000) and sepal by A+E genes (Ditta et al., 2004). This new "quartet model" (Theissen and Saedler, 2001) directly linked floral organ identity to the action of transcription factor complexes composed of the A, B, C and E classes of MADS box proteins (Figure 1.3).

With the different classes of genes partaking in floral organ identity established, the next question that arose was how the A, B, C and E class genes interacted to specify different floral organs. Riechmann et al. (1996) initially proved that the four MADS-box proteins, AP1, AP3, PI and AG, acted in a combinatorial manner to specify the identity of *Arabidopsis* floral organs. Immunoprecipitation data indicated that these four proteins were capable of interacting with each other and exhibited partner-specificity for the formation of DNA-binding dimers since only AP1 homodimers, AG homodimers and AP3-PI heterodimers were capable of binding to the CArG-box sequence. Recently, it was demonstrated that protein complexes AP3-PI-AP1 and AP3-PI-SEP3 were sufficient to activate the *AP3* promoter (Honma and Goto, 2001). As both AP1 and SEP3 formed homodimers, it was suggested that dimers provided the activation domain and the tetramers (consisting of two dimers) increased the DNA-binding affinity (Theissen and Saedler, 2001). Thus, MADS-box genes were shown to be combinatorial transcription factors, that is, they derived their regulatory specificity from other DNA-binding or accessory factors (Theissen and Saedler, 2001). The combinatorial nature of transcriptional control in eukaryotic cells allows for the generation of regulatory diversity by a limited number of transcription factors (Singh, 1998).

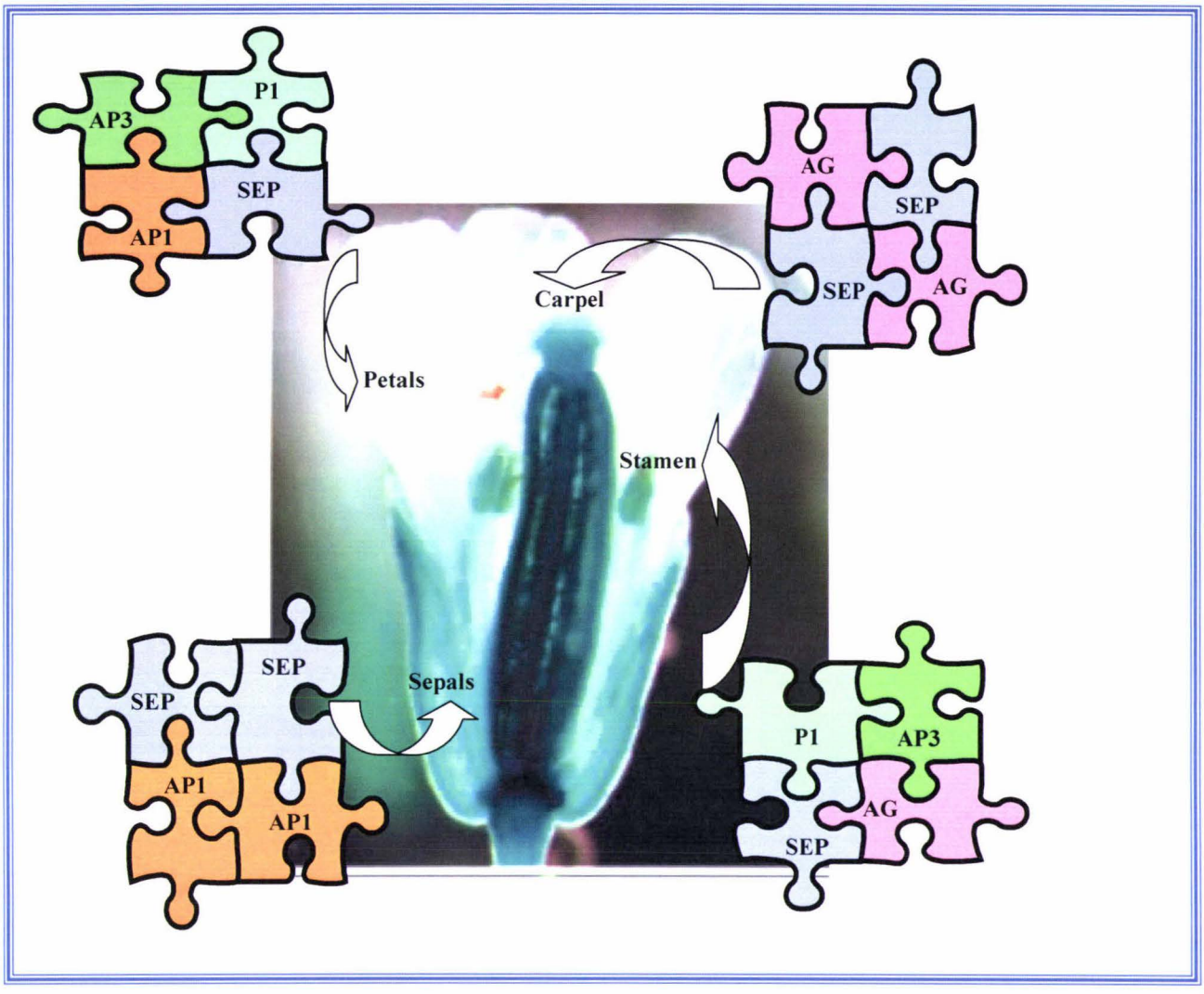

Figure 1.3 The quartet model for floral organ specification in *Arabidopsis*.

The identity of the four floral organs (sepals, petals, stamen, carpel) is determined by combinations of MADS-box proteins. Sepal identity is specified by A+E class genes (AP1, SEP), petal identity is specified by A+B+E class genes (AP1, AP3, PI, SEP), stamen identity is specified by B+C+E class genes (AP3, PI, AG, SEP) and carpel identity is specified by C+E class genes (AG, SEP).

AG- AGAMOUS; **AP**- APETALA; **PI**- PISTILLATA; **SEP**- SEPALLATA

1.5.5 Function of MADS-box genes outside the flower

In addition to flower development, MADS-box genes have a vast number of roles varying from control of flowering time to fruit development to the development of non-floral plant structures such as leaves, stems, roots, seed coat and siliques (Alvarez-Buylla et al., 2000a; Nesi et al., 2002; Rounsley et al., 1995). A number of MADS-box genes were shown to participate in a network of different pathways that lead to flowering (Lee et al., 2000; Reeves and Coupland, 2000). For example, the *Arabidopsis* MADS-box gene, *FLOWERING LOCUS C* (*FLC*), played a role in the vernalization pathway and also regulated two key flowering-time genes, *FT* (*FLOWERING TIME*) and *SOC1* (*SUPPRESSOR OF CONSTANS 1*) (Lee et al., 2000; Michaels and Amasino, 1999; Samach et al., 2000; Sheldon et al., 1999). *MADS AFFECTING FLOWERING 1* (*MAF1*) also behaved similarly to *FLC* (Alvarez-Buylla et al., 2000a; Ratcliffe et al., 2001) while *FRUITFUL* (*FUL*) was involved in the control of floral transition by acting redundantly with *AP1* and *CAL* (Ferrandiz et al., 2000; Gu et al., 1998). In addition to controlling flowering time, *FUL* also participated in the development of fruit valves, while a second set of genes, *SHATTERPROOF 1* (*SHP1*) and *SHP2*, were essential for the differentiation of the valve dehiscence zone (Gu et al., 1998; Liljegren et al., 2000). The gene, *SHORT VEGETATIVE PHASE* (*SVP*), had mutants that caused early flowering in *Arabidopsis* (Hartmann et al., 2000).

MADS-box genes were also involved in the development of non-reproductive structures. For example, the expression of the gene, *ANR1*, was restricted to the roots and was positively regulated by nitrate, resulting in internal root growth (Zhang and Forde, 1998). Another root specific gene, *NMH7*, found in *Medicago sativa* had its expression specifically induced in root nodules upon infection by *Rhizobium* (Heard et al., 1997). The MADS-box gene, *TRANSPARENT TESTA 16* (*TT16*), in *Arabidopsis* was shown to control seed coat pigmentation and endothelial cell specification in seeds (Nesi et al., 2002). In the absence of functional data, functions of some MADS-box genes were inferred from their expression pattern. Examples of these are *AGL12*, *AGL14* and *AGL19*, which were found exclusively in root cells and were thought to be involved in root development (Alvarez-Buylla et al., 2000a; Rounsley et al., 1995). The gene *AGL16* was shown to have a generalized expression in plants but was mostly associated with stomatal guard cells and trichomes in the leaves and also the epidermal cells of the roots, thus implicating its

function in these structures (Alvarez-Buylla et al., 2000b). In addition, *AGL18* and *AGL15* were shown to be expressed in the seed and therefore thought to function in seed development (Alvarez-Buylla et al., 2000b). *AGL15* was also expressed in the embryo in the 8-cell stage and throughout embryo development (Perry et al., 1996), which suggested a possible role of *AGL15* in regulating embryogenesis.

1.5.6 Behavior of MADS-box genes

Since MADS-box genes arose by duplication (Svensson et al., 2000) and phylogeny classifies genes into particular clades based on their sequence similarity, genes arising by duplication have very similar sequences and thus are usually grouped together in phylogenetic trees (De Bodt et al., 2003b; Martinez-Castilla and Alvarez-Buylla, 2003; Parenicova et al., 2003). There are a number of phenomena observed among the members of the MADS-box gene family that has its roots in gene duplication.

For example, it is observed that members of closely related clades usually have similar expression patterns (Alvarez-Buylla et al., 2000a; Alvarez-Buylla et al., 2000b; Nesi et al., 2002a; Purugganan, 1997; Rounsley et al., 1995; Theissen and Saedler, 1995). An example is the B and B_{sister} (B_S) clades of the $MIKC^c$ type genes. The genes of the DEF (DEFICIENS) and GLO (GLOBOSA) clades, which in *Arabidopsis* are represented by *AP3* and *PI*, respectively, together make up the B-class genes. These genes are expressed in the second and third whorl of the *Arabidopsis* flower and are involved in the development of petals and stamens (Goto and Meyerowitz, 1994; Jack et al., 1992). The clade, GGM13, which is represented by *TT16* in *Arabidopsis*, is thought to represent the sister clade of the B genes and thus is called the B_{sister} (B_S) clade. *TT16* is involved in the specification of endothelial cells of the ovule as well as in the control of the flavonoid biosynthesis in the seed coat (Nesi et al., 2002). Therefore, in contrast to B genes, which are mainly expressed in the male reproductive organs (and angiosperm petals), B_S genes are mainly expressed in female reproductive organs, especially in ovules. This clearly demonstrates a link between classification into clades based on sequence similarity and expression pattern where closely related clades have similar expression patterns or are expressed in related organs.

A second pattern observed among MADS-box genes is the strong correlation between the expression pattern and function of most well-characterized genes (Becker and Theissen, 2003; Goto and Meyerowitz, 1994; Gu et al., 1998; Jack et al., 1992; Liljegren et al., 2000; Mandel et al., 1992; Sommer et al., 1990; Theissen and Saedler, 1995; Trobner et al., 1992; Weigel et al., 1992; Yanofsky et al., 1990). The floral organ identity genes of the A, B and C classes are a perfect example of this phenomenon since the genes function in the precise domains/locales in which they are expressed (Bowman et al., 1991; Coen and Meyerowitz, 1991). For example, the A-class genes play a role in sepal and petal identity and their expression is detected only in the whorls that give rise to these two organs (Goto and Meyerowitz, 1994; Irish and Sussex, 1990; Jack et al., 1992; Mandel et al., 1992).

Functional redundancy is a third phenomenon observed among MADS-box genes, with the four *SEP* genes (*SEP1/2/3/4*) being an ideal example. Single gene knockout mutations in any one of the *SEP1/2/3* genes produced subtle phenotypes (Pelaz et al., 2000), an observation consistent with the possibility of gene redundancy. Generation of a triple *sep1/2/3* mutant, however, confirmed the existence of redundancy since all the organs in the mutant flower resembled sepals (Pelaz et al., 2001). This result was complimented by a gain-of-function analysis where overexpression of the *SEP1/2/3* genes, together with the A and B function homeotic genes, were sufficient to convert the vegetative rosette leaves into petals (Pelaz et al., 2001). Furthermore, generation of a quadruple *sep1/2/3/4* loss-of-function mutant resulted in flowers composed entirely of leaf-like structures (Ditta et al., 2004). There are numerous other examples of closely related MADS-box genes with overlapping expression patterns that display functional redundancy, such as the *SHP1/SHP2* and *AP1/CAL* (Ferrandiz et al., 2000; Kempin et al., 1995; Liljegren et al., 2000; Pelaz et al., 2000).

Finally, as discussed above, MADS-box genes are known to act in a combinatorial manner by forming homodimers and heterodimers, which in turn combine to form functional units of transcription factors that regulate gene expression (Coen and Meyerowitz, 1991; Ditta et al., 2004; Pelaz et al., 2000; Shore and Sharrocks, 1995; Theissen and Saedler, 2001; Thiessen et al., 2001).

1.6 MADS-box genes and evolution- what is the link?

The diverse range of functions described for MADS-box genes demonstrates that throughout plant evolution, this family of transcription factors has been recruited as a regulator in the development of a wide variety of plant structures. Since MADS-box genes control development, understanding the evolution of MADS-box genes can facilitate our understanding of the origin and evolution of plant architecture. For example, floral organ identity totally depends on the activity of MADS-box genes. In this respect, the **origin** of the hallmark of the flowering plants, i.e. the flower, can only be understood in the context of MADS-box gene evolution (Henschel et al., 2002). The existence of MADS-box genes (specifically type II genes) in gymnosperms (including the floral homeotic genes), ferns, lycophytes, and mosses (Figure 1.4), none of which form flowers, implies an ancient role of MADS-box genes in lower plants (Hasebe et al., 1998; Henschel et al., 2002; Krogan and Ashton, 2000; Mouradov et al., 1998; Munster et al., 2002; Sundstrom et al., 1999; Svensson and Engstrom, 2002; Tanabe et al., 2003; Tandre et al., 1995; Winter et al., 1999). The question then arises as to **when** the MADS-box genes arose during evolution and what functions they exerted before they attained the specific roles observed in present-day higher plants.

1.6.1 History of land plant life cycle: alternation of generation

One of the most remarkable features of land plant life cycle is the alternation of generation between two phases: the haploid sexual phase (gametophyte) and the diploid asexual phase (sporophyte). In evolution, the development of sex is believed to have preceded alternation of generation (Blackwell, 2003). Thus, the gametophyte generation, which is the sexual generation, is considered to be older than the sporophyte generation. The gametophyte generation likely evolved in an aquatic environment or at least required the presence of water for fertilization. In contrast, it is believed that the entire sporophyte phase of the plant life cycle evolved in a terrestrial environment (Kenrick and Crane, 1997) and is thus thought to be the "key" to terrestrial plant development and their success in the terrestrial environment (Blackwell, 2003).

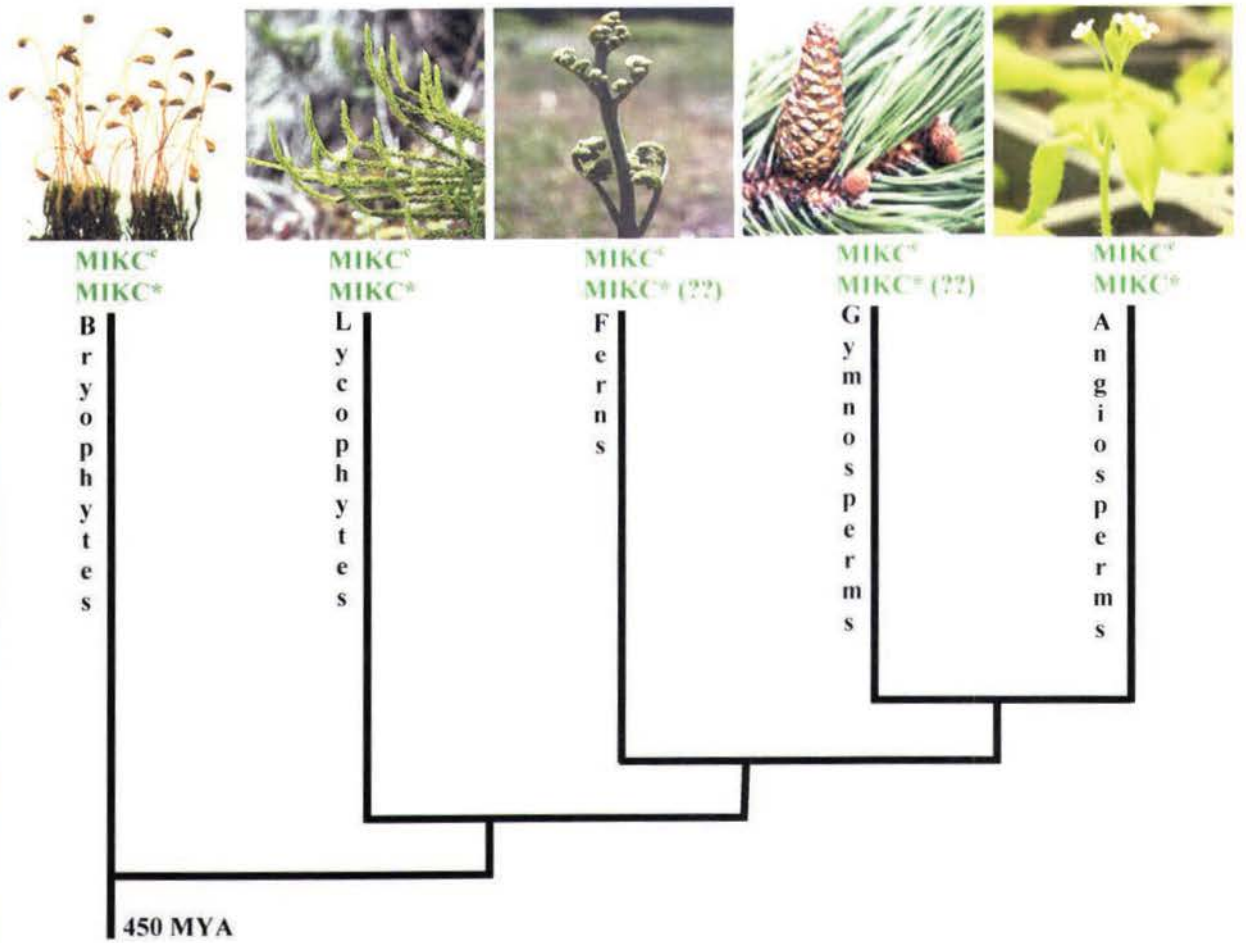

Figure 1.4: MIKC^c and MIKC* type genes in land plants.

Type II MADS-box genes of the MIKC^c and MIKC* type are present in various taxa of land plants. A duplication event is thought to have given rise to these two distinct type II MADS-box genes before the separation of lineages leading to extant mosses and vascular plants >450 million years ago. MIKC^c and MIKC* genes have been reported in bryophytes, lycophytes and angiosperms. Question marks denote uncertainty regarding the presence of MIKC* genes in ferns and gymnosperms. Since both the MIKC^c and MIKC* are present in the basal land plants (bryophytes), it is thought that either the MIKC* genes have been lost or has escaped isolation in ferns.

1.6.1.1 Two theories on the origin of sporophyte or alternation of generation

There are two competing theories on the evolution of alternating generation in land plants, specifically the origin of sporophytes:

- The "homologous theory" considers the sporophyte to be a direct modification of the gametophyte but with a specific function of spore production. Support for this theory comes from the existence of algae such as *Ulva*, which has isomorphic gametophytes and sporophytes. Such ancestral algae, with both a gametophyte and a sporophyte phase already present in its life cycle, is believed to have given rise to land plants (Blackwell, 2003).
- The "antithetic theory", on the other hand, suggests that only the gametophyte was present initially in the algal life cycle and the sporophyte arose subsequently and in correspondence with (or soon after) land occupancy by the gametophyte. It suggests that the sporophyte appeared in the life cycle by amplification of the zygote where a delay in meiosis, followed by a series of mitotic division, produced a mass of cells that gave rise to the new diploid sporophyte generation (Blackwell, 2003).

1.6.2 Life cycle of lower land plants

Based on molecular data, the charophycean algae are considered to be the closest relatives of modern land plants, with the charalean algae being an example of an extant member of this group (Delwiche et al., 2002; Karol et al., 2001). In the life cycle of the *coleochaete*, which forms a sister-group to the charalean algae, the sperm fertilizes the egg, which is attached to the female gametophyte, to form a diploid zygote (Graham and Wilcox, 1983). This zygote accumulates nutrients and undergoes post-fertilization growth to form a zygospore. After a period of dormancy, the zygospore undergoes meiosis to produce 8-32 haploid cells that are released as zoospores. The zoospores then settle on a favorable substrate and germinate into multicellular gametophytes through mitotic divisions (Graham and Wilcox, 1983).

The bryophytes, which are the most basal of land plants (Nickrent et al., 2000; Nishiyama and Kato, 1999), have a life cycle that is obligatorily tied to an aquatic environment. In bryophytes (liverworts, hornworts and mosses), fertilization takes place when the sperm swims from the male gametophyte to the female gametophyte. The resulting diploid zygote undergoes several

mitotic divisions to produce a multicellular sporophyte, which in turn produces a sporangium that undergoes meiosis to form haploid spores (Haig and Wilczek, 2006). These spores later germinate into separate multicellular male and female gametophyte. Similar to *coleochaetes*, bryophytes have a dominant multicellular gametophyte phase. However, unlike the *coleochaetes*, bryophytes possess a *multicellular* sporophyte. In both these taxa, the sporophyte is nutritionally dependent on the gametophyte.

Lycophytes are thought to be the closest relatives of extant vascular plants. The life cycle of lycophytes and ferns are similar in that the gametophyte generation is small and inconspicuous. They usually produce eggs and sperms on different parts of the same gametophytic structure, thus minimizing the need for the presence of water for fertilization. Fertilization produces a diploid zygote that develops into a mature multicellular sporophyte and attains nutritional independence from the gametophyte during the course of its development. The sporophytes produce haploid spores that germinate into individual multicellular gametophytes. Lycophytes and ferns represent a middle ground between the life cycles of mosses and seed plants. In contrast to the mosses, the multicellular sporophyte in lycophytes and ferns are the dominant generation that develops independently of the gametophyte. This switch in the dominance of the gametophyte and sporophyte generation is maintained in higher seed plants.

The haploid gametophyte generation in higher plants, such as the gymnosperms and angiosperms, is drastically reduced in stature and usually embedded in the sporophyte. As mentioned above, in these taxa, the sporophyte is the dominant generation. They produce multicellular haploid gametophytes that are dependent on the sporophyte during development. The gametes, upon fertilization, give rise to a diploid zygote that develops into a multicellular embryo while contained within the seed. After this brief dependence of the embryo (sporophyte) on the gametophytic tissues for nourishment, the diploid sporophyte develops independently following seed germination.

Taking into account the reproduction modes of land plants, from the basal (mosses) to the higher plants (angiosperm), the dominance between the gametophyte and sporophyte generation in plant life cycle and the multicellularity of these two generations is observed to have changed dramatically, with the gametophyte becoming reduced, embedded, and physiologically dependent on the sporophyte in higher plants.

1.6.3 Expression pattern of MADS-box genes in land plants

The change in the dominance of the two generations in land plants has been accompanied by changes in the expression pattern of MADS-box genes. Whereas the MADS-box genes in lower land plants show a broad expression pattern, the genes in higher plants exhibit very specific expression in the plant tissues in which they function (Kofuji et al., 2003).

MADS-box genes of the MIKC^c type have been isolated from algae where they were expressed in the gametophyte generation (Tanabe et al., 2005). Since, algae are the closest relatives of land plants (Delwiche et al., 2002; Karol et al., 2001), the implication was that MADS-box genes originally functioned in the haploid gametophyte phase and were recruited into the diploid sporophyte generation later as the land plants evolved (Tanabe et al., 2005). Mosses, being the most basal of land plants, possess two different types of MADS-box genes, the MIKC^c and MIKC* type (Figure 1.4), belonging to the type II lineage (Krogan and Ashton, 2000). This suggested that the most recent common ancestor of mosses and vascular plants possessed at least one MIKC^c and one MIKC* type gene. A duplication event is thought to have given rise to these two distinct type II MADS-box genes before the separation of lineages leading to extant mosses and vascular plants >450 million years ago (Henschel et al., 2002). In contrast to algae, the MIKC^c type MADS-box genes in mosses showed a broad expression pattern since it was present in both the sporophytic and gametophytic tissues (Krogan and Ashton, 2000). The MIKC* type genes in mosses, on the other hand, were detected only in the gametophytic tissues (Henschel et al., 2002).

Lycophytes (Figure 1.4), express both the MIKC^c and MIKC* type genes in the sporophyte generation (Svensson and Engstrom, 2002; Tanabe et al., 2003). Their expression pattern in the gametophytic tissues has not been studied due to the difficulty in obtaining these tissues (Svensson and Engstrom, 2002; Tanabe et al., 2003). The first MADS-box gene reported in lycophytes, *LAMB1*, belonged to the MIKC* type and was found to be exclusively expressed in the developing strobili of the sporophytic generation in *Lycopodium annotinum* (*L.annotinum*) (Svensson and Engstrom, 2002). The discovery of *LAMB1* was important since this was the first report of the existence of a MIKC* type gene (Svensson and Engstrom, 2002). Genes of the MIKC^c type in lycophytes are expressed in a wide range of vegetative and reproductive tissues (Svensson and Engstrom, 2002). In ferns, only genes belonging to the MIKC^c type have been isolated (Hasebe et al., 1998; Munster et al., 1997). These genes

showed a broad expression pattern in the gametophyte and sporophyte generation (Hasebe et al., 1998; Munster et al., 1997). So far, there have been no reports of MIKC* type genes in ferns, thus implying that either the MIKC* genes have been lost in ferns (Kofuji et al., 2003) or they have escaped identification as a result of difficulties encountered in obtaining gametophytic tissues of ferns due to its small and inconspicuous stature.

A plethora of MADS-box genes involved in the development of various structures in both the sporophyte and the gametophyte generation of angiosperms has been identified (section 1.5). These genes show very precise temporal and spatial expression patterns compared to the lower plants, thus indicating that the MADS-box genes in angiosperms acquired highly specialized functions (Kofuji et al., 2003). In angiosperms, most MIKC^c type genes, except *AGL18* (Alvarez-Buylla et al., 2000a), are expressed in the sporophyte, usually in specific tissues of the vegetative and reproductive structures. The MIKC* genes, on the other hand, are found predominantly in the gametophytic tissues of angiosperms (Honys and Twell, 2004; Kofuji et al., 2003; Pina et al., 2005; Verelst et al., 2007).

The above facts indicate that dynamic changes have occurred in gene expression patterns between the gametophyte and sporophyte generation during land plant history. The expression of MADS-box genes appear to have evolved from a broad expression pattern in lower land plants to more discrete expression patterns in higher plants where they control development of specific tissues. These changes are thought to be correlated with the evolution of body plan (Kofuji et al., 2003) and maybe even changes in the dominance and multicellularity of the gametophyte and sporophyte generation of land plants.

1.7 Can plant evolution be explained in the context of the expression pattern of the MIKC^c and MIKC* type MADS-box genes?

MIKC^c type genes have been detected in a broad range of taxa, from algae to angiosperms, where they exhibit a broad expression pattern in the gametophytic and sporophytic tissues and control the development of a wide variety of structures. The MIKC* type genes, on the other hand, have so far been found in mosses, lycophytes and angiosperms only and display very specific expression patterns (Henschel et al., 2002; Honys and Twell, 2004; Kofuji et al., 2003; Pina et al., 2005; Svensson and Engstrom, 2002; Verelst et al., 2007). In lycophytes, MIKC*

genes are expressed only in the reproductive structures of the sporophyte generation (Svensson and Engstrom, 2002) while in mosses and angiosperms, these genes are found predominantly in the gametophyte (Henschel et al., 2002; Honys and Twell, 2004; Kofuji et al., 2003; Pina et al., 2005; Verelst et al., 2007), which is the sexual or reproductive phase of a plant life cycle (Blackwell, 2003).

The preservation of the expression pattern of the MIKC* type genes between the most basal (moss) and highest (angiosperm) plant taxa is intriguing since the physical features of the gametophyte generation of extant mosses and angiosperms differ drastically. While the moss gametophyte is a dominant multicellular generation, which differentiates into shoot- and root-like structures, the gametophyte of angiosperms is highly reduced and embedded, thus very inconspicuous. The male gametophyte in angiosperm is composed of a three-celled pollen grain while the female gametophyte is a seven-celled embryo sac which is embedded and completely dependent on the maternal sporophyte. Such comparisons of the physical features of moss and angiosperm gametophytes makes it difficult to envision a common role of MIKC* type genes in their development. However, the specific and predominant expression of the MIKC* genes in the gametophyte of these two taxa cannot be ignored.

The gametophyte is the sexual generation of land plants. Higher plants, with its reduced but complex gametophyte, have been very successful throughout evolution. Is it possible that mechanisms governing the differentiation of reproductive structures in higher and lower plants are fundamentally similar, even though the structures *per se* are physically distinct? To answer such questions, functional characterization of MIKC* genes is important. Since *Arabidopsis* is used as a model species for investigation into the workings of higher plants, a large amount of data on its genome is currently available. Thus, the function of MIKC* type genes in the angiosperm gametophyte can be readily investigated. The results will not only enable comparison with the ancestral function of MIKC* genes in mosses, which in turn would give clues to the evolution of MADS-box genes and subsequently, the evolution of land plant forms, but also give insight into the development of higher plants, specifically their gametophytes.

1.8 Gametophyte development in angiosperms

1.8.1 Male gametophyte

The formation of the gametophyte from the sporophyte in higher plants is the result of two sequential processes: sporogenesis and gametogenesis. Sporogenesis starts with the differentiation of the hypodermal cells of the ovules and anthers to form the archesporial cells, which later differentiates into the female megasporocytes or the male microsporocyte, respectively and undergoes meiosis (Yang and Sundaresan, 2000).

Microsporogenesis (male sporogenesis) and microgametogenesis (male gametogenesis) takes place within the anther locules, or the pollen sac, and depend on complex interactions between sporophytic and gametophytic cells. This interaction is very important since sporophytic mutations are known to disrupt pollen development and cause male sterility (Chaudhury, 1993; Sanders et al., 1999; Taylor et al., 1998; van der Meer et al., 1992). During microsporogenesis, the archesporial cells divide mitotically to form an outer, primary parietal layer and an inner, sporogenous layer. The parietal layer undergoes periclinal and anticlinal division, giving rise to several concentric layers that differentiate into the endothecium, middle layers and innermost tapetum, whereas the sporogenous layer gives rise to the microsporocytes or the pollen mother cells (PMC) (Twell, 2002). The tapetum is identified to be an important tissue since it secretes enzymes required for callose dissolution (for microspore release); controls the supply of nutrients required by the pollen; and controls the deposition of secondary metabolites required for elaboration of the unique pollen wall, UV protection, pollinator attraction and pollen fertility by undergoing polar degeneration to deposit its cell contents within the exine cavities to form the pollen coat (Murgia et al., 1991).

1.8.1.1 Microspore development

Microgametogenesis may be considered to be comprised of two phases: microspore development, which terminates with the completion of PMI, and pollen development, which culminates in the release of the sperm cells into the ovules. Microspore development starts with meiosis within the PMC and results in a tetrad of microspores, each of which are enclosed in a thick callose wall and is thus isolated from each other and the diploid parent (Figure 1.5). This enables independent development of the individual spores, which at this stage is

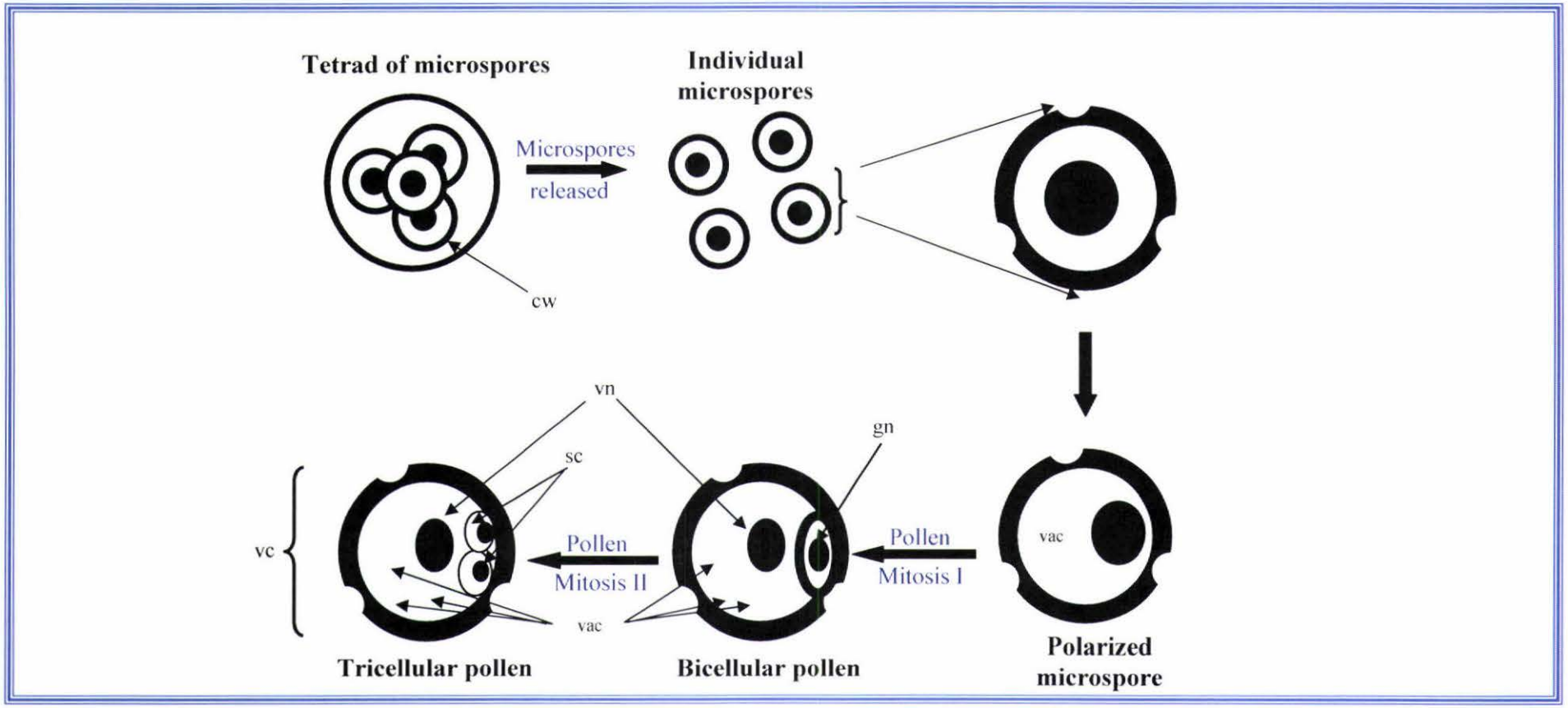

Figure 1.5 Microgametogenesis in *Arabidopsis*.

Microgametogenesis is initiated when meiosis of pollen mother cells give rise to a tetrad of uninucleate microspores. Each microspore then undergoes mitosis to form the bicellular pollen containing the vegetative and generative nuclei. The generative nucleus further divides to form 2 sperm cells and these are contained within the vegetative cell. In *Arabidopsis*, the mature pollen is tricellular, containing 2 sperm cells imbedded inside the vegetative cell.

cw- callose wall; **gn**- generative nucleus; **sc**- sperm cells; **vac**- vacuole; **vc**- vegetative cell; **vn**- vegetative nucleus

uninucleate (one nucleus). During meiosis, cytoplasmic reorganization leads to the degradation of most of the PMC RNA and ribosomes, and the dedifferentiation of plastids and mitochondria. This erases macromolecules from the sporophytic phase and thus provides a neutral environment to establish the gametophytic program. In addition, the connections between the PMC, tapetum and the wall layers are severed at this point (Twell, 2002). The isolated microspores enlarge and each undergoes an asymmetric mitosis, called pollen mitosis I (PMI,) to form two cells of different sizes: a larger vegetative cell (VC) and a smaller generative cell (GC).

Disruption of PMI affects pollen development as indicated by the isolation of several mutants (Twell, 1999; Twell and Howden, 1998a; Twell and Howden, 1998b). The gene, *SPOROCYTELESS* (*SPL*) (Yang et al., 1999), encodes a MADS-box transcription factor whose expression is restricted to the sporogenous cells and PMC, suggesting a primary role in formation of microsporocytes. In *spl* mutants, subepidermal cells of the anther are able to form the sporogenous and parietal cells, but subsequently fail to form microsporocytes. In addition, the formation of the anther walls and tapetum is disrupted (Yang et al., 1999). It is thought that the primary role of *SPL* may be to promote the formation of the sporocyte and the defect in anther wall development is indirect. These results imply that the development of anther walls and tapetum is dependent upon signals from the microsporocytes, without which sporophytic development of the anther cannot proceed (Yang et al., 1999).

1.8.1.2 Pollen development

PMI signifies the termination of microspore development and the initiation of pollen development (male gametophyte development). The two cells of the bicellular pollen grain have different fates. The VC does not divide again. It constitutes the bulk of the pollen cytoplasm, is transcriptionally active, and eventually forms the pollen tube. The GC, on the other hand, is engulfed inside the cytoplasm of the vegetative cell, contains relatively few organelles and stored metabolites, and is transcriptionally repressed (Yang and Sundaresan, 2000). This represents a "cell within a cell" structure that has evolved to enable internalization and transport of the male gametes within the pollen tube.

The asymmetric division during PMI is important for the fate of daughter cells as revealed by differences in cell cycle progression and cell-specific gene expression. When the asymmetry is

disrupted by *in vitro* colchicine treatment in tobacco, both daughter cells adopt VC fate (Eady et al., 1995). The mutant, *gemini pollen 1* (*gem 1*), also displays a range of pollen phenotype depending on equal, unequal or partial divisions at PMI (Park et al., 1998). *GEM1* is thought to function in the spatial coordination of nuclear and cytoplasmic division. Cell fate analysis in this mutant revealed a quantitative relationship between VC fate and division symmetry such that the nuclear/cytoplasmic ratio was a factor in the determination of cell fate (Park et al., 1998). The nuclear division in *gem1* was often uncoupled from cytokinesis, producing enucleate cytoplasmic compartments, indicating that cytokinesis (division of cytoplasm) and karyokinesis (division of chromosomes), although tightly linked, could operate independently (Park et al., 1998). Other mutants, such a *sidecar pollen* (*scp*) (Chen and McCormick, 1996) and *two-in-one pollen* (*tio*) (Twell and Howden, 1998a), were also defective in asymmetric division and show altered generative cell fate, thus confirming that microspore polarity during PMI is very important for generative cell differentiation. An extensive number of other gametophytic mutants have been described that affect meiotic chromosome pairing, segregation and cell division, some of which are specific to male meiosis (Yang and Sundaresan, 2000).

All known seed plant species produce pollen that follow one of two general pathways that are distinguished on the basis of the timing of PMII. In the majority of the species, pollen is shed in a bicellular condition and the GC divides to form two sperm nuclei within the pollen tube. In tricellular species (such as *Arabidopsis*), the GC divides before pollen is shed from the anther (Twell, 2002).

In *Arabidopsis*, the primary function of the three-celled male gametophyte (mature pollen grain) is the transport of the two sperm cells to the embryo sac of the ovule, where they participate in double fertilization. One of the sperm cells will fertilize the egg cell to form the zygote that develops into an embryo while the other will fertilize the diploid central cell to form a triploid endosperm (McCormick, 2004). A key structural assemblage formed within the male gametophyte or mature pollen grain is known as the male germ unit (MGU), which consists of a physical association between the two sperm cells (SC) and the vegetative cell (VC) (Twell, 2002). The major activity of the vegetative cell of the male gametophyte is the synthesis and assembly of the pollen tube wall (Mascarenhas, 1993) in order to achieve its ultimate goal of gamete transfer to the ovule for successful fertilization.

Pollen germination and pollen tube elongation signifies the last stages of pollen development and terminates with the release of the sperm cells into the embryo sac (Twell, 2002). The pollen tube is a highly regulated and homeostatically governed cell type which is capable of growing on minimal requirements and enduring differences in the concentration of external solutes. During growth, pollen tubes face the mechanical barriers of the stigma and style in order to deliver the sperm cells into the ovule for fertilization. The nature of pollen tube growth is different in that growth is restricted to the tip region only, as opposed to growth taking place over the entire surface of most other types of plant cells (Mascarenhas, 1993).

A fundamental aspect of tip growth in pollen tubes is the continuous deposition of new cell wall and plasma membrane at the tube apex (Feijo et al., 2001; Mascarenhas, 1993), a process called exocytosis. The wall precursors are contained in Golgi-derived vesicles which are transported by the actin cytoskeleton (Feijo et al., 2001; Mascarenhas, 1993). The Golgi vesicles are organized in an inverted cone, with the base of the cone adjacent to the apical plasma membrane and the tip of the cone extending into the tube. Elongation of the pollen tube is achieved by the fusion of Golgi vesicles with the plasma membrane and the release of its contents. This fusion is significantly regulated by the transport of inorganic ions, such as Ca^{2+} and K^{+}, across the plasma membranes of pollen and/or pollen tubes (Fan et al., 2001; Feijo et al., 2001). Earlier studies had determined that the extracellular flux of Ca^{2+} was inwardly directed at the tip of the growing tube where a steep gradient of Ca^{2+} existed (Taylor and Hepler, 1997). It was also evident that ion entry was restricted to a small region at the extreme apex where growth occurred and membrane deformation was maximal (Taylor and Hepler, 1997). It is now commonly accepted that there is a breaking mechanical point based either in the yielding of the wall or in the increase of internal turgor, which leads to stretching of the plasma membrane and results in the consequent opening of the Ca^{2+} channels to allow Ca^{2+} to flow in and promote vesicle fusion leading to tube elongation (Feijo et al., 2001).

1.8.2 Female gametophyte

As in the male gametophyte, sporogenesis and gametogenesis are the two processes that lead to the formation of the female gametophyte from the sporophyte in higher plants. The transition between the sporophyte and gametophyte in females occur in the ovule, which is the site of meiosis, female gametogenesis and double fertilization. Ovules consist of tissues derived from both generations of the plant life cycle, the diploid sporophyte and the haploid gametophyte.

The formation of ovules harboring the megagametophyte is a key step in sexual reproduction. Several processes, such as primordia initiation; specification of identity; pattern formation; and eventually morphogenesis and cellular differentiation contribute to ovule development (Grossniklaus and Schneitz, 1998; Skinner et al., 2004). Megasporogenesis (female sporogenesis) in the ovule starts with the enlargement and differentiation of a hypodermal cell, found at the distal end of the primordium (nucellus), into archesporial cells, which differentiate into the female megasporocyte or megaspore mother cell (MMC) (Yang and Sundaresan, 2000). The MADS-box gene, *SPL*, also plays a role in megasporogenesis since *spl* mutants fail to form MMC (Yang et al., 1999).

Shortly after the MMC becomes visible, two integuments originate from the chalazal nucellus, each of which consists of two epidermally derived cell layers. They eventually cover the ovule leaving only a small cleft (micropyle) at the distal end through which the pollen tube enters to effect fertilization (Grossniklaus and Schneitz, 1998). Outer integument elongation is asymmetric with extensive growth occurring at the outer side. This causes the characteristic curvature of the ovule and results in the placement of the micropyle next to the funiculus in *Arabidopsis*. The funiculus, with an internal vascular strand, connects the ovule to the placenta (Grossniklaus and Schneitz, 1998).

Integuments are important for the progression of the meiotic divisions of the megasporocyte since defects in integument development in *Arabidopsis* result in meiotic arrest. This is demonstrated by *aintegumenta* (*ant*) and *bell* mutants where the integuments are either missing (*ant*) or abnormal (*bell*) (Klucher et al., 1996; Reiser et al., 1995). Recently, a mutant, *sterile apetala* (*sap*), was isolated in which megasporogenesis was arrested during or just after the first meiotic division even though integument development is normal (Byzova et al., 1999). The *SAP* gene is initially expressed in the nucellar region and later confined to the developing integuments. The *sap* phenotype and the expression pattern of *SAP* could thus indicate that signals from the integuments, regulated or mediated by the *SAP* gene product, are required for the meiotic progression of the megasporocyte.

Meiosis marks the start of megagametogenesis (female gametogenesis) (Figure 1.6), which is highly coordinated with the development of the surrounding sporophytic tissues of the ovule. The MMC undergoes two rounds of meiosis forming a tetrad of haploid megaspores. Out of these, only the chalazal megaspore becomes functional while the other three undergo

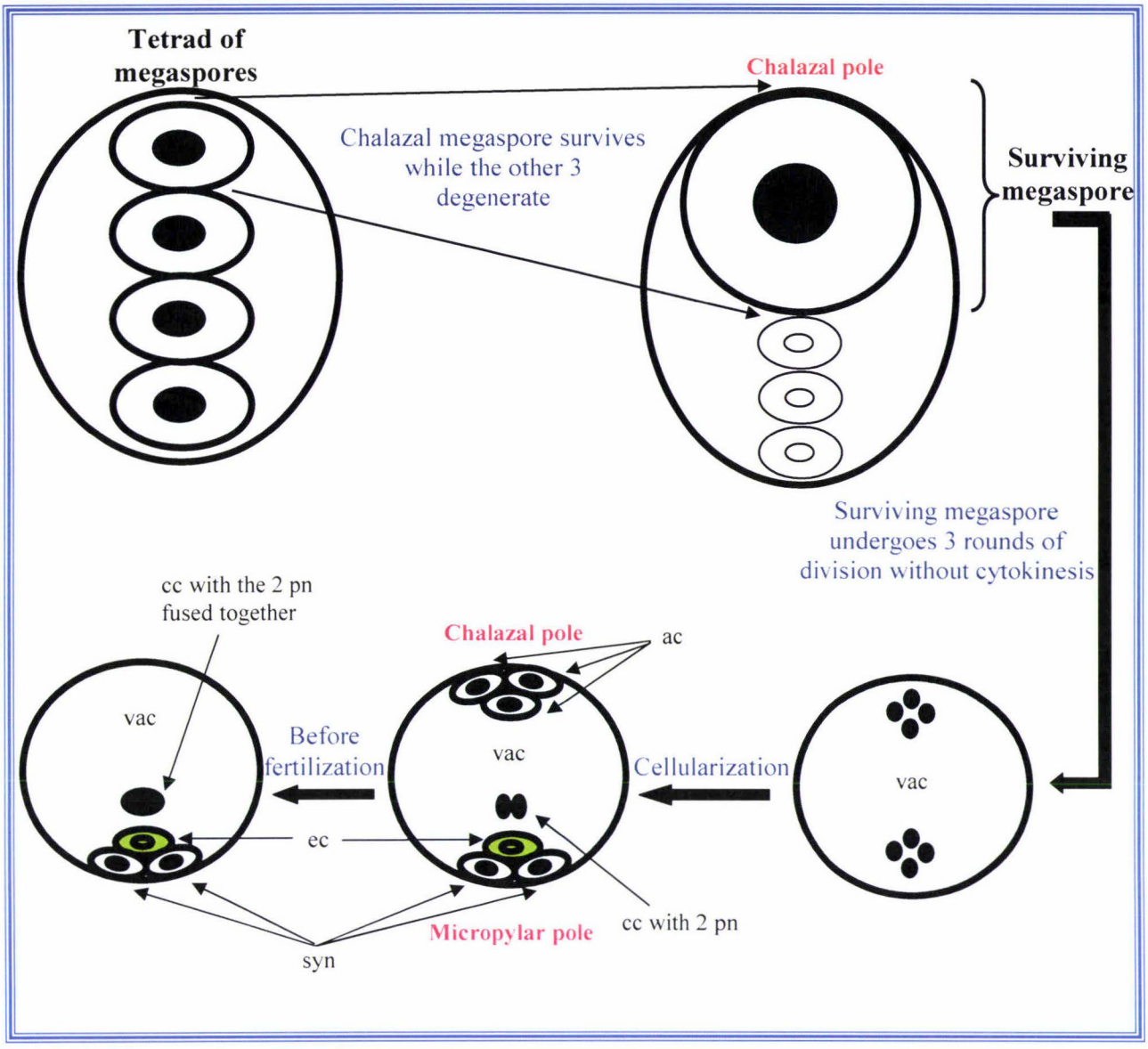

Figure 1.6: Megagametogenesis in *Arabidopsis*.

Megagametogenesis is initiated when meiosis of the megaspore mother cells give rise to a tetrad of megaspores. Only the chalazal megaspore survives while the other three degenerate. The surviving megaspore undergoes 3 rounds of division without cytokinesis to form an eight-nucleate cell. Cellularization gives rise to three antipodal cells on the chalazal pole, two synergids and an egg cell at the micropylar pole and a central cell with two nuclei. This eight-nucleate, seven-celled structure comprises the female gametophyte or embryo sac. Before fertilization, the antipodal cells degenerate and the two nuclei of the central cell fuse.

ac- antipodal cells; **cc**- central cell; **ec**- egg cell; **pn**- polar nuclei; **syn**- synergids; **vac**- vacuole

programmed cell death (Grossniklaus and Schneitz, 1998; Yang and Sundaresan, 2000). The surviving megaspore enlarges and its nucleus divides 3 times without cytokinesis. This eight-nucleate female gametophyte (embryo sac) contains 4 nuclei at both poles that are separated by a large central vacuole (Grossniklaus and Schneitz, 1998). Cellularization leads to the formation of 3 antipodal cells at the proximal pole, an egg and 2 synergids (together called the egg apparatus) at the distal pole, and a large central cell containing 2 nuclei called the polar nuclei (Grossniklaus and Schneitz, 1998). In a mature female gametophyte, the central cell occupies most of the volume, is located at the chalazal pole, contains a large vacuole at its chalazal pole, and has a relatively large nucleus at its micropylar pole (Huang and Russell, 1992; Yadegari and Drews, 2004). This eight-nucleate, seven-celled embryo sac represents the mature female gametophyte. In *Arabidopsis* and many other species, the polar nuclei fuse to form the diploid central cell nucleus (secondary nucleus) and the antipodal cells degenerate before fertilization (Portereiko et al., 2006).

1.9 Where does *AGL104* fit in?

1.9.1 Importance of *AGL104* as a MIKC* type gene in mosses and angiosperms

AGL104 is one of the six MIKC* type genes found in *Arabidopsis* (Honys and Twell, 2004; Kofuji et al., 2003; Verelst et al., 2007). To date, none of the MIKC* have been functionally characterized, either in angiosperms or in lower land plants. Thus, characterization of a MIKC* type gene such as *AGL104* will shed light on the function of MIKC* type genes and their role in the development of the gametophyte generation in land plants. This would enable comparative analysis with other species, particularly the lower plants, and facilitate understanding of the changes in the gametophyte generation throughout evolution.

1.9.2 Importance of *AGL104* as a transcription factor in pollen

Another noteworthy feature of *AGL104* is that, being a transcription factor, it is expressed very late in the *Arabidopsis* male gametophyte. Transcriptome analyses of mature pollen revealed an over-representation of the type I and MIKC* type genes in pollen when all other types of transcription factors were under-represented (Pina et al., 2005). This increased the curiosity surrounding MIKC* type genes. The four MIKC* type genes, *AGL104*, *AGL66*, *AGL30* and

AGL65 (out of the five examined) were pollen enriched (high number of transcripts) and three of these (*AGL104, AGL66* and *AGL30*) were selectively (exclusively) expressed in pollen (Pina et al., 2005). Expression of these same genes in mature pollen was also detected using Macroarray and Northern Blot analysis by Kofuji et al. (2003). In addition, Honys and Twell (2004) reported the expression of *AGL104* exclusively in pollen, with initial low levels of expression at the uninucleate stage that grew stronger as pollen development progressed. The strongest level of expression was again detected in mature pollen (Honys and Twell, 2004 Supplementary Data).

It had previously been shown that mature pollen continuously accumulated mRNA molecules from the stage of microspore mitosis through to the stage of pollen maturation. Some of these transcripts were expected to encode proteins involved in pollen germination and tube growth (Mascarenhas, 1990). Recently, transcriptome analysis of mature pollen verified expression of genes with proposed functions in signal transduction, cell wall biosynthesis, vesicle trafficking, transport across membranes, and cytoskeletal dynamics (Becker et al., 2003; da Costa-Nunes and Grossniklaus, 2003; Feijo and Moreno, 2004; Hepler, 2001; Honys and Twell, 2003; Honys and Twell, 2004; Pina et al., 2005), thus indicating a commitment of mature pollen towards pollen germination and growth of the tubes.

The expression profile of *AGL104* reported by Honys and Twell (2004) is intriguing since it implicates a role of *AGL104*, a transcription factor, in pollen germination and pollen tube elongation (Mascarenhas, 1990). It is commonly accepted that pollen germination and tube elongation is independent of transcription (Mascarenhas, 1975; Mascarenhas, 1990). Why then, is a transcription factor being expressed in mature pollen? Is it possible that contrary to popular belief, the process of pollen germination and tube growth involves active transcription? Even though the general trends and changes during the development of the *Arabidopsis* gametophytes (male and female) have been identified (Grossniklaus and Schneitz, 1998; McCormick, 2004; Skinner et al., 2004; Twell, 2002; Yadegari and Drews, 2004; Yang and Sundaresan, 2000), precise information on the developmental steps leading to a mature gametophyte is lacking in the current literature. For example, it is not known what mechanisms are used by the mature pollen grain to facilitate pollen tube germination and tube elongation or the precise role played by the mature ovule during pollen tube guidance to ensure fertilization (Hulskamp et al., 1995; Kandasamy et al., 1994). Thus, characterization was *AGL104* was also

expected to shed light on gametophyte development in *Arabidopsis*, especially the role of transcription factors during pollen germination and pollen tube elongation.

1.10 Hypothesis and aims

Prior to our study, the only published data available on *AGL104* was the preliminary expression analysis carried out by Parenicova et al. (2003), who detected *AGL104* expression in the inflorescence, particularly in the inside of whorl 1/sepal (in early stage 3 flowers) and in young developing anthers, petals and within the septum and developing ovules of the carpel. With a limited amount of data available, we thus considered the generals trends of MADS-box genes (section 1.5.6) when formulating a hypothesis for our project.

1.10.1 Hypothesis

Taking into account the behavior of MADS-box genes (section 1.5.6), phylogenetic trees can thus be utilized as a valuable tool when postulating functional hypotheses for uncharacterized genes (Martinez-Castilla and Alvarez-Buylla, 2003). *AGL104* is most closely related to the members of the B_s clade (Becker et al., 2003; Martinez-Castilla and Alvarez-Buylla, 2003) (Figure 1.7). Considering that genes of the B_s clade functions in the female reproductive organs (Nesi et al., 2002) and those of the B clades function in the male reproductive organs (Goto and Meyerowitz, 1994; Jack et al., 1992), it was most likely that *AGL104* functioned in the reproductive structures of *Arabidopsis*. Incorporating this information with the expression analysis results of Parenicova et al. (2003), which showed *AGL104* expression in sepals, petals, anther and carpel, the following hypothesis was proposed:

- *AGL104* has a role in the development of the reproductive structures of the *Arabidopsis* flower, most probably in the development of the male and female gametophyte.

The validity of our hypothesis was further reinforced by preliminary expression analyses carried out in our laboratory that demonstrated that *AGL104* was expressed only in the inflorescence of *Arabidopsis* (Ambrose, pers comm). In addition, Northern Blot analysis by Kofuji et al. (2003) had also detected expression of *AGL104* in pollen.

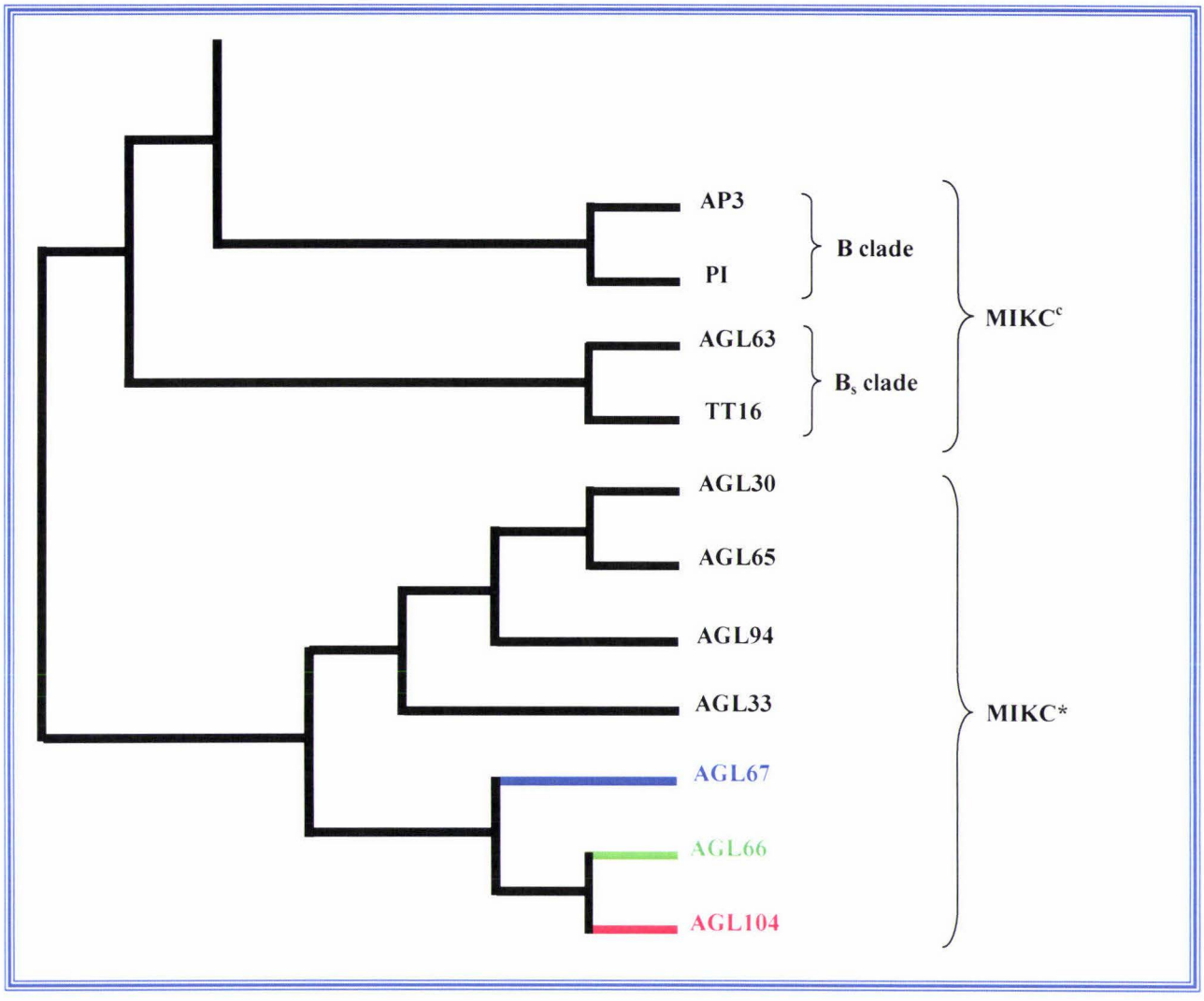

Figure 1.7: Phylogenetic relationship between the MIKC^c (B and B_s) and MIKC* clades

The MADS-box genes of the B clade are expressed in the male reproductive structures and genes of the B_s clade are expressed in the female reproductive structures and developing seeds. Since closely related genes usually exhibit similar expression patterns and most well characterized MADS-box genes function in the structures in which they are expressed, the close relationship of *AGL104* (and other MIKC* genes) to the B and B_s clades led to our hypothesis that *AGL104* most likely functioned in the reproductive organs of the *Arabidopsis* flower. (The expression of *AGL104* in the reproductive structures of *Arabidopsis* had been confirmed earlier by Parenicova et al (2003)).

AGL- AGAMOUS-LIKE; **AP**- APETALA; **PI**- PISTILLATA; **TT**- TRANSPARENT TESTA

1.10.2 Aims

Every approach or method that is currently being utilized by the scientific community for the purpose of functional characterization has its advantages and disadvantages (Chapters 3, 4 and 5). Each method reveals specific information that can be complemented or verified by other methods. In order to successfully characterize the function of *AGL104*, we chose to perform three different types of analyses that were intended to complement the results: expression analysis of the gene, gain-of-function analysis and loss-of-function analysis. To achieve optimum results, we formulated the following aims for this study:

- Establish the specific temporal and spatial expression pattern of *AGL104* in the *Arabidopsis* flower and identify the precise structures/cells in which it is expressed. We opted to use GUS assaying method for this purpose (Chapter 3).
- Generate gain-of-function and loss-of-function mutant lines to phenotypically analyze the function of *AGL104* (Chapters 4 and 5).
- Determine the role of *AGL104* in the development of the male (pollen) gametophyte of *Arabidopsis* (Chapters 4 and 5).

The results of this study were expected to give insight into the function of *AGL104* in the gametophyte of *Arabidopsis* and allow future comparison with its role in the lower land plants, which would shed light on the evolution of MADS-box gene function and thus evolution of land plant body form.

Chapter 2

GENERAL MATERIALS AND METHODS

2.1 Media

All media were prepared using milliQ water supplied from MILLI-Q® Reagent grade water system (MILLIPORE®, Billerica, MA USA) and sterilized by autoclaving (120 °C and 15 psi, for 20min). Solid media (e.g. LB-agar and MS-phyta agar plates) were cooled to 45 °C before antibiotics were added. Liquid media (e.g. LB-broth) were cooled to room temperature before the addition of antibiotics. The final concentration of the antibiotics in both solid and liquid media was 1X that is, 50ug/ml Ampicillin, 100ug/ml Spectinomycin, 50ug/ml Kanamycin, 20ug/ml Hygromycin, 50ug/ml Rifampicin,

2.1.1 Luria-Bertani (LB) broth and agar plates (for selection of resistant *E.coli* colonies)

LB media was prepared following the recipe in Molecular Cloning Laboratory manual (Sambrook J. 2001). LB-broth contained 1g tryptone, 0.5g yeast extract (MERCK, Frankfurter Str. Darmstadt, Germany) and 0.5g NaCl per 100ml media. The pH was adjusted to 7.5 using 1M NaOH. LB – agar media contained additional 1.5g bactoagar per 100ml media

2.1.2 Murashige Skoog (MS) phyta agar plates (for selection of resistant *Arabidopsis* seedlings)

MS agar plates were prepared following recipe provided by The Arabidopsis Information Resource (TAIR) web page (http://www.arabidopsis.org/). The MS plates contained 0.44g MS salt and 1g sucrose per 100ml of media. These were dissolved in milliQ water and the pH was adjusted to 5.6-5.8 using 1M KOH. Then 0.8g of phyta agar per 100ml of the media was added and autoclaved. The appropriate antibiotic was added to the cooled (45°C), autoclaved MS-phyta agar solution and poured into plates. Plates were stored upside down at 4°C. The antibiotic, Hygromycin (final concentration 20ug/ml), was used when *A.tumefaciens* contained the pCAMBIA binary vector and Kanamycin (final concentration 50ug/ml) was used when pART27 was used as a binary vector.

2.2 Buffers and solutions

MilliQ water was used to make solutions and autoclaved. In the instance where solutions were not autoclaved before use, autoclaved milliQ water was used. Unless otherwise stated, all the buffers and solutions were prepared according to Molecular Cloning Laboratory manual recipe (Sambrook J. 2001). See Appendix 2 for the recipe of buffers and solutions.

2.3 Bacteria

2.3.1 *E.coli*

2.3.1.1 Making competent E.coli *cells* (Sambrook J. 2001)

E.coli cells were cultured in a liquid media (LB-broth) for 3-4 hours or until the OD_{600} of 0.3-0.5 was reached. The cells were centrifuged (Sorvall GSA rotor) at 8000 rpm, 4°C for 10min. The supernatant was discarded and the pellet was washed with 10ml of cold 100mM calcium chloride solution and resuspended in 20ml of cold 100mM calcium chloride. The cells were centrifuged again under same condition, the supernatant discarded and the pellet was resuspended in a cold solution containing 250µl of 80% glycerol and 750µl of 100mM calcium chloride. The heat shock competent *E.coli* cells were snap frozen using liquid nitrogen and stored in 50µl aliquots at -80°C and thawed before use in transformation.

2.3.1.2 Strains and growth conditions

The *E.coli* strain used for transformation in the lab was XL1-Blue. Unless stated otherwise, *E.coli* was grown on LB-agar plates (solid) or LB-broth (liquid culture) (section 2.1.1) with an appropriate antibiotic (final concentration of 1X) for selection and incubated at 37°C for 16-18 hrs. Liquid culture was incubated tilted at an angle on a shaker at 200 rpm.

2.3.1.3 Transformation of E.coli through heat shock and plating

All transformation of the *E.coli* was done using the heat shock method (Sambrook J. 2001). 50µl of the competent *E.coli* cells were thawed on ice. Then the ligation product (section 2.5.1) (usually 15-20µl) was added to the cells and left on ice for 20 min. The cells were then placed

in a 42°C water bath or dry bath for 90 sec. The cells were transferred to ice for 1 min before 1ml of autoclaved LB-broth was added to it. This mixture was then incubated at 37°C for at least 1h to enable the *E.coli* cells to recover. After the recovery period, the transformed *E.coli* cells were spun down in the centrifuge at high speed for 20 sec so that the cells formed a loose pellet at the bottom of the Eppendorf tube. The LB supernatant was pipetted out leaving only about 300µl in the tube. Then the cells were re-suspended by gentle vortexing. These re-suspended cells were then plated on LB-agar plates containing an appropriate antibiotic, which depended on the antibiotic resistant gene present in the vector that was used to transform the *E.coli*.

2.3.1.4 Blue-white selection

Where appropriate, the blue-white selection was used to distinguish antibiotic resistant *E.coli* colonies which were transformed with just the vector from the colonies transformed with the recombinant vector (vector + insert) i.e. the successful ligation product. However, this was only possible for vectors that had the lacZ operon (e.g. pGEM-Teasy, pBSKS II, pART27). This operon is designed such that when the plasmid is transcribed, the product of the lacZ gene, an enzyme is able to utilize X-gal and IPTG to give a blue coloured product. If the ligation is successful (i.e. the insert is successfully incorporated into the vector), the lacZ operon is disrupted (since the multiple cloning site (MCS) is engineered within the lacZ operon and thus the insert is incorporated within the lacZ gene, thus making it non-functional) and the *E.coli* is not able to make the lacZ enzyme to enable it to use X-gal and IPTG to give the final blue product. However, if the ligation is not successful and the digested ends of the vector re-ligate without incorporating the insert DNA, then the lacZ operon is still functional and thus *E.coli* transformed with just the empty vector can result in a blue colony when the *E.coli* grows on the plates. Thus, for each transformation where a lacZ vector was used, blue-white selection method was used to differentiate between *E.coli* transformed with the successful (vector + insert) and unsuccessful (empty vector) ligation products. Before the transformed *E.coli* was plated, 40µl of 2% X-gal and 10µl of 20% IPTG (Sambrook J. 2001) was spread over the plate and placed at 37°C for 5 min to enable the above two to chemicals to soak into the agar. The transformed cells were then placed on the plates and incubated as usual (section 2.3.1.2).

.

2.3.2 *A.tumefaciens*

2.3.2.1 Making electro competent A.tumefaciens *cells* (Diethard Mattanovich et al., 1989)

A single colony of *Agrobacterium* was inoculated in LB media containing Rifampicin (10ug/ml) at 30°C for 2 days. 2 ml of the pre-innoculum was added to 200ml fresh LB medium containing 10ug/ml Rifampicin and cultured for 12 – 16 hours until OD_{600} reaches 0.9. Cells were then harvested by centrifuging (Sorvall GSA rotor, Life Technology™) at 8000 rpm, 4°C for 10minutes. The supernatant was discarded and the pellet was washed with 25ml water. Centrifugation and washing was repeated 3 times to remove salt precipitates and finally cells were resuspended in 1ml 20% glycerol. The electro-competent Agrobacterium cells were then snap frozen with liquid nitrogen and stored in aliquots of 100µl at -80°C and thawed before use in transformation.

2.3.2.2 Strains and growth conditions

The *A.tumefaciens* strain used was LBA4404 (Invitrogen). Unless stated otherwise, *A.tumefaciens* was grown on LB-agar plates or LB-broth (section 2.1.1) under appropriate antibiotic selection and incubated at 30°C for 48 hrs (16 hrs if used for transformation of plants). Liquid culture was incubated on a shaker at 300 rpm.

2.3.2.3 Transformation of A.tumefaciens through electroporation (Diethard Mattanovich et al., 1989)

All transformation of *A.tumefaciens* was done using electroporation. 100µl of the electro-competent *A.tumefaciens* cells were thawed on ice. 500ng-1µg of plasmid DNA was added to the *A.tumefaciens* cells and mixed thoroughly by pipetting. The mixture was placed inside the special cuvette between the two steel plates and placed on ice for 5-10min. The cuvette was carefully wiped dry and placed in the holder of the PULSE CONTROLLER electroporation machine (Bio-Rad, Albany Auckland New Zealand). A 2.5V current was passed through the cells to make them electro-competent. 1ml of LB-broth was added to the cuvette and the mixture was transferred to an Eppendorf tube and allowed to recover at 30°C for 2-3hrs before plating on an appropriate antibiotic plate.

2.3.2.4 Preparation of A.tumefaciens before transformation into plants

A 5ml culture with appropriate antibiotics was made from glycerol stock (Appendix 2) and incubated at 30°C for 48hrs on a shaker. The culture was poured into a 2.5l flask containing

500ml of LB-broth and incubated for 12-16hrs (overnight) for transformation into plants. On the day of plant transformation, the overnight culture was centrifuged at 6000 rpm for 10min and the pellet was suspended in the infiltration media (Appendix 2), ready for use.

2.4 Nucleic Acid extraction (DNA or RNA)

Where applicable, standard molecular biology protocol was used for all nucleic acid extractions (Sambrook J. 2001). The protocols outlined below may or may not contain variations from standard protocol used in molecular laboratories.

2.4.1 Boiling Lysis (Mini Prep) for plasmid DNA extraction (Quigley, 1981)

This DNA extraction procedure was used to extract recombinant vector DNA from *E.coli* grown in liquid culture. 1-1.5ml of an o/n culture was placed into an Eppendorf tube and centrifuged at high speed for 15-20sec to pellet the bacteria. The LB supernatant was discarded and the pellet was re-suspended in 200µl of STET (Appendix 2). 3µl of RNAse (10mg/ml) (for every 200µl of STET) was added to the cells followed by 10µl of lysozyme (10mg/ml). The tubes were inverted gently to mix. The mixture was placed in boiling water (100°C) for 1min and centrifuged at high speed (13000 rpm) for 10min. The slimy pellet was removed with a sterile toothpick and discarded. 0.7 volume of Isopropanol (at rt) was added to the supernatant to precipitate DNA. The mixture was centrifuged again at high speed for 10min and the supernatant was discarded. DNA should have formed a colorless or white pellet at the bottom of the tube. The pellet was washed with 70% ethanol, air dried and re-suspended in 20µl TE (Appendix 2) or sterile water. DNA was either used immediately for restriction endonuclease digest or stored at -20°C.

2.4.1.1 Phenol Chloroform wash and DNA precipitation

Alternatively, an equal volume of 1:1 phenol chloroform (chloroform denatures the proteins in the mixture) was added to the supernatant after removal of the slimy pellet upon boiling. Upon addition of phenol chloroform, the mixture was centrifuged at high speed for 5min and the supernatant (containing clean DNA) was moved to a new tube. The DNA was precipitated by adding 0.1 volume 3M sodium acetate (pH 5.2) and 2.5volume 100% ethanol and placing on ice for 10min. After centrifugation at high speed for 10min, the pellet was washed with 70%

ethanol, air dried and re-suspended in 20µl TE (Appendix 2) or sterile water. DNA was either used immediately for restriction endonuclease digest or stored at -20°C.

2.4.2 Plasmid DNA isolation using a Quantum Miniprep kit

Plasmid DNA extraction using a Quantum Prep Plasmid Miniprep Kit (Bio-Rad, Albany Auckland New Zealand) for automatic sequencing of the target DNA sequence was carried out following the manufacturer's instructions. An overnight culture (1.5 ml *E. coli*) was centrifuged at 13,000rpm for 1 minute and the pellet re-suspended in 200 µl of cell re-suspension solution. To the suspension, 250 µl of cell lysis solution and 250 µl of neutralization solution were added. The solution was mixed then centrifuged at 13,000rpm for 5 min to remove cell debris. The supernatant was transferred into a spin filter inserted in an Eppendorf tube. To the supernatant, 200 µl matrix solution was added and mixed thoroughly, and centrifuged at 13,000 x g for 30 seconds and the filtrate discarded. The matrix was washed twice with 500 µl wash solution and DNA in the matrix was eluted by the addition of 100 µl sterilized water to the matrix and centrifuged for 1 minute.

2.4.3 Genomic DNA extraction

700µl of DNA extraction buffer (Appendix 2) was added to every 100µg of homogenized plant tissue (section 2.7.2) and the mixture was vortexed briefly at rt to mix. Then 700µl of phenol-chloroform-isoamyl alcohol (25:24:1) was added and vortexed again. The mixture was centrifuged at 13000 rpm for 10min and the supernatant was transferred to a new Eppendorf tube. 0.1 volume 3M sodium acetate (pH5.2) and 500-700µl Isopropanol was added to the supernatant and centrifuged as above. The supernatant was discarded and the pellet was washed with 70% ethanol, air dried and re-suspended in 40µl TE (Appendix 2). Genomic DNA extractions were stored at –20°C until use.

2.4.4 Total RNA extraction

Total RNA was extracted as per protocol for TRIzol® reagent purchased from Invitrogen, Carlsbred, CA, USA. Plant tissue (about 1 g fresh weight) were ground to a fine powder in liquid nitrogen, and mixed with 1ml of Trizol. After thawing, the mixture was centrifuged at 12,000 rpm for 10 minutes at 4°C. The supernatant was mixed vigorously with 0.2 ml chloroform, kept at room temperature for 3 minutes, then centrifuged at 12,000 rpm for 15 minutes at 4°C. The RNA in the aqueous phase was precipitated by adding 0.5 ml Isopropanol

and recovered by centrifuging at 12,000 rpm for 10 minutes at 4°C. The RNA pellet was washed with 10 ml of 75 % ethanol, air-dried, and re-suspended in 20µl DEPC-treated water. RNA was stored at -80 °C until use.

2.5 Nuclei Acid manipulation (DNA or RNA)

Where applicable, standard molecular biology protocol was used for all nucleic acid manipulations (Sambrook J. 2001). The protocols outlined below may or may not contain variations from standard protocol used in molecular laboratories.

2.5.1 Ligation reaction (DNA)

DNA ligation was performed to incorporate the insert DNA (e.g. the gene) into the vector. The quantity (µg) of the vector and insert DNA used depended on the amount present in the extracted samples. Generally, an appropriate amount was used to make the ratio of the vector DNA to the insert DNA at least 1:2 in order to drive the ligation reaction, that is, to ensure that the insert DNA was incorporated into the vector. On some occasions, a number of different ligation reactions with differing ratios of vector to insert DNA (1:3 to 1:10) were set in order to ensure the success of the ligation reaction. The vector and insert DNA was combined with an appropriate amount of a 2X or 5X ligation buffer (Appendix 2) to bring the final concentration of the buffer to 1X. 1-2µl of the T4 DNA ligase (PROMEGA, Madison, WI USA) was used in the reaction depending on the amount of vector and insert DNA. The ligation mixture was allowed to stand at 14°C or at room temperature for 2-16hrs.

2.5.2 Restriction endonuclease digestion (DNA)

Restriction endonuclease digests of plasmids were performed for diagnostic purposes to determine the success/failure of ligations of vectors and inserts and the "direction" of the incorporation of the inserts into a vector. All restriction endonucleases were purchased from Roche (Roche, Indianapolis, IN, USA). Unless otherwise stated, 3µl (usually 1-2µg) of plasmid DNA was digested with 10U (1µl) of an appropriate restriction endonuclease enzyme(s). The appropriate commercial restriction enzyme buffer (Roche, Indianapolis, IN, USA) (final concentration 1X) was used. The digestion reaction mixture was incubated for 2hrs at the recommended temperature suited for the restriction endonuclease enzyme being used. Digested DNA was then run and separated by agarose (1%) gel electrophoresis (section

2.5.3). In the event where two different restriction endonuclease enzymes required two different reaction temperatures, the restriction endonuclease enzyme requiring the buffer with a lower salt concentration was used first.

2.5.3 Agarose gel electrophoresis (DNA and RNA)

Agarose gel electrophoresis was used to separate DNA or RNA following PCR, restriction endonuclease digestion and extraction. The appropriate amount of agarose was added to 0.5X TBE buffer (Appendix 2) to make a 1% (w/v) gel. The solution was heated until the agarose had completely melted (usually 60-90 sec at high power on a standard microwave). The solution was allowed to cool to 45°C before an appropriate amount of ethidium bromide was added to bring the final ethidium bromide concentration to 0.5µg/ml. The gel solution was then poured into the gel running apparatus (casting trays) with combs and cooled until it set. The gel was then covered with 0.5X TBE buffer before the DNA or RNA samples were loaded. An appropriate amount of the 10X loading dye (Appendix 2) was added to the DNA sample to reach a final concentration of 1X before the samples were loaded onto the agarose gel. The agarose gels were run at 100-110V for 1h, or until the desired separation had been achieved. The DNA was then visualized using the ethidium bromide fluorescence using a short-wavelength UV trans-illuminator and digital photographs were obtained using Quantity One software (Bio-Rad, Albany Auckland New Zealand). The size of the DNA fragments were determined by comparing the mobility of the DNA samples with that of the 1kb Plus DNA ladder (Appendix 2) that was loaded next to the samples.

When loading total RNA on an agarose gel, for every 1µl of total RNA, 0.6µl of 10X MOPS, 6µl of formamide, 1.8µl of ~37% formaldehyde and 2.6µl of sterile water was added. The mixture was placed at 65°C for 2min. Then 1.2µl of RNA dye and 1µl of 0.1mg/ml ethidium bromide was added before loading the RNA on the gel. The gel was run at 100-110V for 1h, or until the desired separation had been achieved.

2.5.4 Gel purification (DNA)

DNA (PCR or restriction endonuclease digest products) was purified after running on a 1% agarose gel for an appropriate time to achieve the desired separation of DNA sizes. The desired DNA band was cut under 366nm UV light using a scalpel. The gel pieces were placed in an apparatus (made by placing a 0.5ml Eppendorf tube, which has a small pinhead-sized hole made at the bottom and covered with glass wool, into a 1.5ml Eppendorf tube) and centrifuged

at 6000 rpm for 10min to collect the solution containing DNA. A phenol chloroform wash and DNA precipitation was carried out as above (section 2.4.1.1).

2.5.5 Nuclei Acid Quantification

2.5.5.1 Gel quantification (DNA)

The amount of DNA present in each extracted sample was quantified before ligation reactions were set. 3-4µl of both the vector and insert DNA was loaded onto separate wells in the gel and electrophoresed for ~15 minutes. The quantity of the DNA in the two samples was estimated by comparing the brightness of the bands (seen on gel). The number of molecules in each of samples was estimated using the above information and the known size of the vector and insert fragment. For example, if the vector was 4kb in length and the insert was 1kb, and the brightness of the two samples on the gel was the same, it implied that there were about 4 times as many insert molecules present in 1µl of the insert sample than in 1µl of the vector sample. This information was then used to determine the amount/ratio of vector and insert DNA used in the ligation reactions.

2.5.5.2 Fluorometer quantification (DNA)

DNA was quantified using a fluorometer (usually for sequencing). A Hoefer DyNA Quant 200 (Amersham, Greenlane Auckland New Zealand) was used. 2ml of 1X TNE (Appendix 2) containing Hoechst dye was placed in a cuvette and calibrated to zero. 2ul of a standard (containing 100ng/µl of DNA) was added to the TNE buffer and calibrated to 100. Fresh TNE was placed in the cuvette and 2µl of the sample (test) DNA was added. DNA concentration in ng/µl is recorded.

2.5.5.3 Nanodrop quantification (RNA)

RNA was quantified by measuring the absorbance at 260 and 280 nm using a Shimadzu UV spectrophotometer. Samples were usually diluted so that the absorbance reading was between 0.1 and 1.0. Good quality RNA samples should have a ratio of A260/A280 greater than 1.8. The concentration of RNA was calculated by the following equation: [RNA] (µg/ml) = 40 x A260 x dilution factor.

2.5.6 Polymerase Chain Reaction (PCR)

PCR reactions were carried out using an Eppendorf Mastercycler® (Eppendorf South Pacific, North Ryde NSW Australia). To optimize PCR reaction conditions, annealing temperatures were varied depending on the melting temperature (Tm) of the primers. Extension times of 1 min/kb were used. Mg^{2+} concentration, primer and DNA template concentration were adjusted accordingly. PCR products were analyzed by gel electrophoresis and visualized by UV trans-illuminator (section 2.5.3). Standard PCR reactions (50 µl in 0.2 ml tube) consisted of 1x *Taq* DNA polymerase buffer (Roche, Indianapolis, IN, USA), 50 µM of dNTP, 200 nM of each primer, and 0.02 unit/µl of *Taq* DNA polymerase (Roche, Indianapolis, IN USA), Taq:Pwo (9:1) DNA polymerase (Roche, Indianapolis, IN USA) or Expand™ High Fidelity polymerase (Roche, Indianapolis, IN USA). The thermocycle conditions were 1 cycle of 2 minutes at 94°C, 35 cycles of 1 minute at 94°C, 1 minute at 55–62°C (depending on the Tm of the primers) and 1-3 minutes (1 min/1 kb) at 72°C, and 1 cycle of 5 minutes at 72°C.

The primers (Appendix 1) were ordered from Gibco BRL, Sigma Corporation or Life Technologies. Each primer was re-suspended in TE buffer to a stock concentration of 500 pmol/ul and stored at -20°C. The primers were diluted to a working concentration of 10pmol/ul for PCR reactions and 3.2 pmol/ul for sequencing reactions. Appendix 1 shows a list of primers used.

2.5.7 cDNA synthesis and RT-PCR

cDNA used in PCR reactions were synthesized using either SuperScript™ III First-Strand Synthesis System for RT-PCR or the SuperScript™ III One-Step RT-PCR System with Platinum® *Taq* DNA Polymerase kits (Invitrogen Carlsbad, CA USA).

2.5.7.1 SuperScript™ III First-Strand Synthesis System for RT-PCR (Cat. No: 18080-051)

cDNA used for PCR reactions was synthesized using a kit following the manufacturer's protocol. 1µl of 10mM dNTP was mixed with 1µl of 50µM oligo$(dT)_{20}$ primer, 3-4µl of template RNA (depending on concentration, up to 5µg in total) and enough DEPC treated water to bring the final volume to 10µl. This mixture was incubated at 65°C for 5min and placed on ice for 1min after the incubation period. In the meantime, a cDNA synthesis mix was prepared by adding the following chemicals (supplied in the kit) in a separate tube: 2µl of 10X

RT buffer, 4µl of 25mM MgCl2, 2µl of 0.1M DTT, 1µl of (40 U/µl) RNase Out and 1µl of (200U/µl) SuperScript III RT. 10µl of this cDNA synthesis mix was added to the incubated mixture and collected by brief centrifugation. The mixture was then incubated at 50°C for 50min followed by incubation at 85°C for 5min to terminate the reaction. The mixture was then placed on ice. 1µl of RNase H was added to the reaction tube and incubated at 37°C for 20min. The cDNA was then either used immediately for PCR or stored at -20°C.

2.5.7.2 SuperScript™ III One-Step RT-PCR System with Platinum® Taq *DNA Polymerase* (Cat. No: 12574-018)

The RT-PCR reaction mix was prepared by mixing the following reagents supplied in the kit and the total RNA extracted from samples into a tube placed on ice: 25µl of 2X Reaction Mix, 1µl of 10µM Sense primer, 1µl of 10µM Anti-sense primer, 2µl of SuperScript™ III RT/Platinum® Taq Mix, 5-10µl of template RNA (0.01pg-1µg) and autoclaved distilled water to bring the total volume to 50µl. The mixture was then put in a thermal cycler. The following PCR program was used: 1 cycle at 45-60°C for 15-30min; 1 cycle at 94°C for 2min; 40 cycles at 94°C for 15sec, 55-65°C for 30sec, 68°C for 1min; and 1 cycle at 68°C for 5min.

2.5.8 Cloning and Subcloning DNA

Genes were cloned using the appropriate primers (Appendix 1) and PCR. The Taq® polymerase used in the PCR reaction leaves an adenosine nucleotide overhang (AA), which was used to subclone *AGL104* (and other gene fragments) into various intermediate vectors (section 2.6) depending on the type of genes/promoters required in the construct. To enable transformation of the construct into the *Agrobacterium* and eventually the plant, the constructs were subcloned into the binary vector. We used the binary vectors pCAMBIA1381Xc and pART27 for our constructs. These binary vectors incorporate any DNA fragment (insert DNA) placed between their left and right borders into the plant genome. The incorporation of the insert DNA into each vector was checked repeatedly using restriction digests (section 2.5.2).

2.5.9 Alkaline phosphatase, shrimp (Roche, Indianapolis, IN USA)

Alkaline phosphatase was used to dephosphorylate vector DNA before ligation reactions. Dephosphorylation prevents the endonuclease-digested ends from ligation together before the insert DNA is ligated into the vector. The following reagents were mixed into a Eppendorf tube placed on ice: 0.5µl (50ng) of vector DNA, 1µl of restriction endonuclease, 0.7µl of corresponding buffer, 4.8µl of double distilled water. The mixture was incubated at 37°C for

10min. After this period, the restriction endonuclease was inactivated by incubating the mixture at 65°C for 15min. 1µl of the alkaline phosphatase and 0.9µl of 10X dephosphorylation buffer was added to the mixture and incubated at 37°C for 10min. The dephosphorylated vector DNA is then used in the ligation reactions or stored at -20°C.

2.5.10 Automated DNA sequencing

DNA was sequenced at Alan Wilson Center (AWC) (www. awcmee.massey.ac.nz/) by the dideoxynucleotide chain termination method, using Big-Dye chemistry (PE Applied Biosystems, Foster City, CA, USA) and oligonucleotide primers synthesized by Invitrogen (Invitrogen, Carlsbred, CA, USA). The products were separated on an ABI Prism 377 sequencer (Perkin-Elmer, Albany, Auckland, New Zealand) for detection. Sequences were assembled into contigs using Sequencher™ Version 3.1.1 (Gene Codes Corporation, Ann Arbor, MI U.S.A).

2.6 Generation of constructs

The constructs are engineered to contain the insert DNA of interest and an appropriate antibiotic resistant gene to enable selection of *E.coli*, *Agrobacterium* and plants containing the construct only. In our study, we used the antibiotics Ampicillin (50ug/ml), Spectinomycin (100ug/ml), Rifampicin (50ug/ml) for selection of transformed *E.coli* and *Agrobacterium* while Kanamycin (50ug/ml) and Hygromycin (20ug/ml) was used for selection of transformed seeds/seedlings. Only bacteria and plants containing the construct, and thus the appropriate antibiotic resistant gene, were able to grow on media containing the antibiotic.

2.6.1 *AGL104::GUS* construct

A 1.7kb fragment of the upstream region (from the start codon) of the *AGL104* promoter was translationally fused to the coding region of the *uidA* gene that encodes the β glucuronidase (GUS) enzyme (Figure 3.1) (present in the pCAMBIA1381Xc vector). The BamHI enzyme was used in the restriction digest to obtain the promoter fragment. The initiation codon (ATG) of *AGL104* was in frame with the second codon of the GUS enzyme. The 35SCaMV promoter constitutively expressed the *AGL104::GUS* construct in the plant tissues. However, the GUS protein is only transcribed in tissues/cells where the *AGL104* promoter is active, thus

indicating the location of the endogenous expression of *AGL104*. Upon GUS staining/assaying (see below), a blue precipitate is observed at the site of the GUS enzyme activity.

2.6.2 35S::*AGL104* construct

The *AGL104* gene was cloned using PCR (section 2.5.6) using the appropriate primers (Appendix 1). The PCR product was purified using agarose gel (section 2.5.4) and the fragment was confirmed to be an exact match to the *AGL104* sequence on the database. The fragment was then subcloned into various vectors as shown in Figure 4.2. The Taq® polymerase used in the PCR reaction leaves an adenosine nucleotide overhang (AA), which was used to subclone *AGL104* into the intermediate vector, pGEM-Teasy®. The *AGL104* fragment was then subcloned into the overexpression vector, pART7 (containing the 35S promoter of CaMV). The *AGL104* gene was excised out of pGEM-Teasy® using the EcoRI restriction site and ligated into the pART7 vector that had also been digested with the same restriction endonuclease. The overexpression fragment (35S promoter + *AGL104*) was then cloned into the binary vector, pART27, using the restriction NotI enzyme (section 2.5.8). The binary vector enabled transformation of the construct into *Agrobacterium* and eventually the plant genome.

2.6.3 Double 35S (35S::35S::*AGL104*) construct

The *AGL104* fragment was excised out of the overexpression construct (pART7+*AGL104*) generated earlier. The XhoI and SmaI restriction endonucleases were used. This fragment was ligated to the pBINJIT (containing two 35S promoter regions) vector that was digested with SalI and SmaI. Since XhoI and SalI are compatible sites, the *AGL104* fragment was able to ligate to pBINJIT.

2.6.4 dsRNAi construct

The full fragment of *AGL104* was cloned into the vector, pRNA69, in the antisense direction using the EcoRI restriction digest site. Then the IKC region of *AGL104* (i.e. *AGL104* gene fragment without the MADS-box domain) was cloned using PCR and the XbaI restriction site was engineered into this fragment. Using the XbaI restriction site, the IKC fragment was cloned in the sense direction into the pRNA69 vector that already contained the *AGL104* in the antisense direction, thus completing the dsRNAi construct. The fragment containing both the *AGL104* and the IKC region was then subcloned using the NotI restriction site into the binary vector, pART27, which enables transformation into *Agrobacterium*.

2.7 *Arabidopsis* whole plant and plant tissue preparations

For wild type (wt) *Arabidopsis* plants, the Colombia ecotype was used. T-DNA insertion lines for *AGL104* (SALK 013898) was obtained from SALK institute (www.signal.salk.edu/).

2.7.1 Plant growth conditions

Arabidopsis plants were grown under constant light at 18°C. The plants were watered twice a week to ensure soil was kept constantly moist according to instruction given by TAIR (http://www.arabidopsis.org/). The *Arabidopsis* generation time is approximately 4-6 weeks (from seeds to adult flowering plants bearing dehisced siliques).

2.7.2 Homogenization of plant tissue (Ichiro Kasajima et al., 2004)

For nucleic acid extraction, 2 to 3 rosette leaves were harvested and snap frozen in liquid nitrogen. An autoclaved mortar and pestle (or Eppendorf tube for small quantities) was used. A styrofoam base was placed under the mortar to prevent it from warming to rt. Liquid nitrogen was poured on the mortar and pestle and allowed to evaporate 3 times in order to cool them. The tissues were placed in the cold mortar, 5-10 ml of liquid nitrogen was added to the tissues (enough to cover the tissues completely) and ground with the pestle to a fine powder. The homogenized tissue was stored in batches of 100µg at -80°C or used immediately for DNA or RNA extraction.

2.7.3 Emasculation

Mature plants were emasculated in order to facilitate generation of multiple mutants or *in vivo* pollen tube growth experiments. Flowers at stage 10 or 11 (Smyth et al., 1990) of development are lightly pressed/constricted at the base of the sepals in order to expose the floral organs. Caution is exercised not to apply too much pressure as this would damage the flowers. With a sharp pair of forceps, the anthers are carefully removed without damaging the other organs of the flower, especially the stigma and ovary. In instances where the pollinated flowers were used for *in vivo* pollen germination studies, sepals and petals were also removed to facilitate better observation under the microscope.

2.7.4 Pollination

Emasculated flowers were pollinated after 1-2d once the flower has reached stage 12 of flower development (Smyth et al., 1990). This is when the stigmatic papillae had fully developed and is receptive to pollen. Freshly dehisced anthers containing mature, viable pollen were picked from flowers at stage 13 of development (Smyth et al., 1990) and gently rubbed on the stigmatic papillae of the live emasculated flowers. The procedure was carried out under the dissecting microscope. When generating multiple mutant lines (below), the pollinated stigmas were allowed to develop to maturity and the seeds were collected. When pollinations were done for the purpose of *in vivo* pollen tube growth studies, the pollinated stigmas were "harvested" at 2hr and 4hr time intervals and fixed (section 2.9.2).

2.7.4.1 Generating multiple mutants

When generating double mutant lines, we emasculated one of the single overexpressing lines and pollinated it with pollen from its corresponding single overexpressing line i.e. if we emasculated a 35S::*AGL66* flower, then we pollinated its stigma with 35S::*AGL104* pollen (or vice versa) to generate the 35S::*AGL66/104* double mutant. The 35S::*AGL67/104* line was generated in a similar manner. When generating triple mutant lines, a double mutant (containing any two overexpressing constructs e.g. 35S:: *AGL67/104*) stigma was pollinated with pollen from the third single overexpressing mutant line (35S::*AGL66*). The pollinated stigmas were allowed to grow into mature siliques and the resulting seeds were selected in antibiotic plates (section 2.8.3). The generation of the plants (T_1, T_2, T_3 etc) used for the crosses was not taken into consideration since all lines used for crosses contained the desired construct(s) as confirmed by RT-PCR.

2.7.5 Embedding plant tissues

In order to embed plant tissue in paraplast to facilitate sectioning, the tissues were prepared as follows. The tissues were first harvested and chemically fixed in FAA solution (Section 2.3.3) overnight at 4°C. Then the tissues were passed through an ethanol series of 50%, 70%, 85%, 95% for 1-1.5hr each. The tissues were placed in 100% ethanol for a total of 4.5hr, changing the 100% ethanol every 1.5hr. The tissues were then passed through different concentrations of xylene (mixed with ethanol since xylene does not mix with water): 25%, 50%, 75% for 1hr each and finally in 100% xylene for a total of 3hr, changing 100% xylene every 1hr.The tissues were placed in a container holding half volume of 100% xylene and half volume of paraplast

chips. The container was placed in a 60°C oven to melt the paraplast. Tissues were moved to 100% melted paraplast and the paraplast was changed several times until all traces (smell) of xylene were removed. The tissues were then embedded in paraplast on the embedding machine.

2.7.6 Sectioning

A 8-10μm thick sections of the embedded tissues were made using the microtome. The sections were placed in 42°C water bath and mounted on slides. The slides were dried overnight on a warmer and the paraplast was removed using histoclear or xylene before staining the tissues or before observing under the microscope.

2.7.7 Plant transformation

Plants were prepared for transformation using the floral dip method (Bent, 1998). Bolting plants (~ 3 weeks old) were decapitated 5-7d before the actual transformation in order to allow more branches and flowers to form. The plants were watered a day prior to the transformation to prevent the soil from soaking the infiltration solution (Appendix 2). On the day of the transformation, the plants were clipped to remove all siliques and flowers older than stage 16 of development (Smyth et al., 1990). This is done to minimize non-transformed seeds in the batch. The plants (still in the pots) are placed upside down on a wire net to allow flowers to be dipped in the infiltration solution containing the re-suspended *A.tumefaciens* with the desired construct. The plants are dipped for 45min. The plants are then placed on their sides (since they are too weak to stay upright) on a tray, covered with plastic (to maintain humidity) and left out of direct light for 12-36hrs. The transformed plants are then placed upright in the growth light and allowed to grow/develop normally. Seeds are collected as usual from dried siliques.

2.8 *Arabidopsis* seeds and seedlings

2.8.1 Seed germination

Seed germination protocol was carried out following TAIR's instruction (http://www.arabidopsis.org/). Seeds were placed in an Eppendorf tube with water and left at 4°C o/n. This allows stratification and vernalization of seeds so that they germinate at the same time. The seeds are then placed on the surface of the soil. Usually 5-10 seeds are placed in a

single pot. Humidity was maintained by placing the pots on a tray of water and covering with a plastic bag. The pots are placed under the growth light and grown under condition mentioned in section 2.7.1 above. The plastic bags are removed when the germinated seedlings has produced at least 2-3 true leaves, usually after 5-7d. The soil is kept moist by adding water to the base of the tray.

2.8.2 Sterilization and plating of seeds

The seeds were sterilized in a laminar flow following protocol of TAIR (http://www.arabidopsis.org/). 1ml of seeds was vernalized at 4°C o/n. The seeds are then washed in 50ml of 70% ethanol for 5min and then in 50ml of 40% bleach for 10min to surface sterilize. The seeds are washed several times with autoclaved water. The seeds are divided into 3 portions and 30ml of cool but molten 0.3% autoclaved agar is added to each portion. 10ml of the agar containing the seeds are plated on MS-phyta agar (section 2.1.2) plates, which had the appropriate antibiotic for selection of resistant plants. The MS-phyta agar plates are sealed with parafilm and placed under the growth light and allowed to germinate.

2.8.3 Selection of resistant seedlings and transplanting

Resistant seedlings (containing the construct) can be detected after 7-10d after plating following protocol of TAIR (http://www.arabidopsis.org/). The roots of the resistant seedlings are able to grow and elongate along the base of the plates (rather than on the surface of the agar) by penetrating the agar whereas non-transformed seedlings are killed by the antibiotics included in MS agar. The seedlings are carefully removed from the agar, the roots are lightly washed with water to remove traces of the antibiotic and transferred to soil. The seedlings are allowed to grow as normal.

Occasionally, it is possible for non-transformed seedlings to "escape" the antibiotic by growing on top of other resistant seedlings. In addition, resistant seedlings may contain only one copy of the construct and may segregate in the next generation. Therefore, even though antibiotic selection is a good indicator of transformed seedlings, it is recommended that the presence of the construct/gene of interest is checked in each generation of plants before they are utilized for phenotypic analysis. This can be achieved through RT-PCR.

2.9 Pollen

2.9.1 Pollen germination *in vitro*

Mature pollen for each mutant line (approximately 8 – 10 individual plants per line) were collected and germinated *in vitro* in the pollen germination media and examined after 24hrs under the Light Microscope at 10X. 3 slides were prepared for each tube containing the germinated pollen. Photographs of an average of 20 different fields of view were taken for each slide. The number of germinated and ungerminated pollen was counted and recorded. Pollen with any visible protrusion and fully elongated pollen tubes were considered "germinated".

2.9.2 Pollen germination *in vivo*

In vivo pollen germination experiments were carried out to determine the rate of germination (length of the pollen tube at certain time interval). Pollinated pistils (section 2.7.4) were removed from the plants at 2- and 4-hap and fixed in 3:1 ethanol:acetic acid for 2hrs at rt. The pistils were then washed three times with distilled water and allowed to stand in 8M NaOH for 8-10hrs or o/n (for clearing the tissues). The pistils were finally washed with distilled water and stained with Analine Blue (section 2.10.3).

Since the length of individual tubes could not be measured separately, data on pollen tube length was collected by measuring the tubes from the stigma to the longest visible pollen tube. We set multiple pollinations for each mutant type. Therefore, the lengths of pollen tubes reported represented the average length of pollen tubes in all pistils of a particular mutant type.

2.10 Staining

2.10.1 GUS

Plant tissue containing the GUS construct were prefixed in 90% cold acetone and allowed to stand for 20min at rt. The tissues were washed with cold water and placed in cold GUS staining solution (Appendix 2) and infiltrated on ice till the tissues sank. In the instance where the tissue did not sink after 2hrs, the tissues were placed under vacuum to allow infiltration. The tissues were then incubated in the dark with the same GUS staining solution at 37°C o/n or

until GUS staining could be visualized. Care was exercised not to incubate for too long since it could lead to unspecific staining. GUS staining of pollen and inflorescence required 8-10h of incubation and ovules required ~48h. The tissues were then passed through an ethanol series (15%, 30%, FAA, 70%, 85%, 95%, 100%) to dry. FAA was used to postfix. Tissues were only dried till 70% ethanol if they were observed as a whole mount. Tissues were passed through the ethanol series till100% and then passed through different concentrations of xylene if they were intended for embedding and sectioning (section 2.7.5 and 2.7.6). The whole mount was viewed under a dissecting microscope (Leica MZ12) while immersed in ethanol to prevent drying. Sections were viewed under a normal light microscope (Olympus BX51).

2.10.2 DAPI

DAPI stain was used to visualize the number of nuclei in the pollen to determine stage of pollen development. 200µl of the DAPI stock solution (0.1mg/ml) was mixed with 49.8ml of the DAPI buffer to make a working solution with a DAPI concentration of 0.4µg/ml. Sections of *Arabidopsis* flowers on a slide was cleared of the paraplast by dipping twice in 100% xylene or histoclear for 5min or until the paraplast washed off. Once the paraplast was removed, the slides were serially dipped in 100%, 95%, 70%, 50% and 30% ethanol for 2min each. The slides were then washed in distilled water. Finally, the slides were placed in the DAPI working solution for 10min in the dark. The slides were drained, the sections were covered with a drop of fluorescence mounting media, then covered with a coverslip and sealed with nail polish. The slides were viewed under a fluorescence microscope (Olympus BX51) using a DAPI filter. The DAPI dye is excited at a wavelength of 372nm and emits a blue fluorescence at 456nm. DAPI stained slides could be stored at 4°C for up to a week in the dark before the dye faded.

2.10.3 Analine Blue

Pollen tubes were stained with Analine Blue to facilitate *in vi*vo pollen tube growth studies. The cleared pistils were stained in 0.1% Aniline Blue solution (Aniline Blue dissolved in 0.1M K_2HPO_4-KOH buffer (pH 11)) for 3-5hrs in the dark. The pistils were then placed on a slide, covered with a drop of glycerol and observed under a fluorescence microscope (Olympus BX51) at 10X magnification using a DAPI filter. The Analine Blue dye is excited at 330-385nm and emits light at 420nm. The blue fluorescence faded after ~90sec.

CHAPTER 3

GUS EXPRESSION ANALYSIS OF *AGL104*

3.1 Background

Even though the spatial and temporal expression of a particular gene is not necessarily an indication of its function in a tissue, it is noted that all well-characterized MADS-box genes are predominantly expressed in the tissues in which they function (Becker and Theissen, 2003; Goto and Meyerowitz, 1994; Gu et al., 1998; Jack et al., 1992; Liljegren et al., 2000; Mandel et al., 1992; Sommer et al., 1990; Theissen and Saedler, 1995; Trobner et al., 1992; Weigel et al., 1992; Yanofsky et al., 1990). Since *AGL104* is a MADS-box gene, it is reasonable to assume that it would function in the tissues in which it is expressed. Therefore, the first step towards the functional characterization of *AGL104* was to determine its domain of expression. A number of methods are available for determining the expression pattern of genes. We chose to use the popular method of GUS assaying for our purposes.

The *uid*A gene, which encodes the β glucuronidase (GUS) enzyme was initially developed as a gene fusion marker in the bacterium, *E.coli*, and the nematode, *C.elegans*, but has more recently been used extensively to uncover the spatial gene expression pattern in plants (Bottino, 2005). The coding sequence for the GUS protein is placed under the direction of the regulatory sequences of another gene, usually the promoter of the gene under investigation. Provided that all the necessary regulatory elements are present in the promoter sequence, the GUS gene is expressed (transcribed and translated) in the tissues where the promoter is active, thus indicating the location/tissues where the gene under investigation is endogenously expressed. The GUS enzyme acts on a substrate and results in a colored precipitate. The best substrate currently available for the histochemical localization of GUS enzyme activity in tissues and cells is 5-bromo-4-chloro-3-indolyl glucuronide (X-Gluc), which produces a blue precipitate at the site of its activity. Thus, upon assaying for GUS activity, the spatial expression of the gene under investigation can be visually deciphered. Since the GUS enzyme

is very stable and is able to tolerate many detergents and widely varying ionic conditions (Bottino, 2005), GUS assaying has become a very useful tool in the study of gene expression.
A second method, *in situ* hybridization, is also commonly used in gene expression studies. It detects the presence of endogenous mRNA transcripts of the gene in question. Tissues are prepared into thin slices and probed with an antisense transcript (labeled with DIG or radio-active elements), which then hybridizes to the endogenous mRNA transcripts present at the location/tissues where the gene is expressed. Thus, *in situ* hybridization and the GUS assays are examples of two different methods commonly used for determining the temporal and spatial expression pattern of genes.

There are a number of other methods that determines the location/tissues where genes are expressed. Microarray/transcriptome analysis is one of these methods. It detects the presence of mRNA transcripts in sample tissues and upon comparison with transcripts in other (control) tissues, changes in the number of and type of transcripts is determined, leading to information on which genes are expressed in a particular tissue or how the expression pattern of these genes change between different tissue types or between different developmental stages. Thus, Microarray analysis provides information on the temporal expression pattern of genes. Northern Blot analysis is another method that detects endogenous mRNA transcripts in a tissue. It provides information on the presence/absence and level (quantity) of expression of the gene of interest in a particular tissue.

Prior to our study, *in situ* hybridization analysis had shown that *AGL104* was expressed in the reproductive structures of the *Arabidopsis* flower, particularly in the inside of whorl 1/sepal of early, stage 3 flowers; in young developing anthers; petals; and within the septum and developing ovules of the carpel in older flowers (Parenicova et al., 2003). Thus, the results of Parenicova et al. (2003) provided information on the **spatial** expression pattern of *AGL104*. To further our knowledge on the **temporal** expression pattern of *AGL104* in the *Arabidopsis* flower and to determine the precise tissues in which it was expressed, an *AGL104::GUS* construct was generated by fusing a 1.7kb fragment from the upstream region of the *AGL104* promoter to the coding region of the GUS protein (Figure 3.1). Assuming that all the necessary regulatory elements required for the expression of *AGL104* were present in the 1.7kb fragment, the GUS protein was expected to be expressed in the tissues where *AGL104* is endogenously expressed and thus, upon GUS assaying, these tissues could be identified. The *AGL104::GUS*

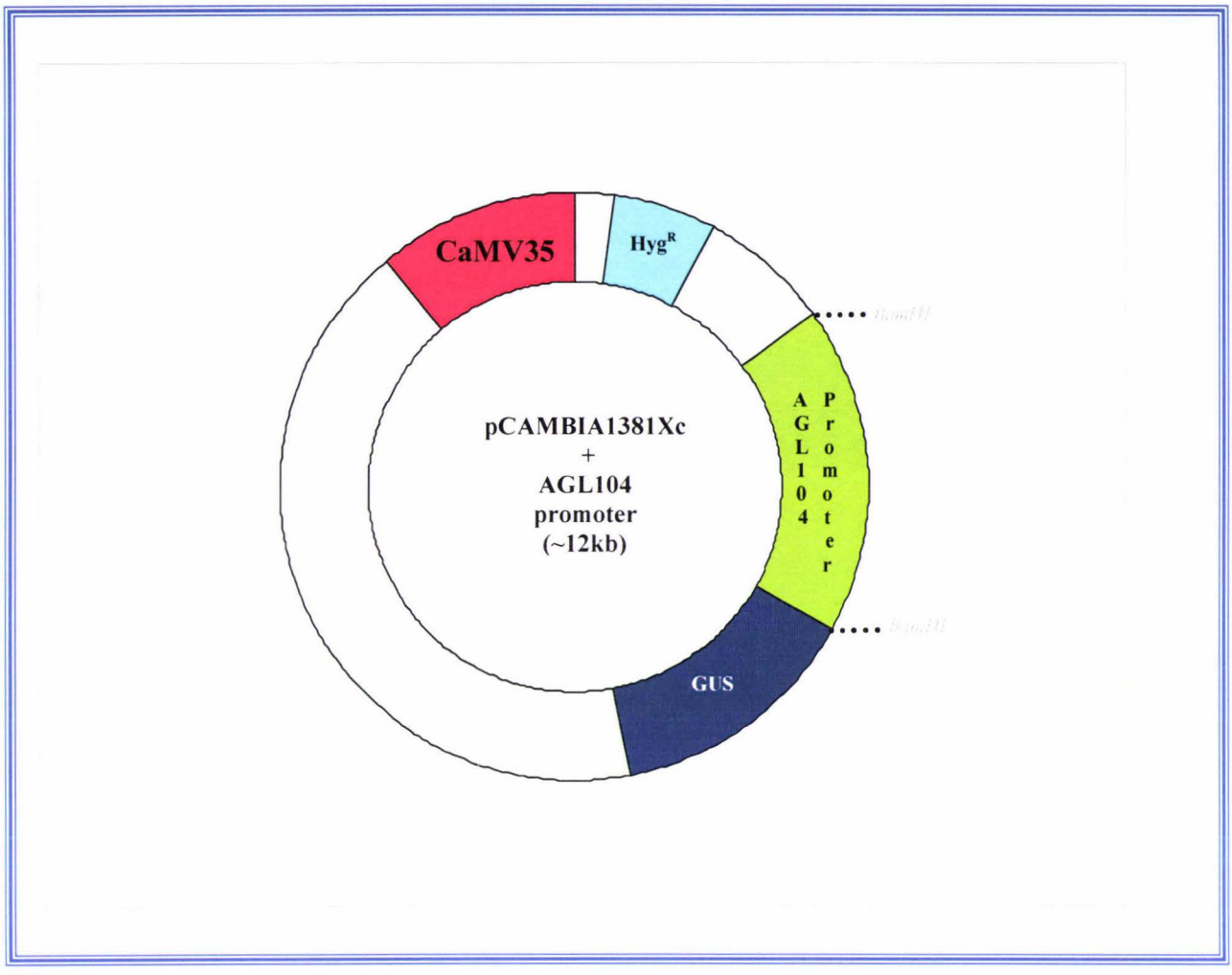

Figure 3.1 ***AGL104::GUS*** **construct** (Prepared by K. Prasad).

The 1.7kb fragment of the upstream region of the *AGL104* promoter was translationally fused to the coding region of the GUS protein. The GUS protein was synthesized in cells/tissues in which the *AGL104* promoter was endogenously active.

GUS- *uidA* gene encoding the β glucuronidase (GUS) enzyme; **Hyg^R**- hygromycin resistance; **CaMV35S**- constitutive promoter from Cauliflower Mosaic Virus; **kb**- kilo bases; **BamHI**- restriction enzyme.

Hyg^R- used for selection of transformed plants.

construct was transformed into wt Arabidopsis plants (Columbia eco-type) and the transformants were assayed to determine the expression pattern of *AGL104*. A total of fifteen independent lines were generated, all of which had a similar expression pattern. Out of these, three representative lines were used for detailed analysis. These were designated *AGL104GUS-1*, *AGL104GUS-2*, *AGL104GUS-3*.

3.2 Results

3.2.1 *AGL104::GUS* Construct

The 1.7kb promoter of *AGL104* was translationally fused to the coding region of the GUS protein (Figure 3.1) so that the initiation codon (ATG) of *AGL104* was in frame with the second codon of the GUS enzyme. This construct was prepared in our lab by Dr. K. Prasad and analyzed *in planta* by myself (see section 2.6.1 for details on generating the construct).

3.2.2 Phenotypic Analysis

3.2.2.1 GUS expression in young developing anthers and mature pollen of the Arabidopsis flower.

GUS assaying of whole inflorescences revealed the expression of *AGL104::GUS* in developing anthers and mature pollen but not in the meristem region or in flowers younger than stage 12 of development (Figure 3.2). GUS expression was most prominent (very intense staining) in the mature pollen grains (Figure 3.2 A, B, D, E), particularly in pollen of flowers older than stage 12 of development (Smyth et al., 1990). In these flowers, the petals were open indicating that anthesis had occurred, thus the pollen were mature. Pollen grains could be seen as individual stained spots in the whole mount inflorescences (Figure 3.2 A, B). GUS expression was also observed in developing anthers of younger flowers that had not undergone anthesis (Figure 3.2 A). GUS expression was not observed in the meristem region (Figure 3.2 C) or in flowers younger than stage 12 of development (Figure 3.2 D). In stage 12 flowers, the petals are as high as the long stamen and the sepals still enclose the inner floral organs (Smyth et al., 1990).

The mutants plants used for GUS assaying were segregating and therefore not all pollen contained the *AGL104::GUS* construct. This resulted in some mature pollen not showing GUS

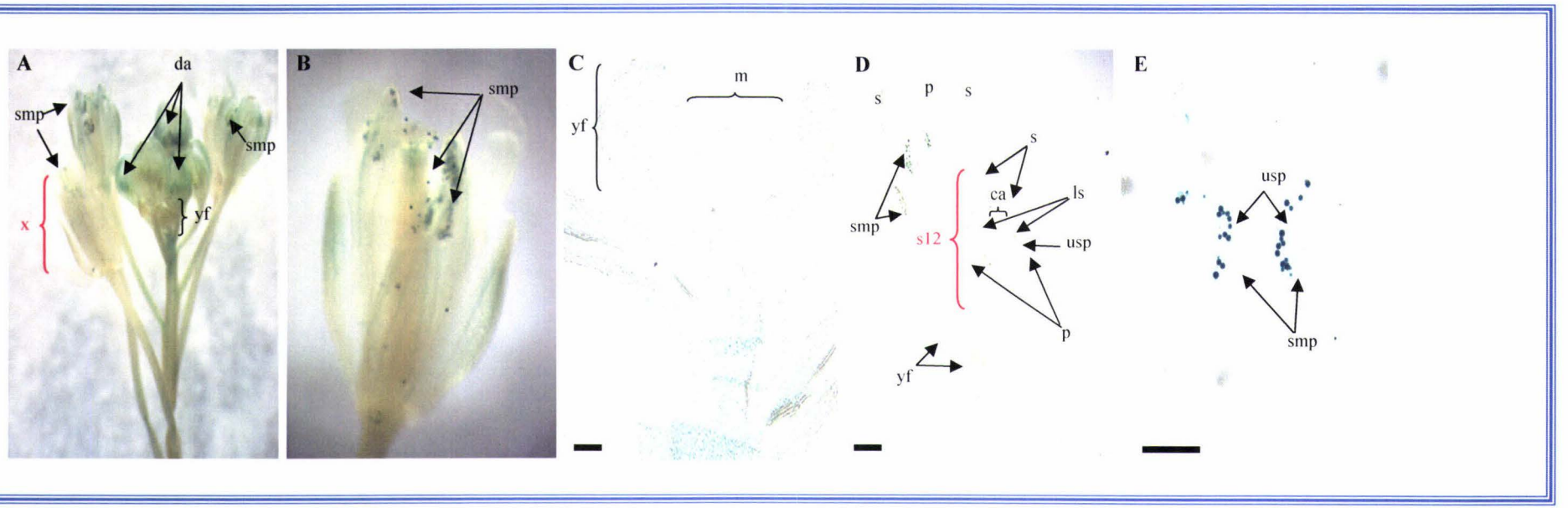

Figure 3.2 Histochemical localization of *AGL104::GUS* expression in developing anthers and mature pollen.

A: Whole inflorescence showing GUS expression in developing anthers and mature pollen. No expression is observed in anthers of young flowers. **B:** Close-up of a single flower showing GUS expression in mature pollen (petals are open and thus anthesis has occurred at this stage, implying that the pollen are mature). **C:** Cross section of an inflorescence showing the meristem (10X). There is a lack of GUS expression in the meristem region and very young flowers (as seen in A). **D:** Cross section of an inflorescence showing flowers at difference stages of development (10X). GUS expression is absent in young and stage 12 flowers (petals are as high as the long stamen and sepals are closed) and comes on in older flowers with the petals and sepals open. **E:** Cross section of an anther showing GUS expression in mature pollen (20X). Unstained pollen is also seen since the plants used for GUS assaying were segregating and thus not all pollen contained the *AGL104::GUS* construct.

In the whole inflorescence, GUS expression is observed in developing anthers and mature pollen only. GUS expression is absent in the meristem region and from anthers of very young flowers. Pale blue coloring/staining seen in the veins of sepals, petals and stems in panels A and B are believed to be due to unspecific staining or lack of proper washing since this staining is not observed in all the inflorescences examined and is very pale (see "x" in panel A). The inflorescences and individual flowers shown are representative of the fifteen GUS lines generated.

ca- carpel; **da**- developing anthers; **G**- GUS stained; **ls**- long stamen; **m**- meristem; **p**- petal; **s**- sepal; **s12**- flower at stage 12 of development; **smp**- stained mature pollen; **usp**- unstained pollen; **yf**- young flower(s); x- flower with no staining in the sepal and petal veins or stems.

Scale bar = 100μm

expression/staining while others in the same anther were clearly stained (Figure 3.2 D, E). Pale blue coloring was observed in the veins of sepals and petals and in the stems of some flowers. Since this staining pattern was not observed in all the flowers (Figure 3.2 A) or all mutant lines examined, it was concluded that this coloring was due to lack of proper washing or unspecific staining and not represent *AGL104* expression in the sepals, petals and stems of *Arabidopsis* flowers.

3.2.2.2 GUS expression correlated with the stage of pollen development/number of nuclei

AGL104::GUS expression was observed at a specific stage of pollen development. The developmental stages of pollen is divided into four phases: uninucleate microspores, bicellular pollen, immature tricellular pollen, and mature pollen grains (Honys and Twell, 2004). To determine the developmental stage at which *AGL104::GUS* expression was first detected, the GUS-stained flowers were sectioned and stained with DAPI (4, 6- diamidino-2-phenyl-iodole), a florescent dye that stains the DNA in the nucleus.

Our results showed that *AGL104::GUS* expression was first detected (very weak expression) at the uninucleate (1-nucleus) stage of pollen development (Figure 3.3 A, B) and continued through to the bi- and tri-cellular stages (2- and 3- nuclei, respectively) (Figure 3.3 C, D, E, F). *AGL104::GUS* expression was stronger in the latter two stages since the staining was more intense (Figure 3.3 C, E). Therefore, in agreement with Honys and Twell (2004), our results indicated that *AGL104* was expressed very early during pollen development, increased in expression as pollen development progressed and was maintained at a high level in the mature pollen grains.

As mentioned above, the mutants plants used for GUS assaying were segregating and therefore not all pollen contained the *AGL104::GUS* construct, thus resulting in some pollen grains not showing GUS staining (Figure 3.3 A, C, E). Due to the difficulty in detecting florescence from the nuclei of GUS-stained pollen (the blue precipitate/GUS stain may have interfered with florescence detection), the number of nuclei in neighboring pollen of the same anther were used as a guide to determine the developmental stage of a particular pollen since all pollen in a particular anther are always at the same developmental stage.

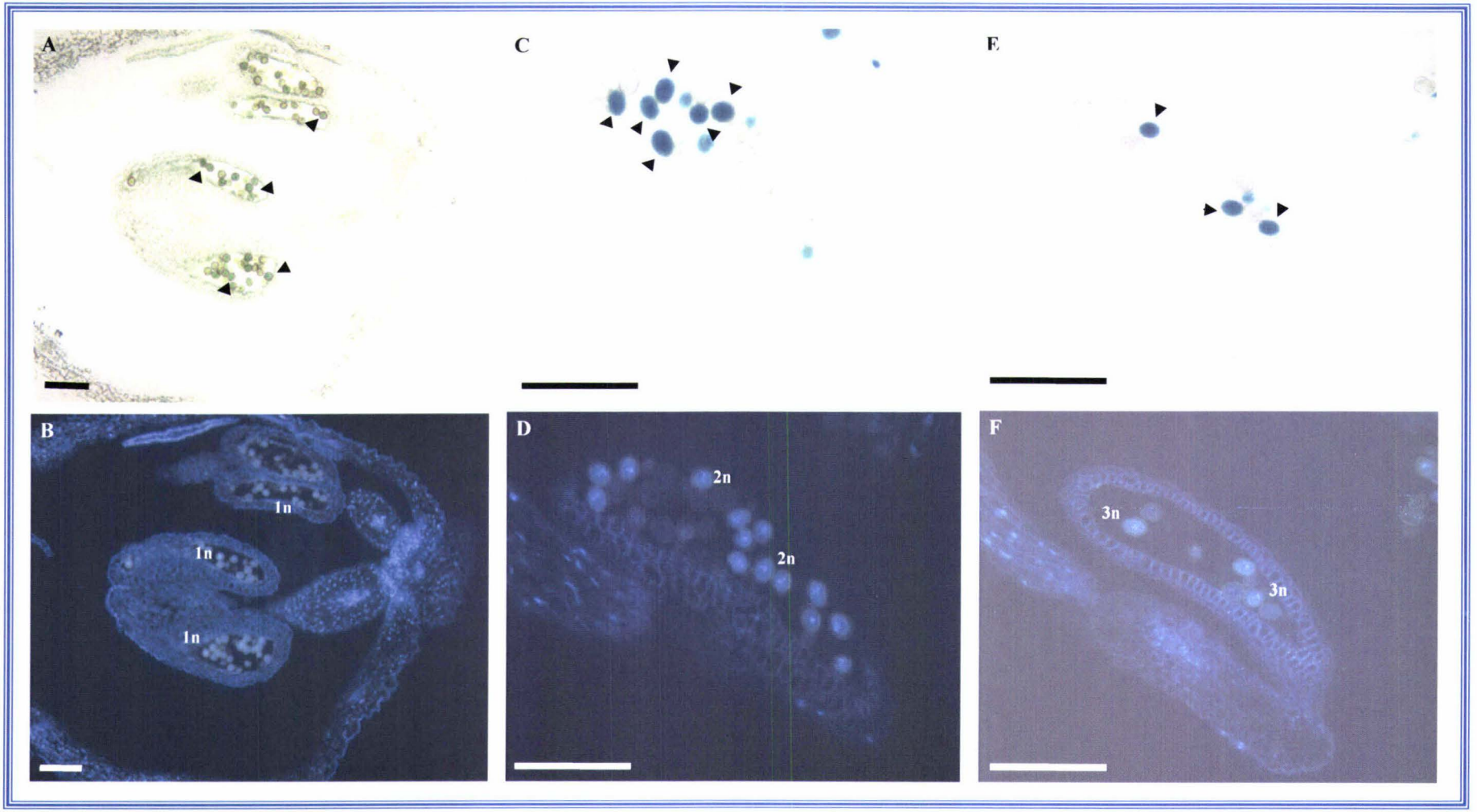

Figure 3.3 *AGL104::GUS* expression during the different developmental stages of pollen.

A, C, E: GUS stained tissue; **B, D, F**: DAPI stained tissue for panels A, C, E, respectively.

AGL104 expression first comes on weakly (A) at the uninucleate stage (B), increases in expression and is maintained at high levels (C and E) through the bicellular (D) and tricellular pollen (F) stages of pollen development. The different developmental stages of pollen were determined by counting the number of nuclei stained with DAPI. Not all pollen in panels A, C and E are stained with GUS since the mutant plants used for GUS assaying were segregating. However, all pollen in a particular anther is at the same developmental stage i.e. has the same number of nuclei. Since it is difficult to visualize the nuclei of GUS stained pollen, the number of nuclei of neighboring pollen in the same anther is indicated (B, D, F).

1n- one nucleus; **2n**- two nuclei; **3n**- three nuclei; **arrow head**- GUS stained pollen
Scale bar = 100μm

3.2.2.3 GUS expression during pollen germination and pollen tube elongation

Our results revealed that *AGL104::GUS* activity was maintained in the mature pollen and appeared to be highest at this stage (judging by the intensity of the GUS stains). In addition, Macroarray and Northern Blot analyses had detected the expression of *AGL104* in mature pollen (Kofuji et al., 2003). Moreover, publications following the commencement of our study reported that pollen transcriptome showed that *AGL104* was pollen enriched (high number of transcripts) and selectively (exclusively) expressed in pollen compared to other tissues (flowers, leaves, seedlings, siliques) (Pina et al., 2005) and exhibited the strongest level of expression in mature pollen (Honys and Twell, 2004). These authors had concluded that *AGL104* played a role in pollen germination and tube elongation (Mascarenhas, 1990).

To verify the spatial expression of *AGL104* during these processes, pollen were germinated *in vitro* and *in vivo* and GUS-assayed to determine if *AGL104::GUS* was expressed during pollen germination and tube elongation. As indicated in Figure 3.4, *AGL104::GUS* expression was maintained as the pollen germinated and the pollen tube elongated down the style *in planta* (Figure 3.4 A, B) and also *in vitro* (Figure 3.4 C). In the *in vitro* germinated pollen, it was noted that the staining was more intense at the growing tips of the pollen tube (Figure 3.4 C) in comparison to the rest of the elongating tube. The slides were focused at different planes and therefore this difference in stain intensity was not due to lack of focus along the growing pollen tube.

3.2.2.4 GUS expression in ovules

AGL104::GUS expression was detected in the ovules of *Arabidopsis* flowers (Figure 3.5 and 3.6). Our results indicated that *AGL104::GUS* expression came on relatively late during ovule development (Figure 3.5 F), at stage 12 or later of flower development when the stigma hairs had fully elongated into bristle-like structures (stigma papillae) and the stigma was receptive to pollen (Figure 3.5 D, E) (Smyth et al., 1990). *AGL104::GUS* expression was not detected (Figure 3.5 C) in ovules of flowers at stage 11 of development (Figure 3.5 A, B). At stage 11, the stigma papillae is in its initial stages of development and their outward growth is limited at first to regions not in contact with the overlapping sepals (Smyth et al., 1990). The stigma papillae at this phase have not elongated fully and the stigma is not yet receptive to pollen

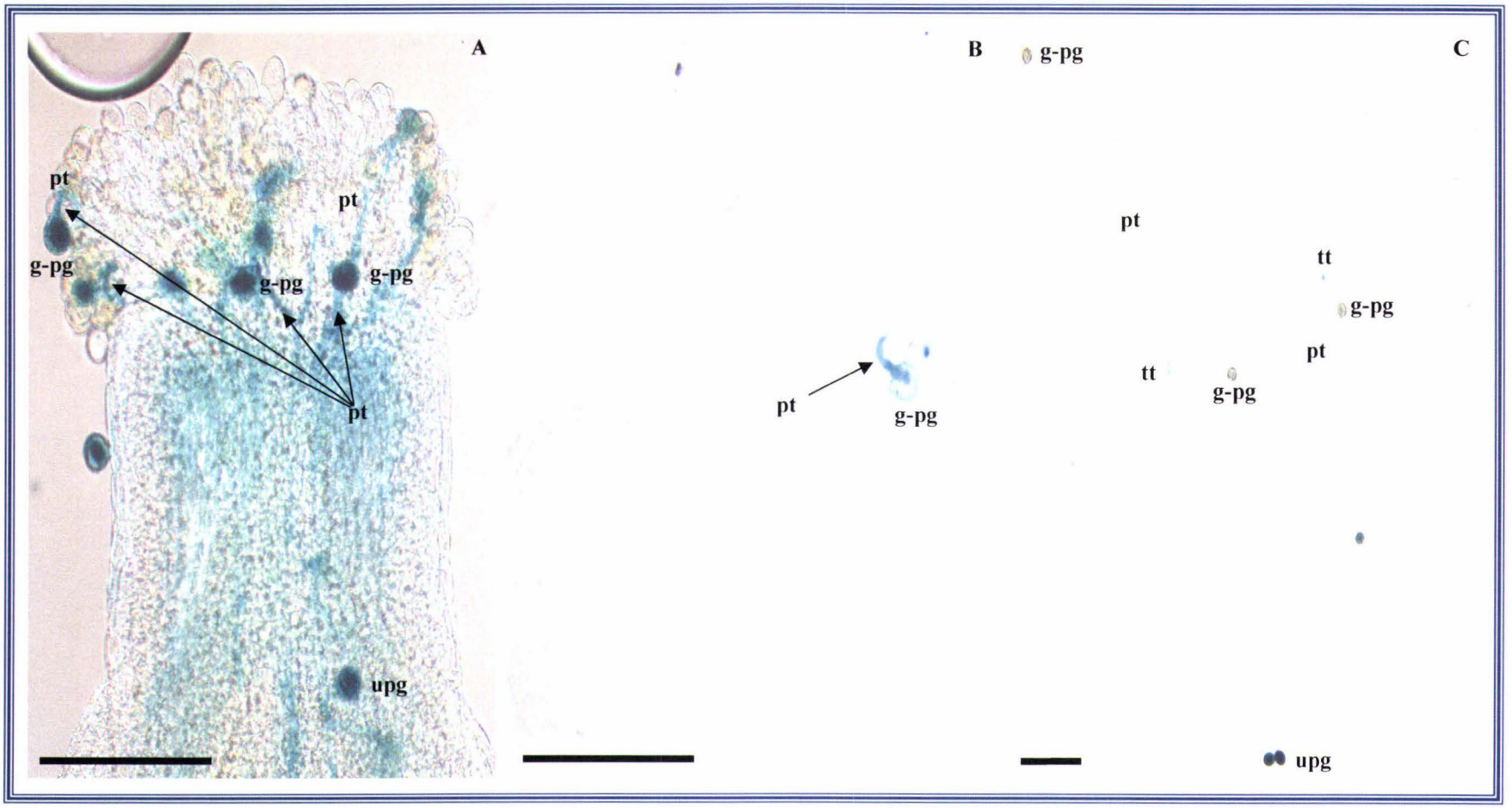

Figure 3.4 Histochemical localization of *AGL104::GUS* expression in germinating pollen and elongating pollen tubes.

A: pollen germinated in planta (40X); **B**: Cross section of a GUS stained germinating pollen (40X); **C**: *in vitro* germinated pollen (10X).

AGL104 expression is maintained in the pollen as it germinates and the pollen tube elongates. GUS expression appears to be higher (since more intense staining) in the tip region of the elongating pollen tubes compared to the rest of the pollen tube (C).

g-pg- germinating pollen grain; **pt**- pollen tube; **tt**- pollen tube tip; **u-pg**: ungerminated pollen grain.
Scale bar= 100µm

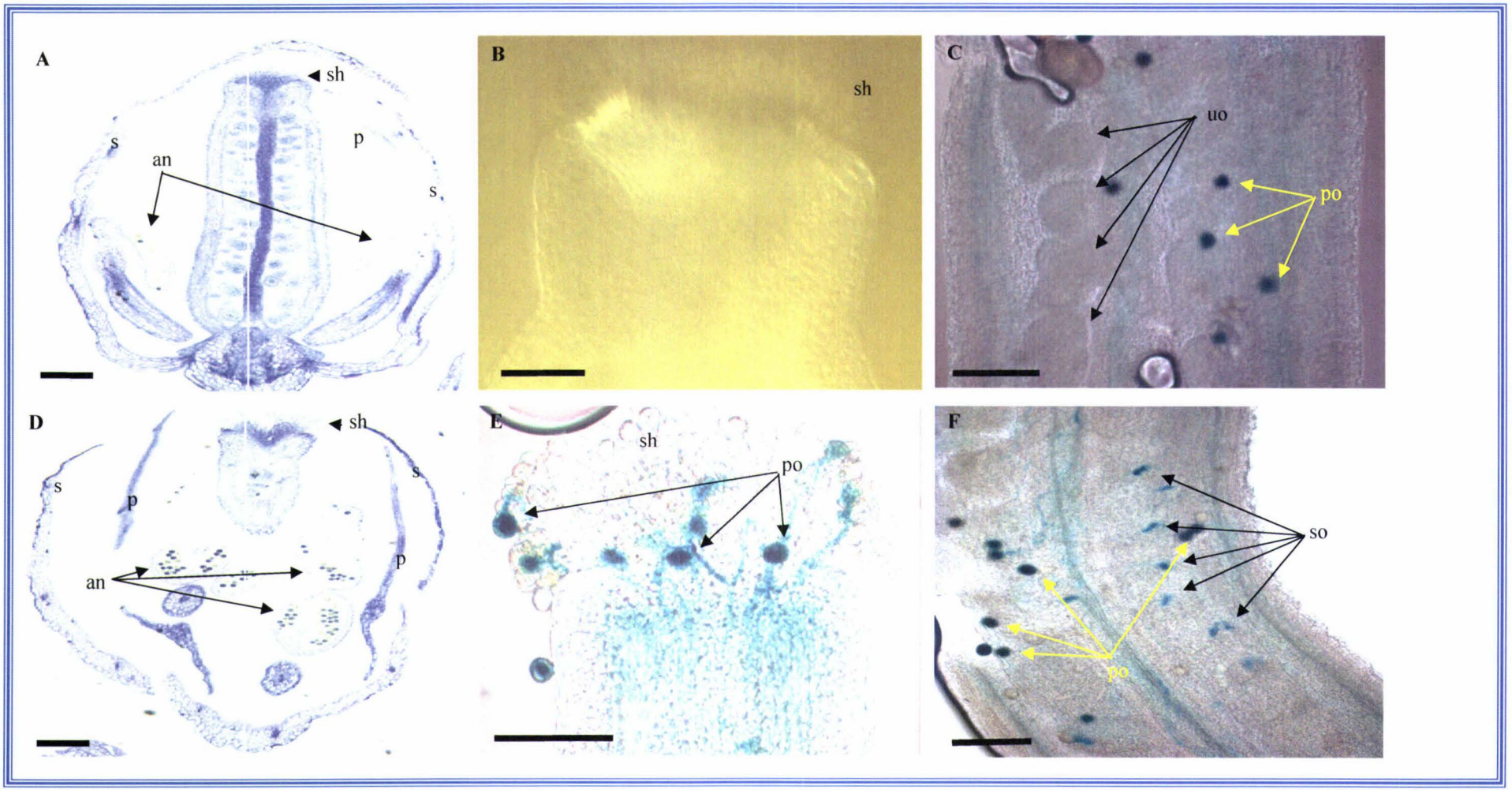

Figure 3.5 Histochemical localization of *AGL104::GUS* expression in the young and older ovules.

A: Cross section of a stage 11 flower (10X), stigma papillae is visible and sepal is closed. The tissue is stained with toluidine blue for visualization. **B**: Stigma hair of young stage 11 carpel (20X). **C:** Young, unstained ovules of carpel shown in panel B (20X). **D:** Cross section of an older flower with the petals open (10X). The tissue is stained with toluidine blue for visualization. **E:** Stigma hair of older carpel (40X) shown in panel D. **F:** Older, stained ovules of carpel shown in panel E (20X).

AGL104 expression comes on at a very late stage in ovule development. Ovules are ready to be fertilized when the stigma is receptive (stigma papillae are fully elongated). This happens at stage 12 or later of flower development. GUS staining is seen in the ovules (F) of flowers with receptive stigma (D-E) whereas there is no GUS staining in ovules (C) of younger flowers (A-B). Pollen seen in panel C and F were washed *over* the carpel specimen during preparation and not present *inside* the carpel.

an- anther; **p**-petal; **po**- pollen washed over the sample; **s**-sepal; **sh**- stigma hair; **so**- stained ovules; **uo**- unstained ovules. Scale Bar= 100μm

(Figure 3.5 A, B). Therefore, in ovules *AGL104::GUS* expression commences at a late stage of development, at stage 12 or later.

Further investigation into the tissues of the ovule in which *AGL104::GUS* expression was observed showed expression was limited to cells at the micropylar pole of the ovule (Figure 3.6 A). Since the embryo sac is located at this pole (Figure 3.6 B), it was concluded that *AGL104* was expressed in the embryo sac. It was also observed that *AGL104::GUS* was not expressed throughout the embryo sac but was limited to certain cells only, specifically the cells closest to the micropyle. These cells are most likely to be either the synergid cells or the egg cell since these are found at the micropylar pole of the embryo sac.

3.3 Discussion

3.3.1 Discrepancies between our GUS assay results and previously published *AGL104* expression results

AGL104::GUS expression analysis showed that *AGL104* was expressed in the developing anthers, mature pollen, elongating pollen tubes and ovules (Figures 3.2 - 3.6). This partly complimented the *in situ* hybridization results obtained by Parenicova et al. (2003) and the Northern Blot analysis by (Kofuji et al., 2003), where the *AGL104* transcript was detected in young developing anthers and ovules, and in pollen, respectively. However, in addition to developing anthers and ovules, Parenicova et al. (2003) detected *AGL104* transcripts in other floral organs such as sepals, petals and the septum of the carpel. Pale blue staining was observed in the veins of sepals and petals and the stems of some of our samples (Figure 3.2 A). However, the staining was very light and did not show consistency in all flowers of a particular mutant line or when flowers of different mutant lines were compared. Therefore, it was concluded that the blue coloring seen in some samples were due to unspecific staining or lack of proper washing and did not represent expression of the GUS protein in these tissues.

In stage 3 flowers, sepal primordia arise and begin to lengthen and curve inward until they overlie the flower primordium (Smyth et al., 1990). GUS expression was not detected in sepals of flowers at this stage of development (Figure 3.2 C) nor in the petals of older flowers and

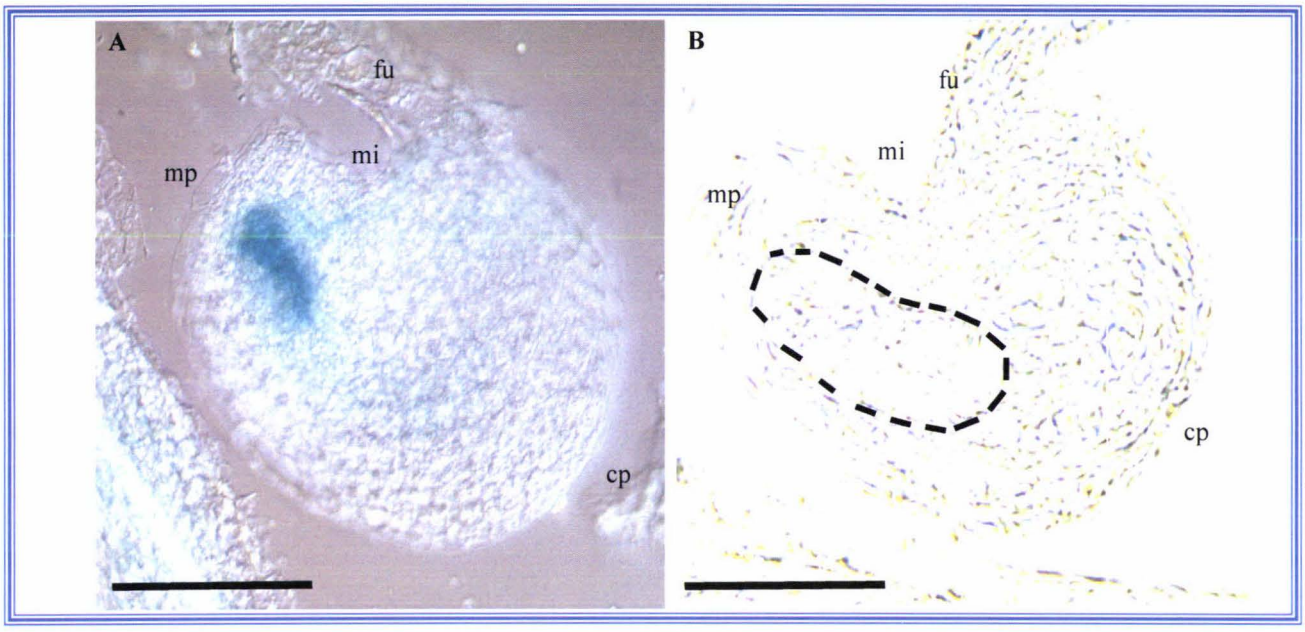

Figure 3.6 Histochemical localization of *AGL104::GUS* expression in the micropylar end of the ovule.

A: GUS stained ovule of a flower older than stage 12 of development (40X) **B**: Cross section of an ovule at a late stage of development (40X). The dotted line shows the location of the embryo sac.

AGL104 expression is seen in certain cells of the ovule only. Since the GUS stain is in the micropylar end, *AGL104* is expressed in the embryo sac, more precisely either in the two synergid cells or the egg cell, both of which are located close to the micropyle.

cp- chalazal pole; **fu-** funiculus; **mi**-micropyle; **mp**- micropylar pole.

Scale Bar= 100μm

septum of the carpel (Figure 3.2 D) as reported by Parenicova et al. (2003). We performed a preliminary GUS assay on whole inflorescences of the fifteen lines initially generated in this experiment, all of which showed a clear and consistent expression in developing anthers, mature pollen and ovules only and none in the sepals and petals. Detailed expression analysis of *AGL104::GUS* in the female reproductive organs was carried out in only three representative mutant lines chosen for this purpose. These lines showed *AGL104::GUS* expression in the embryo sac of the ovules and none in the septum of the carpel (Figure 3.5).

The difference between our GUS assay results and that of Parenicova et al. (2003) may have been due to the fact that the 1.7kb *AGL104* promoter region that was fused to the coding region of the GUS protein in our experiment perhaps did not contain all the DNA (*cis*) elements necessary for the expression of this protein in the sepals, petals and the septum of the carpel, but included only the elements necessary for its expression in the developing anthers, ovules and mature pollen. It was also noted that Parenicova et al. (2003) performed a mass *in situ* hybridization with five different genes but was able to detect the transcripts of only *AGL104* and their positive control, *AGL80*. Thus, there was a possibility that their *in situ* hybridization conditions were not optimal and resulted in unspecific hybridization in the sepals, petals and the septum of the ovules in the *Arabidopsis* flower.

To clarify if *AGL104* is expressed in the developing anthers, pollen and ovules only as indicated by our results, or if it is also expressed in the sepals, petals and the septum of the ovule as proposed by Parenicova et al. (2003), further experiments need to be conducted. GUS constructs containing more sequences of the 5` and 3` untranslated regions (UTR) of *AGL104*, which may include more regulatory elements, need to be generated. We used only a 1.7kb fragment since regulatory elements necessary for gene expression for a majority of the genes are usually present within this distance of the start codon. In addition, the constructs should include the intragenic region of the *AGL104* gene. It was previously shown that intragenic regions were also involved in the regulation of gene expression (Sieburth and Meyerowitz, 1997) and elimination of these regions could affect expression. Analyses of such constructs that include all the possible regulatory regions should clarify the endogenous expression pattern of *AGL104* and further functional characterizations should determine its role in the *Arabidopsis* flower.

3.3.2 Function of *AGL104* in young developing anthers and the meristem

GUS assays of whole inflorescences failed to detect GUS expression in anthers of flowers younger than stage 12 of development and in the meristem region (Figure 3.2). This implied that the *AGL104* promoter was either activated in older anthers at a later stage of pollen development or was repressed in younger anthers and the meristem region. *DEFH125* is a MADS-box gene isolated from the pollen of *Antirrhinum* (Zachgo et al., 1997). Investigation into the expression pattern of *DEFH125* led to implications that it was involved in male gametophyte development (Zachgo et al., 1997). Since the B class genes, *DEF* and *GLO*, in *Antirrhinum* were known to be involved in stamen and petal organogenesis, the question arose if these two genes interacted with *DEFH125*. Verification of expression pattern revealed that the expression pattern of *DEFH125* was complementary to that of *DEF* and *GLO* (Lauri et al., 2006). Whereas *DEFH125* was exclusively expressed in the developing microspores, *DEF* and *GLO* were expressed in the filaments, connective tissues and vasculature of the stamens but were absent from the microspores. Further scrutiny of the promoter region of *DEFH125* led to the discovery that DEF/GLO heterodimers bound to a specific CArG-box on the *DEFH125* promoter and acted as a repressor of *DEFH125* (Lauri et al., 2006). This interaction and resulting expression of *DEFH125* was thought to allow the development of microspores in their specific locales.

Since *AGL104* is also a MADS-box transcription factor isolated from pollen, it is possible that, similar to *DEFH125*, the expression of *AGL104* is excluded from the undeveloped microspore and the meristem regions, thus explaining its absence. Further analysis of the promoter region of *AGL104* would determine the elements that regulate the expression of *AGL104* and the mode of regulation that is used i.e. whether *AGL104* is deliberately activated in the older anthers or simply repressed in the younger ones. This information would shed light on the mode of regulation used by *AGL104* to ensure proper development of the male reproductive organs.

3.3.3 Proposed functions of *AGL104* in pollen development, pollen germination and tube growth deduced from its expression pattern

.

AGL104::GUS expression analysis performed in our laboratory and i*n situ* hybridization results obtained by Parenicova et al. (2003) established the expression of *AGL104* in the

developing anthers and mature pollen (Figure 3.2). Male gametophyte development is a complex phenomenon that is highly regulated. As evident from the phenotypes of various mutants described previously (Chapter 1, section 1.8.1), changes in the expression of key regulatory genes can have drastic effects, even lead to non-functionality of the male gametophyte which can seriously undermine the chances of survival of a species. Thus, accurate spatial and temporal gene expression in the gametophyte is crucial to the reproductive success of a species.

During male gametophyte development, progression through PMI is associated with a dramatically increased level of transcription whereas after PMI there is a major shift in the translatable mRNA population (Mandaron et al., 1990; Schrauwen et al., 1990). Analysis of the accumulation profiles of a number of pollen-expressed and pollen-specific mRNAs revealed two broad classes of genes: the early pollen-expressed genes and the late pollen-expressed genes, both of which may belong to different regulatory or functional groups (Mascarenhas, 1990).

3.3.3.1 Early and late pollen-expressed genes

Early pollen-expressed gene transcripts are first expressed in microspores, but decrease in abundance before pollen maturation. These genes are known to control basic cellular functions such as translation and mitochondrial functions in order to support microspore development (Twell, 2002). Late pollen-expressed genes, on the other hand, are usually expressed after PMI and accumulate to a high level during pollen maturity (Eady et al., 1995), and are thought to control processes such as pollen maturation, germination and pollen tube elongation (Mascarenhas, 1990). Since the late pollen-expressed genes represent a much larger group of gametophytically expressed transcripts, there are additional complexities in the patterns of expression of these genes. For example, some late pollen-expressed genes are in fact first activated (weakly) in the unicellular microspore and are strongly enhanced after PMI (Eady et al., 1995), instead of being activated after PMI at the bi-cellular stage as previously thought.

According to the above definitions, *AGL104* is classified as a late pollen-expressed gene, a fact that was visually confirmed by our GUS assay and DAPI staining results (Figures 3.3). *AGL104::GUS* was weakly expressed very early, at the uninucleate stage of development, and reached a maximum level of expression in the bi- and tri-cellular stage of pollen development.

This is the first known data showing spatial expression of *AGL104* in the different developmental stages of pollen even though temporal expression had previously been reported (Honys and Twell, 2004). As mentioned above, due to their accumulation and storage in the mature dehydrated pollen, the late pollen-expressed genes were assumed to function during pollen maturation, germination and tube growth stages (Mascarenhas, 1990).

To confirm if *AGL104*, being a late pollen-expressed gene and accumulating in mature pollen, was expressed during pollen germination and tube growth, we performed GUS assays on germinating pollen. Our analyses indicated that GUS activity was maintained during pollen germination and tube elongation (Figure 3.4). It was also noted that *AGL104::GUS* activity was concentrated at the growing tips of the pollen tube since a more intense staining was observed at the tips compared to the rest of the pollen tube (Figure 3.4 C). Our results are the only known data showing spatial expression of *AGL104* in the germinating pollen and the tip region of the pollen tube. This discovery was novel since previous studies on *AGL104* had not focused on its expression during pollen germination and tube elongation and thus there had been no reports of *AGL104* expression in pollen tubes. This may have been due to the popular belief that pollen germination and tube elongation occurred independently of transcription (Hoekstra and Bruinsma, 1979; Mascarenhas, 1975b), and thus expression of a transcription factor during these processes may have seemed unlikely and therefore not investigated further. Therefore, our findings were significant as it raised the possibility that pollen germination and tube elongation may not occur independently of transcription as previously understood. This possibility was further strengthened by recent reports of isolation of transcription factors from pollen of various species as described below.

3.3.3.2 Transcription factors expressed in pollen

Initial evidence for the presence of transcription factors in pollen emerged when several different pollen-specific *cis*-regulatory elements were identified in the promoter region of various pollen-specific genes (Bate and Twell, 1998; Eyal et al., 1995). The results indicated that these *cis*-regulatory elements acted as a positive regulator to either enhance expression of these genes in pollen or to activate transcription factors. However, data on the type of transcription factors that bound to these pollen-specific *cis*-elements and the regulatory mechanisms that they utilized were lacking. Search for the protein compliments that bound to

the *cis*-elements thus led to the isolation of several different classes of transcription factors from pollen of various species.

Various genes of the MADS-box family of transcription factors have been isolated from pollen. *DEFH125* was the first report of a MADS-box transcription factor in pollen of *Antirrhinum* (Zachgo et al., 1997). *DEFH125* was expressed at detectable levels in the third whorl (stamen) and only after the completion of PMI, thus *DEFH125* was classified as a late pollen-expressed gene. Detailed expression analysis showed that *DEFH125* was first expressed (weakly) in the uninucleate stage, increased in expression in the cytoplasm of the vegetative cell as the microspore developed, and persisted in the *in vitro* germinated pollen (Lauri et al., 2006). This led to the proposal that *DEFH125* was involved in pollen development in the stamen and during pollen germination and tube elongation. A second gene, *ZmMADS2*, isolated from maize, is another example of a MADS-box transcription factor in pollen. Temporal and spatial expression analyses showed that *ZmMADS2* expression was restricted to pollen and roots (Heuer et al., 2000). RT-PCR detected *ZmMADS2* expression at all stages of pollen development with transcripts being most abundant in mature pollen after dehiscence, thus *ZmMADS2* also represented a late pollen-expressed gene. *In situ* hybridization on *in vitro* germinated pollen localized *ZmMADS2* transcripts in the pollen tubes. These transcripts were translocated into the pollen tubes immediately upon germination and displayed a gradient with increasing concentration towards the tube tip (Heuer et al., 2000). The high expression in pollen after dehiscence and localization of transcripts in pollen tubes implied a function of *ZmMADS2* in pollen tube growth. As previously described (section 3.2.2.3), *AGL104::GUS* exhibited an expression pattern similar to *DEFH125* and *ZmMADS2*.

In addition to MADS-box genes, other types of transcription factors have also been isolated from pollen. Recently, members of the *MYB* family, which represents the largest transcription factor family in plants and regulates diverse developmental processes, have been isolated from tobacco pollen (Yang et al., 2001). The transcripts of the tobacco genes, *NtMYBAS1* and *NtMYBAS2*, were expressed in the sporophytic and gametophytic tissues of the anther and were thus thought to regulate gene expression in both the sporophytic anther tissues and in the male gametophyte i.e. the pollen. In pollen, the *NtMYBAS* transcripts initially accumulated in the uninucleate microspores and, similar to the late class of pollen-expressed genes, increased in expression in the mature pollen. However, further investigation suggested that these proteins were anther-specific transcription factors and therefore, they most likely regulated processes in

the sporophytic, but not the gametophytic tissues of the anther (Yang et al., 2001). Pollen-specific genes encoding zinc-finger proteins of the LIM class represent another class of transcription factors isolated from pollen of sunflower and tobacco (Baltz et al., 1992; Sweetman et al., 2000). An important functional theme of LIM-domain proteins is their ability to interact with themselves, other LIM proteins or unrelated proteins. Thus, LIM domains can be perceived as protein-binding modules interacting with a wide range of other proteins. In tobacco, these proteins were detected immediately before PMI and reached a maximal level in mature pollen, which was maintained in germinating pollen (Sweetman et al., 2000). This high level of developmentally regulated expression specifically in pollen suggested an important role for LIM proteins during pollen maturation.

Based on the expression of transcription factors in pollen described above; the high level of expression of *AGL104::GUS* in the mature pollen; the continued activity of the GUS protein in the germinating pollen and elongating pollen tube; and taking into consideration the results of the pollen transcriptome analyses performed by Pina et al. (2005) and Honys and Twell (2004), who detected an increase in *AGL104* transcript in the mature pollen, we proposed that *AGL104*, being a transcription factor, was involved in pollen development during the bi- and tri-cellular stage and in the regulation of genes controlling pollen germination and tube growth.

Nevertheless, it was noted that the blue precipitate detected at the tips of the pollen tube in our GUS assay represented the **action** of the GUS enzyme and not necessarily its active synthesis in this region. In the *AGL104::GUS* construct, the GUS enzyme was synthesized in the tissues where the *AGL104* promoter was active. Since the GUS protein is very stable (Bottino, 2005), it may have persisted in pollen for a longer period than endogenous proteins. As mentioned previously, prior to our study, the *in situ* hybridization results obtained by Parenicova et al. (2003) were the only known published report on the spatial expression pattern of *AGL104* and thus was heavily relied upon when designing our experiments. Since our interest initially was to confirm the expression pattern of *AGL104* in the different tissues of the *Arabidopsis* flower, the GUS assay method suited our purposes.

In view of our GUS assay data that implied that *AGL104* was involved during pollen germination and tube elongation, we were compelled to examine our experimental design more carefully. It is known that the male gametophyte, i.e. the mature pollen, is composed of three cells with the vegetative cell (VC) making up the bulk of the pollen. Since the VC extends to

become the pollen tube, the contents of the VC exits the pollen grain and travels at the tip of the growing pollen tube (McCormick, 2004), i.e. the contents of the mature pollen grain are drained into the pollen tube. Therefore, it was noted that the detection of GUS enzyme activity during pollen germination and at the growing tips of the pollen tube could be due to two scenarios:

- The GUS protein was **actively synthesized** during the processes of pollen germination and tube elongation. This would indicate active synthesis of *AGL104* and thus active gene regulation during these processes.
- The GUS protein was **pre-synthesized** and stored in the mature pollen grain and "drained out" with the cytoplasm into the pollen tube as the vegetative cell gave rise to the pollen tube. This would mean that the GUS protein, whose activity was detected at the tip of the pollen tube, may have been a remnant protein and not necessarily indicate active expression of *AGL104* and its functioning as a transcription factor in this region.

To differentiate whether the *AGL104* promoter was indeed active in the tip region or whether the GUS enzyme represented a remnant protein that had not yet been degraded, it was crucial to determine the localization of AGL104 protein in the VC. Immunoprecipitation studies combined with DAPI staining would have provided conclusive evidence. Localization of the AGL104 protein into the vegetative nucleus (VN) would have implied its activity in the tip region whereas localization to the cytoplasm of the VC would have left open the possibility that AGL104 was a remnant protein and thus may not necessarily have a function during pollen germination and tube elongation. Due to time constraints, we could not perform immunoprecipitation studies and instead compared published literature on the synthesis and activity of other closely related, pollen-expressed MADS-box genes, such as *ZmMADS2* and *DEFH125*, and compared their behavior to *AGL104*.

In maize, the untranslated transcripts of *ZmMADS2* were localized to the pollen tube before the VN moved into the elongating tube (Heuer et al., 2000), implying that the gene was newly translated later during the process of pollen tube elongation. This is in agreement with the results of Honys and Twell (2004) that active translation occurred during pollen germination and tube elongation. In addition, the DEFH125 protein, which has high homology to ZmMADS2, was immunolocalized from the cytoplasm of the VC to the VN during the process

of tube elongation (Zachgo et al., 1997). These data implied that a MADS-box transcription factor was translated and localized to the nucleus **during** pollen germination and tube elongation. Localization of transcription factors into the nucleus is indicative of an active regulation of genes, i.e. active transcription or repression. Since *AGL104* is a MADS-box gene and displays an almost identical expression pattern to *DEFH125* and *ZmMADS2* (Heuer et al., 2000; Zachgo et al., 1997), it was most likely that *AGL104* was also actively synthesized at the tips of the elongating pollen tubes and transported into the VN to regulate gene expression and not represent a remnant protein.

Assuming the above proposition was correct, the fact that a transcription factor was present during the processes of pollen germination and tube elongation had huge implications. As mentioned above, it was generally believed that pollen germination and tube growth occurred independently of transcription (Hoekstra and Bruinsma, 1979; Mascarenhas, 1975b). If these processes were indeed independent of transcription, then what was the role of *AGL104*, a transcription factor that regulated gene expression, in these tissues? The involvement of the processes of transcription and translation in pollen of other species were assessed to shed some light.

3.3.3.3 Transcription and translation during pollen germination and tube growth

In most angiosperms, germination and pollen tube growth occur rapidly, requiring swift activation of synthetic and catabolic processes. Early inhibitor studies had shown that germination and early tube growth were largely independent of transcription but strictly dependent on translation (Hoekstra and Bruinsma, 1979; Mascarenhas, 1975b). Moreover, biochemical analyses showed that the mRNAs, ribosomes, and tRNAs required for germination were synthesized during pollen maturation and persisted in the pollen grain until they were utilized for translation during the germination process (Mascarenhas, 1992). This was further confirmed by Microarray experiments, which indicated that mRNAs encoding translation proteins were enriched in the mature pollen (Becker et al., 2003; Honys and Twell, 2003; Pina et al., 2005).

However, in species such as *Impatiens balsamina*, blocking protein translation did not affect pollen germination or tube growth, indicating that proteins required for these initial processes were present in the mature pollen grain (Shivanna et al., 1974). Further support was provided

by the generation of antibodies which led to the discovery that proteins corresponding to mRNAs that coded for translation complexes in cells were *already* present in the mature, ungerminated pollen (Kim et al., 2002; Muschietti et al., 1994; Muschietti et al., 1998), indicating that some of the genes were not only transcribed but also translated before the mature pollen was shed. Taken together, these results supported the mechanistic notion that mature pollen was charged with functionally specialized mRNAs and preformed translational apparatus (made up of protein complexes) to enable rapid translation of these mRNAs into proteins upon pollen hydration and germination (Honys et al., 2000; Mascarenhas, 1989; Schrauwen et al., 1990).

In spite of such observations, the dependence of pollen germination and tube elongation on pre-synthesized mRNA and translation complexes varies with plant species. In species such as *Typha*, *Lilium* and *Clivia*, it was observed that **new** translations were essential for germination as blocking translation inhibited germination (Franke et al., 1972; Hoekstra and Bruinsma, 1979). This indicated that mature pollen of these species contained only the preformed mRNA but not the preformed protein complexes essential for translation during germination. In other words, these particular species required **active translation** for pollen germination.

Some species are also known to require **active transcription** in order to facilitate either the initial pollen germination or to sustain pollen tube elongation after its initial growth (Mascarenhas, 1966; Mascarenhas and Mermeisten, 1981). This suggests a need for transcription factors during these processes. For example, the RNAs synthesized during pollen germination and tube growth in the species *Tradescantia* appeared to be messenger RNAs or messenger RNA precursors (Mascarenhas, 1966; Mascarenhas and Mermeisten, 1981), indicating that new transcription occurred during pollen germination and the initial tube growth. Further support was provided by the discovery that the tobacco pollen specific gene, *NTPc303*, was transcribed during pollen maturation, pollen germination, and pollen tube growth (Weterings et al., 1992). In addition, plants shedding bicellular pollen are known to be dependent on new mRNA synthesis for the maintenance of pollen tube growth and division of the generative cell into two sperm cells (Mascarenhas, 1975a). Thus, active transcription in some species would necessitate the expression of transcription factors during the germination and pollen tube elongation.

Considering transcription and translation in the different species described above, it was tempting to assume that despite popular belief, active transcription of genes occurred during pollen germination and tube elongation in *Arabidopsis*, thus accounting for the presence of *AGL104*, a transcription factor, during these processes. However, Honys and Twell (2004) showed that *Arabidopsis* pollen germination and tube elongation were relatively independent of transcription but were strictly dependent on translation. Germinating *Arabidopsis* pollen in increasing concentrations of Actinomycin D (a transcription inhibitor) and Cycloheximide (a translation inhibitor) indicated that blocking transcription had only moderate effects on pollen germination and tube growth even at high concentrations while blocking translation led to a 60% decrease in pollen germination and 90% inhibition of pollen tube growth (Honys and Twell, 2004). The expression of *AGL104* during these processes (Figure 3.4) was thus curiosity-provoking. Was it possible that *AGL104* may still be involved in the transcription of genes but these genes did not directly affect pollen germination and tube elongation or their effects were so negligible that blocking transcription did not result in observable phenotypes? On the contrary, since transcription factors regulate gene expression by acting as an activator as well as a repressor, could *AGL104* act as a repressor of gene expression during pollen germination and tube elongation? To answer these questions, functional characterization was carried out (Chapters 4 and 5).

3.3.4 Proposed function of *AGL104* in ovules

AGL104::GUS expression in the ovules was detected after stage 12 of flower development (Figure 3.5) (Smyth et al., 1990). This was in agreement with the results of Parenicova et al. (2003) where *in situ* hybridization detected *AGL104* transcript in the developing ovules of **older** flowers only. Given that *AGL104::GUS* was expressed at this late stage in the ovules, it was unlikely that *AGL104* played a role in ovule development *per se* since the stigma was receptive at stage 12, which indicated that the ovules were ready to be fertilized (Smyth et al., 1990), thus fully developed. Therefore, the expression of *AGL104* this late in the female gametophyte implied a role in processes such as pollen tube guidance or fertilization that occurred late in the mature ovule. Further investigation showed that *AGL104* was expressed specifically in cells found at the micropylar end of the embryo sac (Figure 3.6). Since the egg cell and the synergid cells are found in this location, it was concluded that *AGL104* was expressed in one or both of these two cell types.

To correctly establish the role of *AGL104* in the female gametophyte, it is important to determine the precise type of cell in which *AGL104* is expressed. This could be achieved by staining a pollinated carpel containing a pre-fertilized GUS-assayed ovule with DAPI. Given that the DAPI dye stains the DNA of a nucleus, it would be possible to differentiate between the synergid cells and the egg cell since only the egg cell would be "approached" by one of the DAPI-stained sperm nuclei. Alternatively, the position of the nuclei in the egg apparatus could help differentiate between the synergid and the egg cell. While the cytoplasm and nucleus of the egg cell are located at the chalazal pole of the egg apparatus, nuclei of the synergid cells are located at the microphylar pole (Christensen et al., 1997).

3.3.4.1 Pollen tube guidance

Localization of *AGL104* in the synergid cells of the embryo sac would imply an involvement in pollen tube guidance (Higashiyama et al., 2001). It had been accepted for sometime that the mature female gametophyte, specifically the embryo sac, was the source of a pollen tube guidance cue (Hulskamp et al., 1995). Upon landing on a stigma, the pollen tubes in *Arabidopsis* penetrate the stigmatic papillae cells and travel through the style and septum before emerging onto the surface of the septum. There they adhere to the epidermis and travel along the transmitting tract to the ovules (Lennon et al., 1998). Kandasamy et al. (1994) observed that in immature pistil, the pollen tubes grew in random directions and continued to grow towards the base of the ovary instead of towards individual ovules as they passed down the style. This implied that mature ovules were involved in signaling mechanisms that directed pollen tube growth towards itself. In addition, the observation that only one tube was directed towards an ovule in the mature pistil implicated the existence of negative signals from the ovule to stop other tubes from growing towards it. Hulskamp et al. (1995) obtained similar results. This group noted that in mutants where the embryo sac development was affected, pollen tubes failed to reach the ovules, implying that the embryo sac within the ovule was involved in long-range signaling to guide a pollen tube towards itself.

Pollen tube guidance from the transmitting tract to the ovule is divided into two steps: funiculus guidance and micropyle guidance. The funiculus guidance signal guides a pollen tube from the transmitting tract onto the surface of a funiculus while the micropyle guidance signal leads the pollen tube to the micropyle (Shimizu and Okada, 2000). Earlier morphological observations of the female gametophyte had shown that the synergid cells had

properties of a secretory cell and thus was suspected to be the source of a diffusible guidance signal (Murgia et al., 1993), which was recently confirmed through laser cell ablation studies (Higashiyama et al., 2001). This group discovered that the synergid cells secreted a species-specific (Higashiyama et al., 2006) diffusible guidance molecule involved in micropyle guidance (Kasahara et al., 2005). In maize, this guidance molecule is encoded by the *ZmEA1* gene, which is expressed in both the egg and synergid cells and its product is secreted into the micropyle (Marton et al., 2005). After fertilization, ZmEA1 was no longer detectable. If expression of *AGL104* is localized to the synergid, regulation of genes that encode or degrade such synergid-specific guidance molecules may be a possible role.

Apart from signals from the synergid cells, other tissues are also involved in pollen tube guidance. GABA (γ-amino butyric acid), which is regulated by *POP2*, is a pollen tube guidance molecule known to be expressed in the pistil of *Arabidopsis* with increasing concentrations from the stigma to the micropyle (Palanivelu et al., 2003). In addition, it had previously been reported that a MADS-box gene, *DEFH125*, was first expressed in the mature pollen and immediately after pollination it was localized to the nuclei of cells lining the transmitting tract (Zachgo et al., 1997). Since the pollen tubes adhere to the septum as they grow down the style, it is possible that, like *DEFH125* and *POP2*, *AGL104* may be involved in regulating secretion of molecules involved in pollen tube guidance.

3.3.4.2 Pollen tube reception and fertilization

In addition to pollen tube guidance, the mature embryo sac is also involved in processes such as fertilization (Huang and Russell, 1994; Russell, 1992; Russell, 1993) and initiation of seed development (Chaudhury et al., 1997; Ohad et al., 1996). During fertilization, the pollen tube enters through the funiculus, penetrates the degenerating synergid cells and bursts to release two sperm cells, one of which fuses with the egg to form the diploid zygote while the other fuses with the central cell to form the triploid endosperm (Grossniklaus and Schneitz, 1998). Discovery of the female gametophytic mutants, *feronia* (Huck et al., 2003) and *sirene* (Rotman et al., 2003) have implicated the synergid cells in the control of later steps of pollen tube growth, that is, pollen tube reception (Huck et al., 2003) that involves termination of pollen tube growth. In these mutants, wt pollen tubes entered mutant female gametophytes but failed to cease growth, rupture, and release their contents. Since these processes took place within the synergid cell, the synergid cell-expressed gene products were thought to be necessary for

ensuring the release of the sperm cells (Weterings and Russell, 2004). A transcription factor, *MYB98*, specific to synergid cells was recently isolated from *Arabidopsis*, thus implying existence of a regulatory mechanism in the synergid cells (Kasahara et al., 2005). If *AGL104* is shown to be expressed in the synergid cells, it is possible that it is involved in pollen tube reception and fertilization.

3.3.4.3 Seed formation and embryogenesis

Upon double fertilization, the ovule develops into a seed harboring both fertilization products: the egg and the endosperm. Cell ablation results indicate that the egg cell and the central cell are not directly responsible for attracting a pollen tube (Higashiyama et al., 2001). Seed formation requires the coordinated development of both fertilization products and of the integumental cell layers surrounding them. Since both the mega- and micro-gametophytes develop from a single spore, each pair of nuclei participating in the double fertilization is genetically identical. Yet, the two fertilization products follow very different paths of development that is thought to be determined by cytoplasmically or chromosomally (epigenetic) based differences between the central cell and the egg (Grossniklaus and Schneitz, 1998). Recently, a type I MADS-box gene, *AGL80/FEM111*, has been identified whose disruption led to defects in central cell and endosperm development (Portereiko et al., 2006) since in *fem111* mutants, the nucleolus (site of ribosome biogenesis) and vacuole of the central cell failed to mature properly and were smaller than the wildtype (Portereiko et al., 2006). In addition, endosperm development in *fem111* was not initiated after fertilization although the fertilized egg developed normally (Portereiko et al., 2006). If *AGL104* is revealed to be expressed in the egg cell of the embryo sac, could it be possible that *AGL104* represents the "complementary" gene to *FEM111* and is thus involved in ensuring that the fertilized egg pursues the embryo developmental pathway?

Finally, the sperm cells contained within the pollen tube have been shown to also express genes (*HAP2*) involved in pollen tube guidance and fertilization (Besser et al., 2006). Could the differential expression of genes in the two sperm cells, in addition to differential gene expression in the egg and central cell, be the major determinant of the two developmental pathways leading to the embryo and the endosperm? It would be interesting to determine if *AGL104* is localized to (one or both) the sperm cell nuclei in pollen and what the consequences of fertilization involving these sperm cells would be.

3.3.5 Role of *AGL104* in general gametophyte development

Recently, all members of the MIKC* clade, including *AGL104*, was reported to be highly expressed in the embryonic tissue (Lehti-Shiu et al., 2005). Transcriptome analyses of the male gametophyte, i.e. mature pollen, had revealed an **over-representation** of the type I and MIKC* type genes while all other types of transcription factors were under-represented (Pina et al., 2005). Is there a "parallel" in the expression, and thus function, of type I and MIKC* type genes in the male and female gametophytes of angiosperms? Detailed analysis of the involvement of *AGL104* in ovule development in gain-of-function and loss-of-function mutants (already generated but not analyzed), is required to establish the function of *AGL104* in ovules. Comparison of its function in the female gametophyte with its function in the male gametophyte (Chapter 4 and 5) will help determine the precise role of *AGL104* in gametophyte development as a whole, and comparisons with lower plants may reveal its ancestral role in land plants and thus give insight into evolution of plant architecture.

3.4 Chapter Summary

AGL104::GUS expression was detected in the inflorescence of the *Arabidopsis* flower, particularly in developing anthers and mature pollen (section 3.2.2.1). We did not detect expression in the meristem region or in flowers younger than stage 12 of development suggesting that *AGL104* was either turned on at a particular point in flower development or was turned off in the young developing tissues. Considering that the MADS-box gene, *DEFH125*, was purposely repressed in the sporophytic tissues of the anther but expressed only in developing microspores to allow pollen development (Lauri et al., 2006), we suggested that perhaps the absence of *AGL104* from the meristem region and very young flowers was due to a similar regulation i.e. purposeful repression. Further analysis of the *AGL104* promoter was suggested to determine the elements regulating *AGL104* expression and the mode of expression used in the stamen.

AGL104 expression was predominantly detected in the developing anthers of older flowers. We correlated the expression of *AGL104* in the male gametophyte with the different stage of male gametophyte development (section 3.2.2.2) and discovered that *AGL104* was first expressed weakly in the uninucleate stage and increased in expression as the male gametophyte

development progressed, with the strongest expression being in the mature pollen. Ours is the first report of spatial expression of AGL104 in the different stages of developing pollen. Expression in the mature pollen indicated involvement of *AGL104* in pollen germination and tube elongation. Thus, we performed GUS assays on germinating pollen (section 3.2.2.3) and detected *AGL104* expression in germinating pollen and at the tips of the elongating tubes. This also was the first report of the expression of *AGL104* in pollen tubes. Since the GUS protein is very stable, its detection at the pollen tube tips implied either active synthesis of the GUS protein or the presence of a remnant protein that was expressed in mature pollen and drained out into the elongating tube upon pollen germination. Localization of the AGL104 protein using immunoprecipitation assays and superimposing the results on DAPI-stained pollen tubes was suggested to determine if *AGL104* was active in the tip region. Localization into the nucleus would have been indicative of an active regulatory role in the tip region whereas localization into the cytoplasm may have left open the possibility of it being a remnant protein. Due to time constraints, we opted to compare the synthesis and activity of other similarly expressed MADS-box transcription factors isolated from pollen of *Antirrhinum* and maize (section 3.3.3.2).

Since in these species the transcription factors were actively synthesized and translocated to the nucleus and *AGL104* follows an identical expression pattern, we assumed that *AGL104* was also synthesized and was active during pollen tube elongation. Since *AGL104* is a transcription factor, this data suggested that active transcription was taking place in the germinating pollen and during tube elongation, an implication that was contrary to the accepted notion that pollen germination and tube elongation was independent of transcription but dependent on translation (Hoekstra and Bruinsma, 1979; Mascarenhas, 1975b). Since Honys and Twell (2004) had demonstrated that pollen germination and tube growth in *Arabidopsis* were indeed independent of transcription but strictly dependent on translation, we proposed that *AGL104* was either involved in transcription of genes that had no or negligible effects on pollen germination and tube elongation or acted as a repressor during these processes, thus blocking transcription (as done by Honys and Twell, 2004) did not result in an observable phenotype. Functional characterization was emphasized.

In the female gametophyte, *AGL104* was observed to be expressed very late during its development (section 3.2.2.4), similar to its late expression in the male gametophyte (section 3.2.2.2 and 3.2.2.3). Further scrutiny revealed that *AGL104* was specifically expressed in the

micropylar cells of the female gametophyte i.e. in either the egg or the synergid cells. DAPI staining to visualize the position of nuclei was suggested in order to differentiate between the two types of cells. It had been suggested by other research groups that the mature female gametophyte was involved in processes such as pollen tube guidance (section 3.3.4.1). The synergid cells, which had properties of a secretory cell, were suspected to be the source of a diffusible signal that guided pollen tubes from the funiculus to the micropyle. It was suggested that perhaps *AGL104* played a role in regulating the expression of such signals. In addition, the synergid cells had been implicated in pollen tube reception to ensure fertilization. Thus, regulation of processes that led to pollen tube growth arrest and subsequent release of sperm cells was proposed as a possible role for *AGL104*.

In addition, it was noted that even though each pair of nuclei participating in the double fertilization was genetically identical, the two fertilization products followed completely different developmental pathways, one to develop into an endosperm and the other into an embryo. Type I and MIKC* type genes were reported to be over-represented in the pollen transcriptome (Pina et al., 2005) and recently, expression of all the MIKC* genes have been reported in the embryonic tissue (section 3.3.5). The type I gene, *FEM111*, has been implicated in endosperm development in the female gametophyte. It was suggested that perhaps *AGL104* represented the "complementary" gene of *FEM111* and ensured that the fertilized egg followed the embryo development path. In general, considering the expression and function of type I and MIKC* genes in the male and female gametophyte, the existence of a "parallel" between the development of the angiosperm gametophytes was possible. The importance of functional characterization to answer the above questions is thus emphasized and comparison of the function of these genes in lower plants is expected to shed light on plant evolution.

CHAPTER 4

GAIN-OF-FUNCTION ANALYSIS

4.1 Background

The gain-of-function method of phenotypic analysis, also known as ectopic expression, "overexpresses" the gene of interest outside of its normal domain of expression, that is, in the transgenic plants the gene is expressed in cells/tissues where endogenous transcripts are not normally found. It is expected that this method would alter the fate of the cells/tissues overexpressing the gene of interest, whereby they would likely acquire the fate that this gene normally specifies. However, since MADS-box genes act in a combinatorial manner (Coen and Meyerowitz, 1991; Shore and Sharrocks, 1995), the fate of the cells/tissues would only be altered when all other elements necessary for the specification of this particular fate are also present in these overexpressing cells/tissues. An example is the overexpression of the *Arabidopsis* gene, *AtMYB23* (Kirik et al., 2001), which directs the development of trichomes. Overexpression of *AtMYB23* resulted in trichome formation in unrelated organs like the cotyledon and subepidermal leaf layers.

Overexpression methodologies are principally of two types: the first is where the gene is constitutively active and second where the gene is conditionally active, that is, active only in particular tissues. Constitutively active forms are obtained by fusing the coding sequence of the gene to a strong activation domain, such as the 35S promoter of the Cauliflower Mosaic Virus (CaMV35S) (Gleave, 1992), whereas conditionally active forms are generated by fusing the gene to the heterologous glucocorticoid receptor or other inducible promoters (Simon et al., 1996; Wagner et al., 1999a). A number of genes have been characterized using both the constitutive gain-of-function approaches (Fridborg et al., 1999; Kardailsky et al., 1999; Weigel et al., 2000) and the conditionally active gain-of-function approaches (Schoof et al., 2000; Simon et al., 1996; Wagner et al., 1999a; Wagner et al., 1999b). A well-documented example of the conditionally active approach is the mis-expression of the *Arabidopsis* homeodomain protein WUSCHEL, where the expression was regulated so that the gene was mis-expressed only in a specific region of the meristem (Schoof et al., 2000; Simon et al., 1996; Wagner et al., 1999b).

Overexpression methods are favoured over loss-of-function methods where the loss-of-function alleles are lethal or when knocking out a gene does not result in observable phenotypes. The latter usually occurs when the influence of the gene on a particular phenotype is very slight, or when there are other redundant genes that fulfil the function of the gene that has been knocked out (Kempin et al., 1995). Since redundancy is very common among *Arabidopsis* MADS-box genes due to its duplicated genome, a gain-of-function analysis was necessary to enable successful characterization of *AGL104*. Due to its proven success in previous studies, we opted for a constitutively active promoter approach for this purpose.

4.1.1 Multiple mutants

Taking into account that members of closely related clades usually have similar expression patterns (Alvarez-Buylla et al., 2000a; Alvarez-Buylla et al., 2000b; Nesi et al., 2002; Purugganan, 1997; Rounsley et al., 1995; Theissen and Saedler, 1995), act redundantly (Ditta et al., 2004) and function in a combinatorial manner (Coen and Meyerowitz, 1991; Ditta et al., 2004; Pelaz et al., 2000; Shore and Sharrocks, 1995; Theissen and Saedler, 2001; Thiessen et al., 2001), it was most likely that *AGL104* associated with *AGL66* and *AGL67*, either by acting in a combinatorial manner to form functional dimers or by acting redundantly. The phylogenetic relationship between these three genes was confirmed through comparison of their percentage amino acid sequence similarity (Figure 4.1). These genes share high sequence homology. There is a closer relationship between *AGL104* and *AGL66* (75% amino acid sequence similarity) than between *AGL104* and *AGL67* (47% amino acid sequence similarity) (Figure 1.7 and 4.1) and therefore the interaction between *AGL104* and *AGL66* was expected to be stronger. Moreover, preliminary expression analyses carried out in our laboratory demonstrated that *AGL66* and *AGL67* were expressed in the inflorescence together with *AGL104* (Ambrose, pers. comm.), thus increasing the likelihood of an interaction between these genes.

Assuming that these genes acted in a combinatorial manner, we generated single overexpression mutant lines of *AGL104*, *AGL66* and *AGL67* (35S::*AGL104*, 35S::*AGL66*, 35S::*AGL67*, respectively) and double (35S::*AGL66/104* and 35S::*AGL67/104*) and triple (35S::*AGL66/67/104*) mutants by crossing the respective single overexpression lines (section 2.7.4.1). The multiple mutant lines (double and triple) using *AGL104*, *AGL66* and *AGL67* were

Figure 4.1 Amino acid sequence similarity between *AGL104*, *AGL66* and *AGL67*.

The percentage amino acid sequence similarity between *AGL104* and *AGL66* is 75%, between *AGL104* and *AGL67* is 47% and between *AGL66* and *AGL67* is 49%. *AGL104* and *AGL66* are thus more closely related compared to their relation to *AGL67*. Stars (*) represents conservation of amino acids in all three genes, dots (.) represents change in amino acid in any one of the three genes and dash (-) represents absence/deletion of the respective amino acid. Stars at the end of the sequence marks the end of the gene.

generated to ensure that all elements necessary for the functioning of *AGL104* in a combinatorial manner was present, thus providing an accurate gain-of-function phenotype for *AGL104*.

In addition, we generated double 35S *AGL104* mutant lines (35S::35S::*AGL104*) in order to overexpress *AGL104* at a much higher level than the single 35S::*AGL104* lines. These lines were generated since it had been previously shown that the level of expression of a gene affected mutant phenotypes (Ferrandiz et al., 2000). Mandel et al. (1992) overexpressed the gene *FRUITFUL* (*FUL*) using the single overexpressing promoter (35S) and observed wt phenotype in the mutant lines. Ferrandiz et al. (2000), on the other hand, overexpressed *FUL* using the double 35S promoter (35S::35S) and observed a mutant phenotype. This indicated that the expression level of the gene of interest may affect the gain-of-function phenotype.

We generated a total of 7 and 2 independent insertion lines of the 35S::*AGL104* and 35S::35S::*AGL104* mutant types, respectively. We also generated 2, 4 and 5 mutants of the 35S::*AGL66 104*, 35S::*AGL6⁻ 104* and 35S::*AGL66/6⁻ 104* multiple mutant types, respectively, through crossing. The phenotype of all these mutant types was analyzed. Expression analysis of *AGL104::GUS* (Chapter 3) revealed strong expression in the ovules, mature pollen, germinating pollen, and the elongating pollen tube. This implied that *AGL104* may be involved in aspects of ovule and pollen development. We, therefore, performed *in vitro* and *in vivo* pollen germination experiments using the various mutant lines in order to determine if *AGL104* played a role in pollen germination and pollen tube growth. Time constraints did not permit phenotypic analysis of ovules in these mutants. Our results are presented below.

4.2 Results

4.2.1 Cloning strategy for the 35S::*AGL104* construct

AGL104 was cloned using PCR (section 2.5.6). The fragment was then subcloned into various vectors as shown in Figure 4.2: first into the intermediate vector, pGEM-Teasy®, then into the overexpression vector, pART7 (containing the 35S promoter of CaMV), and finally into the

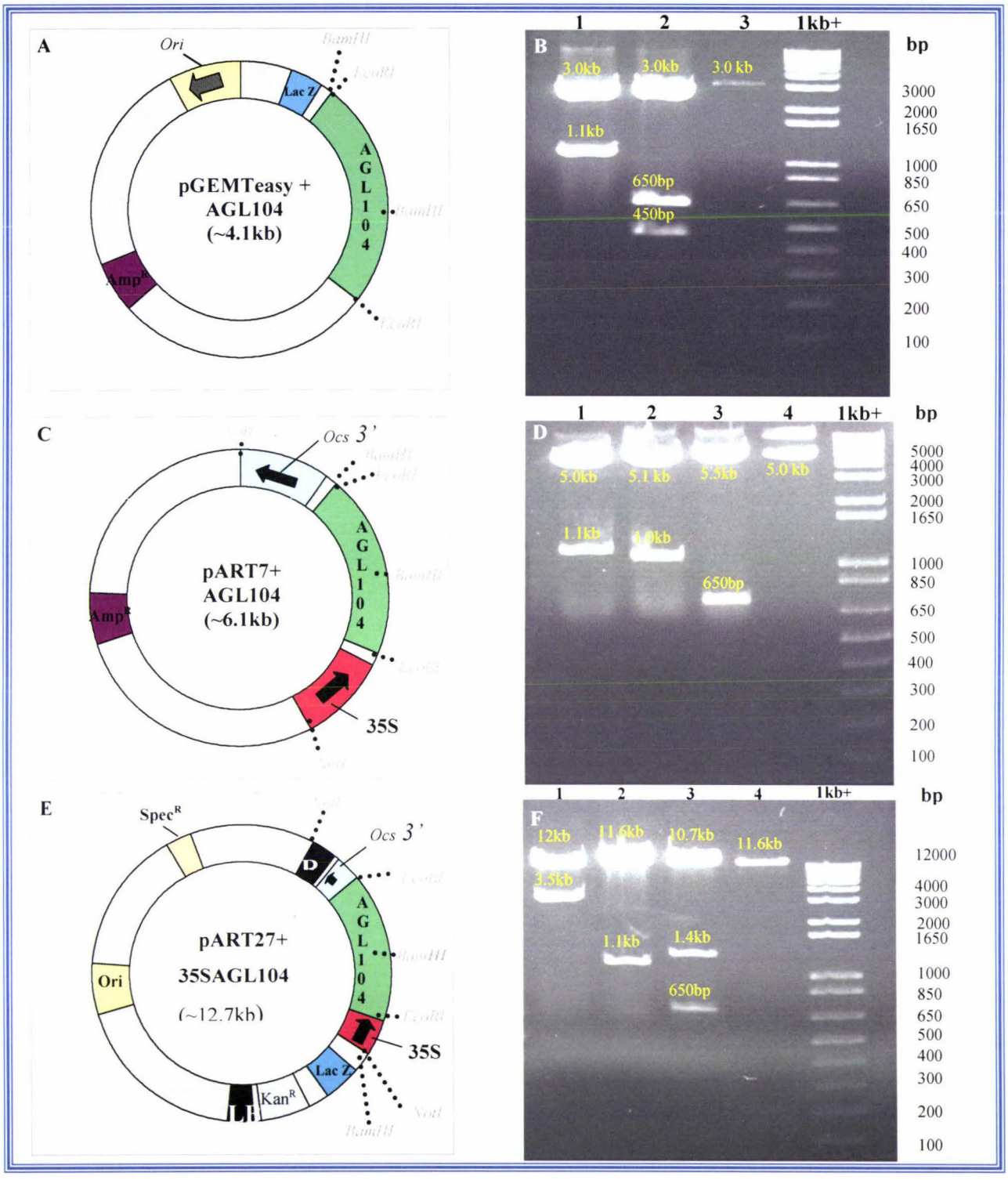

Figure 4.2 Cloning strategy for the single *AGL104* overexpression (35S::*AGL104*) construct.

A,C,E: plasmid maps showing *AGL104* in various vectors. The endonuclease restriction sites shown (grey prints) were used to check the construct or excise and ligate *AGL104* into these vectors.

B, D, F: restriction digests confirming the construction of A, C and E, respectively. Each lane represents the respective construct digested with various enzymes. Fragment sizes are marked. Enzymes used for digestion is listed below. **B:** lane 1- EcoRI; lane 2- EcoRI and BamHI; lane 3- pGEM-Teasy (control). **D:** lane 1- EcoRI; lane 2- HindIII; lane 3- BamHI; lane 4- pART7 (control). **F:** lane 1- NotI; lane 2- EcoRI; lane 3- BamHI; lane 4- pART27 (control).

1kb+- DNA ladder; **Amp^R**- ampicillin resistance; **bp**- base pairs; **35S**- constitutive promoter from Cauliflower mosaic virus; **Kan^R**- kanamycin resistance; **kb**- kilo bases; **Lac Z**-gene used for blue-white selection of transformed *E.coli*; **LB**- left border; **Ori**- origin of replication; **RB**- right border; **Spec^R**- spectinomycin resistance.

binary vector, pART27. The binary vector enabled transformation of the construct into *Agrobacterium* and eventually the plant genome (see section 2.6.2 for details on generating the construct).

The fragment cloned using PCR was confirmed to be *AGL104* through sequencing (section 2.5.10). The sequence of our cloned fragment was identical to the *AGL104* sequence on the database. During the generation of the above construct, the incorporation of the *AGL104* fragment in the correct direction in each vector was verified repeatedly using restriction digests. The restriction enzymes used and the fragments obtained through each restriction digest are presented in Figure 4.2.

4.2.2 Generation of single, double, triple and double 35S overexpressing mutant lines

The *Agrobacterium* containing the above overexpression construct (35S::*AGL104*) was transformed (section 2.7.7) into wt *Arabidopsis* plants to generate single *AGL104* overexpression mutant lines (the 7 lines used for phenotypic analysis were designated *AGL104-1*, *AGL104-2*, *AGL104-3*, *AGL104-4*, *AGL104-5*, *AGL104-6* and *AGL104-7*). The seeds (T_1 generation) collected from the putative transformed plants were germinated on antibiotic (Kanamycin) plates to obtain resistant/mutant seedlings (section 2.8.3), which were then transplanted into soil and allowed to self-fertilize (T_2 generation). Resistant seedlings carrying the construct of interest are viable although some seedlings that are not transgenic are able to escape Kanamycin screening by growing on top of other resistant seedlings. In addition, the self-fertilized population of the transformants (T_2 generation and beyond) segregates in the next generation thus plants grown for phenotypic analysis (or crossing) may or may not contain the overexpression construct. Therefore, to confirm that the plants chosen overexpressed the desired genes, an RT-PCR was performed before utilizing the plants.

Single overexpression lines of *AGL66* and *AGL67* (35S::*AGL66* and 35S::*AGL67*) were generated in a similar manner as above by our lab member, Dr. K. Prasad. The 35S::*AGL66* and 35S::*AGL67* mutant lines were not phenotypically analyzed but were used to generate multiple mutant lines through crossing (section 2.7.4.1). The double 35S *AGL104* construct (35S::35S::*AGL104*), as shown in Figure 4.3, was generated in our lab by Dr B. Ambrose and

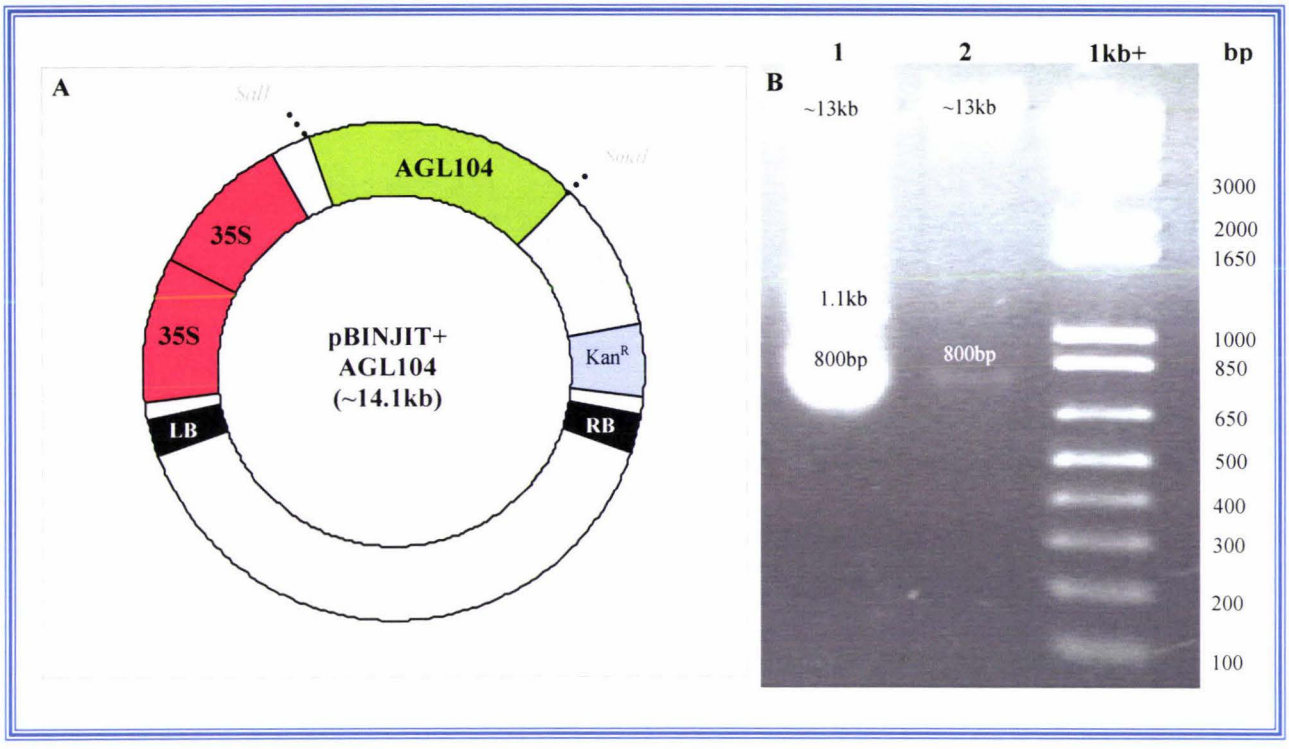

Figure 4.3 Double 35S *AGL104* (35S::35S::*AGL104*) construct (Prepared by B. Ambrose).

A: plasmid map showing *AGL104* in the pBINJIT vector. The endonuclease restriction sites (SalI, SmaI) shown (grey prints) were used to excise and ligate *AGL104* into the vector. **B:** restriction digests confirming the construction of A. Lane 1- construct digested with EcoRI, lane 2- empty vector digested with EcoRI (control). Fragment sizes are marked.

1kb+- DNA ladder; **bp**- base pairs; **35S**- constitutive promoter from Cauliflower mosaic virus; **Kan^R**- kanamycin resistance; **kb**- kilo bases; **LB**- left border; **RB**- right border.

phenotypically analyzed by myself. The 2 double 35S mutant lines used for phenotypic analysis were designated *AGL104-8* and *AGL104-9*.

When generating double overexpression mutants, we crossed the individual single overexpressing lines, each of which were verified through RT-PCR to contain the construct of interest (section 2.7.4.1). For example, when generating 35S::*AGL66/104* double overexpression mutants, we emasculated 35S::*AGL66* flowers and pollinated their stigma with 35S::*AGL104* pollen or vice versa. Likewise, when generating triple overexpression mutants (35S::*AGL66/6⁻/104*), double overexpression mutants (containing combinations of any two of the above three genes) were emasculated and pollinated with pollen from the missing single overexpression mutant. The pollinated stigmas were allowed to grow into mature siliques and the resulting seeds were selected on antibiotic plates. The generation of the plants (e.g. T_1, T_2, T_3 etc) used for either crossing or phenotypic analysis was not taken into consideration since all plants used for these purposes was confirmed by RT-PCR to contain the desired overexpressing construct(s). The multiple mutant lines chosen for phenotypic analysis were designated *AGL66/104-1*, *AGL66/104-2*, *AGL6⁻/104-1*, *AGL6⁻/104-2*, *AGL6⁻/104-3*, *AGL6⁻/104-4*, *AGL66/6⁻/104-1*, *AGL66/6⁻/104-2*, *AGL66/6⁻/104-3*, *AGL66/6⁻/104-4* and *AGL66/6⁻/104-5*.

Preliminary expression studies had shown that *AGL104*, *AGL66* and *AGL6⁻* were expressed in the inflorescence only (Ambrose pers.comm). Since the constitutively active gain-of-function approach used in our study expressed the gene of interest in all tissues of the plant, its expression outside the inflorescence confirmed the successful generation of the mutants. Therefore, we used total RNA extracted from the leaf tissues of putative mutants to test for overexpression of the genes. Total RNA from wt inflorescence and wt leaf tissues were used as positive controls and negative controls, respectively.

Figure 4.4 A shows RT-PCR results confirming the generation of single overexpressing lines of *AGL104*. Bands corresponding to the size of *AGL104* (1.1kb) are marked in the lanes of the Figure. Lanes 1 to 5 represent five 35S::*AGL104* individuals used for phenotypic analysis and crossing. Lane 6 of Figure 4.4 A represents a positive control whereas Lane 7 represents a negative control.

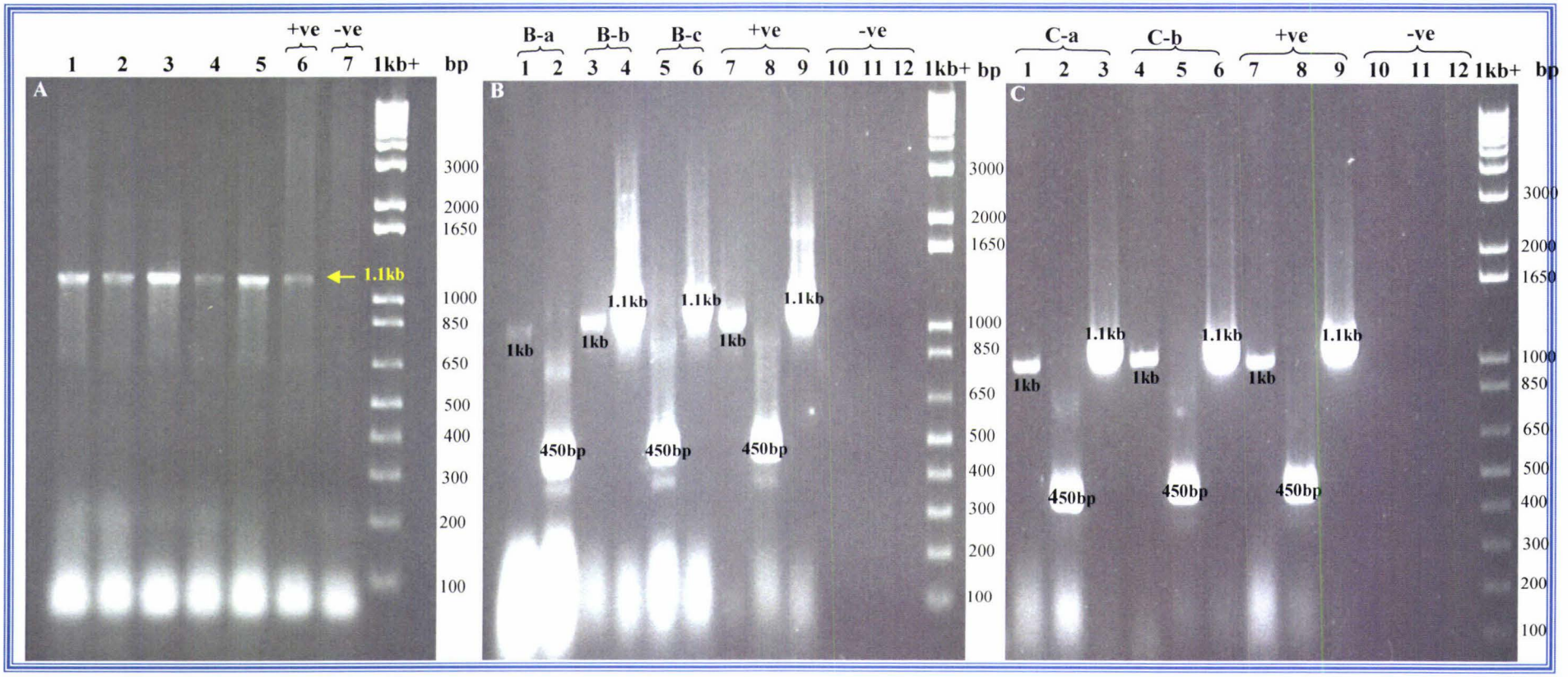

Figure 4.4 RT-PCR showing single, double and triple overexpression lines of *AGL104*, *AGL66*, *AGL67*.

A: Overexpressing *AGL104* lines: lanes 1-5= individual 35S::*AGL104* lines. **B:** Double overexpressing lines: **B-a**= 35S::*AGL66/67*, **B-b**= 35S::*AGL66/104*, **B-c**= 35S::*AGL67/104*. Lanes 1, 3, 7, 10= *AGL66*; lanes 2, 5, 8, 11= *AGL67*; lanes 4, 6, 9, 12= *AGL104*. **C:** Triple overexpressing lines: **C-a** and **C-b**= 35S::*AGL66/67/104*. Lanes 1, 4, 7, 10 = *AGL66*; lanes 2, 5, 8, 11 = *AGL67*; lanes 3, 6, 9, 12 = *AGL104*.

AGL66, *AGL67* and *AGL104* are endogenously expressed in the inflorescence but not in the leaves of the wt plant. The RT-PCR shown above are using total RNA from leaf tissues, therefore confirming the overexpression of these genes in the respective single, double and triple mutant lines.

bp- base pairs; **kb**- kilo bases; **+ve**- positive control (wt inflorescence RNA); **-ve**- negative control (wt leaf RNA); **1kb+** - ladder.

Figure 4.4 B and C shows RT-PCR results confirming the generation of double (35S::*AGL66/67*, 35S::*AGL66/104*, 35S::*AGL67/104*) and triple (35S::*AGL66/67/104*) overexpressing mutants, respectively. Bands corresponding to the size of each gene are marked in the lanes of the Figures. Figure 4.4 B lanes 7-9 and Figure 4.4 C lanes 7-9 represent positive controls whereas Figure 4.4 B lanes 10-12 and Figure 4.4 C lanes 10 – 12 represent negative controls. Figure 4.4 B shows RT-PCR results of only one representative line of each of the double mutants generated (i.e. B-a= representative of *AGL66/67*, B-b= representative of *AGL66/104*, B-c= representative of *AGL67/104*) while Figure 4.4 C shows two representative lines of the triple mutant (C-a and C-b= representatives of *AGL66/67/104*).

4.2.3 Phenotypic Analysis

4.2.3.1 Gross morphological analysis of the overexpressing mutant lines

To determine if overexpressing the genes led to an observable phenotype at the whole plant level, we performed gross morphological analyses of the different overexpressing mutant lines and wt plants.

The phenotype of sporophytic structures such as leaves, stems, siliques and inflorescences of the mutants were compared. As shown in Figure 4.5, no obvious phenotypic variation in these structures between the mutants and wt *Arabidopsis* plants were observed. Panels A to F show whole plants of the different mutant types generated: wt, 35S::*AGL104*, 35S::*AGL66/104*, 35S::*AGL67/104*, 35S::*AGL66/67/104*, 35S::35S::*AGL104*, respectively. Since *AGL104*, *AGL66* and *AGL67* are expressed in the inflorescence tissue (Ambrose pers. comm.), the structure of the inflorescences (panels G-L) and individual flowers (M-R) were also examined under the Light microscope. There were no differences in phyllotaxy or spacing of flowers in the inflorescence between the wt and the mutants (panels G-L). The structure and number of the floral organs also did not exhibit any differences in phenotype (panels M-R) i.e. the number of sepals, petals, stamens and carpel and their shape and size in the mutants and wt flowers were similar. Features such as phyllotaxy and floral organ structure are also sporophytic in origin. In addition, it was noted that all overexpressing mutant lines had fruits and seed set similar to wt and therefore fertility did not seem to be affected in any of the mutant lines.

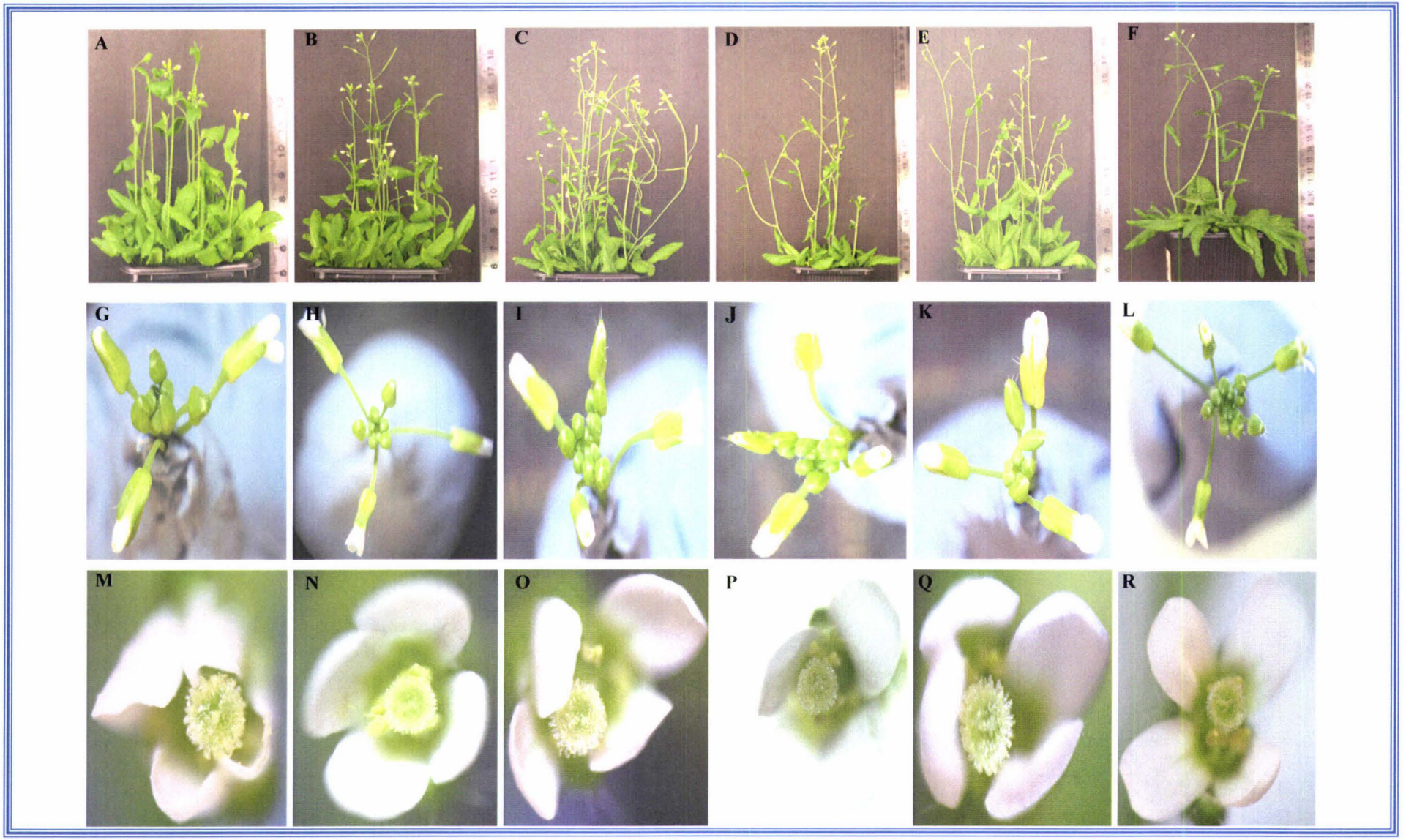

Figure 4.5 Gross morphological phenotype of single and multiple overexpression mutants.

A,G,M- wt; **B,H,N-** 35S::*AGL104*; **C,I,O-** 35S::*AGL66/104*; **D,J,P-** 35S::*AGL67/104*; **E,K,Q-** 35S::*AGL66/67/104*; **F,L,R-** 35S::35S::*AGL104*

Gross morphological phenotype (whole plants, inflorescence and individual flowers) of the single and multiple overexpression mutants were examined. There were no observable changes in the morphology of the mutants compared to the wildtype. In addition, all mutant lines were fertile.

4.2.3.2 In vitro pollen germination

To determine if *AGL104* played a role during pollen germination, we assessed pollen germination frequency of mutant lines overexpressing *AGL104* (35S::*AGL104* and 35S::35S::*AGL104*) and also that of multiple overexpressing mutant lines containing *AGL104*, *AGL66* and *AGL67* (Table 4.1). We collected pollen from stage 13 flowers of all mutant lines and allowed the pollen to germinate *in vitro* for 24 hours in a liquid pollen germination media from Mouline et al. (2002) (section 2.9.1). Following the germination period, 2-3 drops of the germination media containing the pollen were placed on a microscope slide, covered with a cover slip and examined under Light microscope. A minimum of 2 slides were prepared for each sample. An average of 20 digital photographs was taken for each slide. Fully elongated pollen tubes and pollen with any visible protrusion and were scored as "germinated".

Table 4.1 shows that wt pollen germinated at a rate of approximately 38.5%. The various overexpression mutant lines, on the other hand, displayed a much higher germination rate: 35S::*AGL104* = 68.3%, 35S::*AGL66/104* = 50.9%, 35S::*AGL67/104* = 55.3%, 35S::*AGL66/67/104* = 52.6% and 35S::35S::*AGL104* = 52.0%. The results suggested that these genes promoted pollen germination. The highest germination rate of 68.3% was observed for the single 35S::*AGL104* mutant line. Interestingly, the double and triple overexpressing lines containing *AGL66* and *AGL67* (35S::*AGL66/104*, 35S::*AGL67/104* and 35S::*AGL66/67/104*) did not enhance the rate of pollen germination as much as the 35S::*AGL104* line alone. In fact, the pollen germination rates of the double and triple mutant lines were slightly lower (but still statistically higher than wt) compared to the single 35S::*AGL104* line. In addition, the double 35S mutant line (35S::35S::*AGL104*) also had a higher germination rate than the wt but this line, like the double and triple mutant lines, also displayed a slightly lower germination rate than the single 35S::*AGL104* line.

The gross morphology of germinated wt pollen was also compared with that of the mutants. Features such as the number of pollen tubes germinating out of each pollen grain and the difference in diameter (width) of the pollen tubes were examined. The germinated pollen of wt and mutant lines did not display any differences in width or number of pollen tubes protruding out of a single pollen grain (Figure 4.6). Each of the germinated pollen gave rise to a single tube that had an average width of approximately 7μm. This was in agreement with previous

Table 4.1

Percentage pollen germination (*in vitro*) in the different single and multiple mutant lines.

	Mutant type	Number of lines examined	Number of germinated pollen	Total number of pollen	Percentage germination	Average percentage
Wildtype (wt)		**1**	**387**	**1006**	**38.5%**	**38.5%**
Gain-of-function mutants	Single overexpression line (35S::*AGL104*)	2	290	415	69.9%	68.3%
			216	324	66.7%	
	Double overexpression lines (35S::*AGL66/104*)	2	120	224	53.6%	50.9%
			75	156	48.1%	
	Double overexpression lines (35S::*AGL67/104*)	4	164	254	64.6%	55.3%
			83	140	59.3%	
			76	164	46.3%	
			92	180	51.1%	
	Triple overexpression lines (35S::*AGL66/67/104*)	3	122	206	59.2%	52.6%
			80	177	45.2%	
			88	165	53.5%	
	Double 35S overexpression lines (35S::35S::*AGL104*)	2	133	278	47.8%	52.0%
			194	345	56.2%	
Loss-of-function mutants	dsRNAi	5	204	544	37.5%	38.2%
			159	349	45.6%	
			106	307	34.5%	
			169	425	39.8%	
			125	372	33.6%	
	T-DNA	1	223	590	37.8%	37.8%

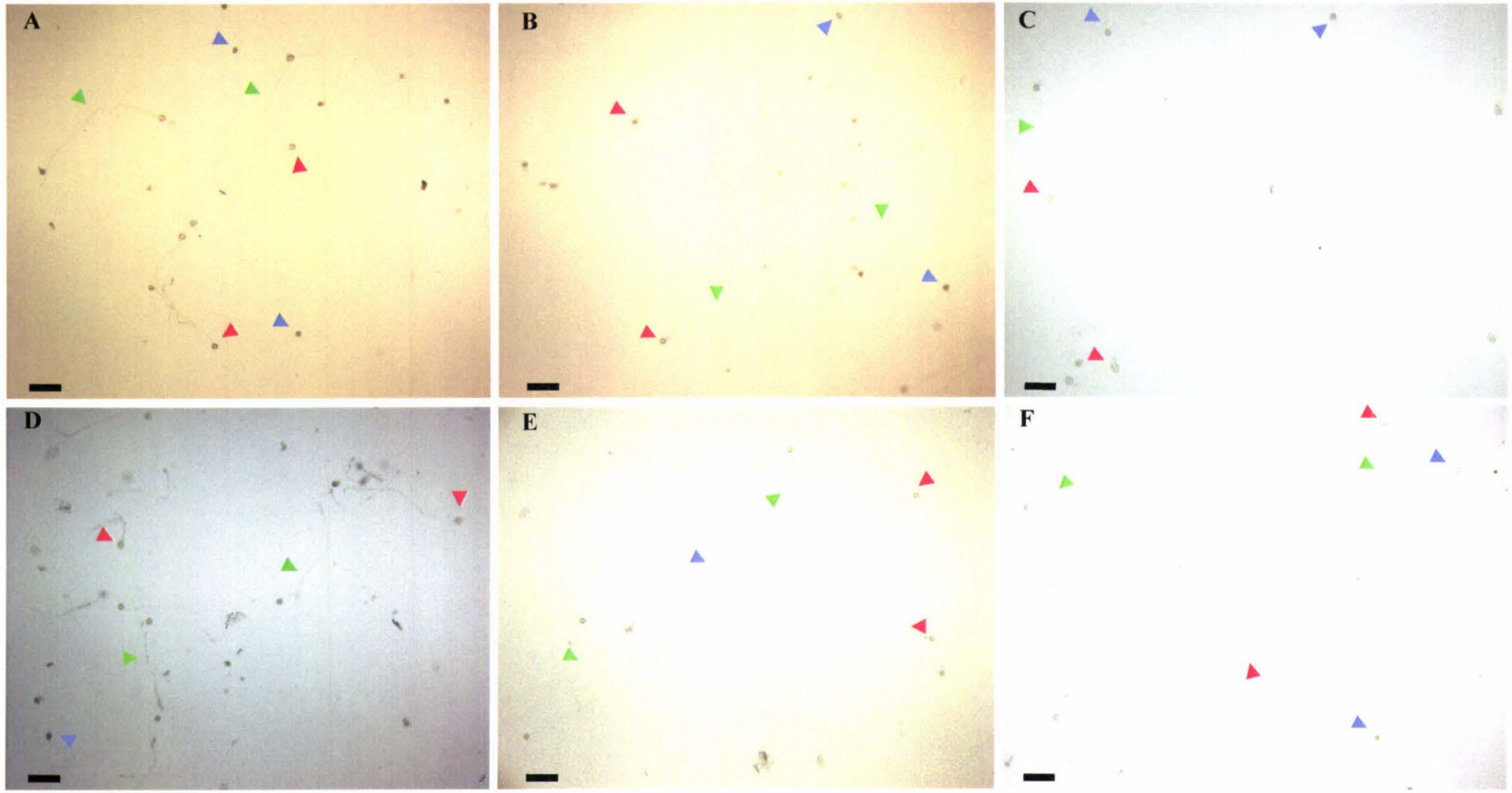

Figure 4.6: Phenotype of i*n vitro* germinated pollen of single and multiple overexpressing lines.

A: wt; **B:** 35S::*AGL104*; **C:** 35S::*AGL66/104*; **D:** *AGL67/104*; **E:** 35S::*AGL66/67/104*; **F:** 35S::35S::*AGL104*.

Phenotype of pollen germinated *in vitro* for 24 hrs (10X). There were no observable changes in the phenotype such as width of the pollen tube or the number of pollen tubes germinating out of each pollen grain. Pollen with visible protrusions or an elongated tube were considered "germinated". A few of the germinated and ungerminated pollen and elongated pollen tubes are marked.

red arrow head- germinated pollen; **blue arrow head**- ungerminated pollen; **green arrow head**- fully elongated pollen tubes.

Scale bar= 100µm

observations of wt *in vitro* germinated pollen, which had straight pollen tubes with a diameter of 5 - 7μm (Derksen et al., 2002).

4.2.3.3 In vivo pollen germination

Pollen germination was investigated in planta (*in vivo* germination) to determine the difference (if any) in the length of the pollen tube between the wt and the various overexpressing mutant lines. Pollinated pistils were excised from the plant after the germination period (2 and 4 hrs), fixed and stained. The Analine blue-stained pistils were examined under the inflorescence microscope and digital photographs were taken. Since the length of individual tubes could not be measured separately, data on pollen tube length was collected by measuring the tubes from the stigma to the longest visible pollen tube in the style. We set multiple pollinations for each mutant type. Therefore, the lengths of pollen tubes reported in Table 4.2 represent the average length of pollen tubes in all pistils of a particular mutant type.

Initially, we germinated the pollen for 2 hrs and measured the length of the pollen tube. As shown in Table 4.2 and Figure 4.7, differences in the length of the pollen tubes ranged from a decrease of 5.8% in the double overexpressing line (35S::*AGL66/104*) to an increase of 9.8% in the triple overexpressing line (35S::*AGL66/67/104*) compared to wt. The single overexpressing line, 35S::*AGL104*, exhibited an increase of 5.8% while the double overexpressing line, 35S::*AGL67/104*, exhibited an increase of 1.8% compared to the wt pollen tube length. Thus, the difference in pollen tube length ranged from a decrease of approximately 6% to an increase of approximately 10%, indicating a considerable amount of change.

We also performed *in vivo* germination for 4 hrs using the 35S::*AGL104* and 35S::35S::*AGL104* lines to determine if the length of pollen tubes between these two mutants and wt differed after a longer germination period. Germination for any period longer than 4 hrs resulted in most pollen tubes reaching the end of the style (results not shown), which made comparison of length more difficult. We excluded the analyses of other mutants after 4 hrs of germination (35S::*AGL66/104*, 35S::*AGL67/104* and 35S::*AGL66/67/104*) due to time constraints and focused on the overexpressing *AGL104* lines only. As shown in Table 4.2 (and Figure 4.8), a significant difference, an increase of 18.8%, was observed in the length of

Table 4.2

Length of pollen tubes germinated *in vivo* in the different single and multiple mutant lines.

	Hours after pollination (hap)	Mutant type	Number of lines examined	Number of pistil pollinated	Average length (μm)	Percentage difference from wt +: increase -: decrease
Gain-of-function mutants	2	**Wildtype (wt)**	-	**10**	**1085**	-
		Single overexpression line (35S::*AGL104*)	7	27	1148	+ 5.8%
		Double overexpression lines (35S::*AGL66/104*)	2	10	1022	- 5.8%
		Double overexpression lines (35S::*AGL67/104*)	2	6	1105	+ 1.8%
		Triple overexpression lines (35S::*AGL66/67/104*)	3	6	1189	+ 9.8%
	4	**Wildtype (wt)**	-	**13**	**1774**	-
		Single overexpression line (35S::*AGL104*)	2	13	2107	+ 18.8%
		Double 35S overexpression lines (35S::35S::*AGL104*)	2	10	1797	+ 1.3%
Loss-of-function mutants	2	**Wildtype (wt)**	-	**10**	**1085**	-
		dsRNAi	5	11	1398	+ 28.8%
		T-DNA	1	4	1148	+ 5.8%

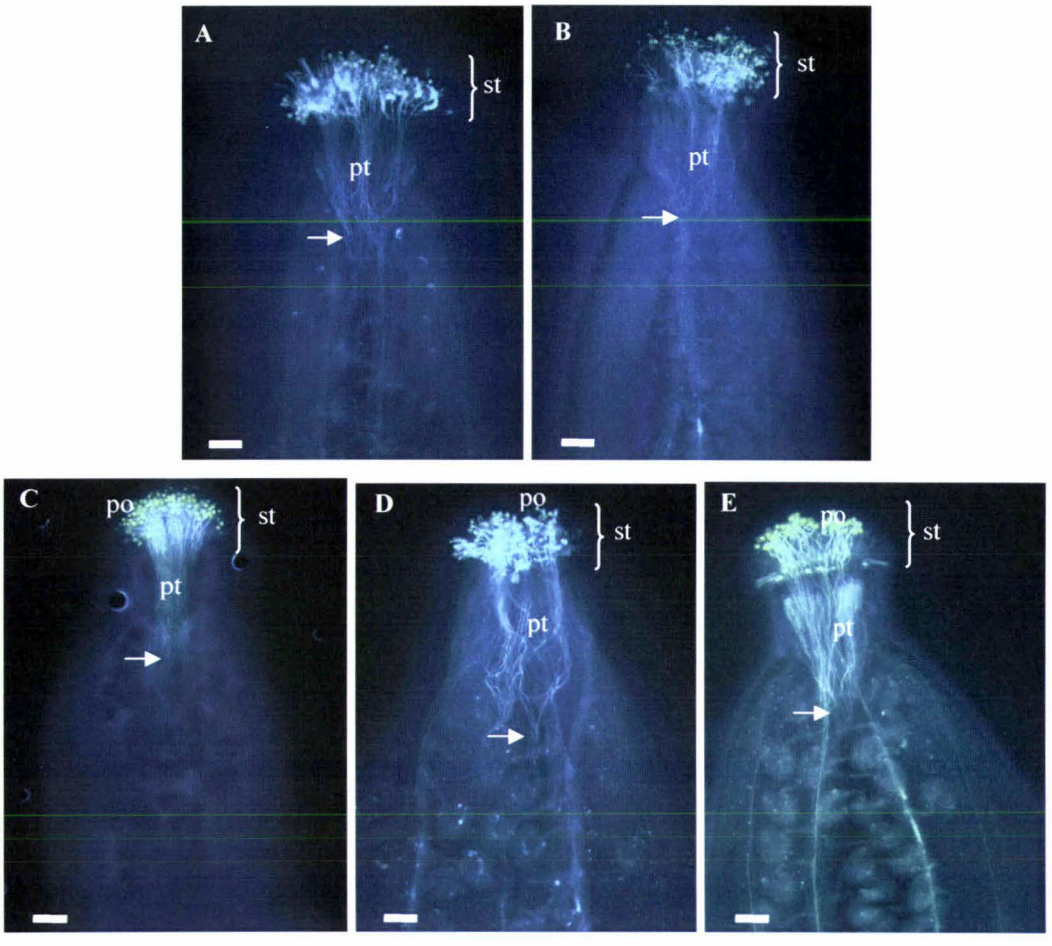

Figure 4.7 Pollen germinated *in planta* for 2hrs

A: wildtype **B:** 35S::*AGL104*; **C:** 35S::*AGL66/104*; **D:** 35S::*AGL67/104*; **E:** 35S::*AGL66/67/104*.

Flowers of respective mutants were emasculated and hand pollinated the following day using pollen from the same plant. Pollen was allowed to germinate on the plant. After 2 hrs, the flowers were picked, fixed and cleared and stained with Analine Blue for 3-5hrs in the dark at rt. Flowers were then examined under the fluorescence microscope (10X) using the DAPI filter. Pollen tubes that had germinated along the style were examined and the length of the longest visible tube was recorded. Multiple flowers were examined for each mutant and an average length of the pollen tubes was calculated.

po- pollen grain; **pt**- pollen tubes; **st**- stigma; **white arrow**- marks the end of the longest visible pollen tube

Scale bar = 100μm

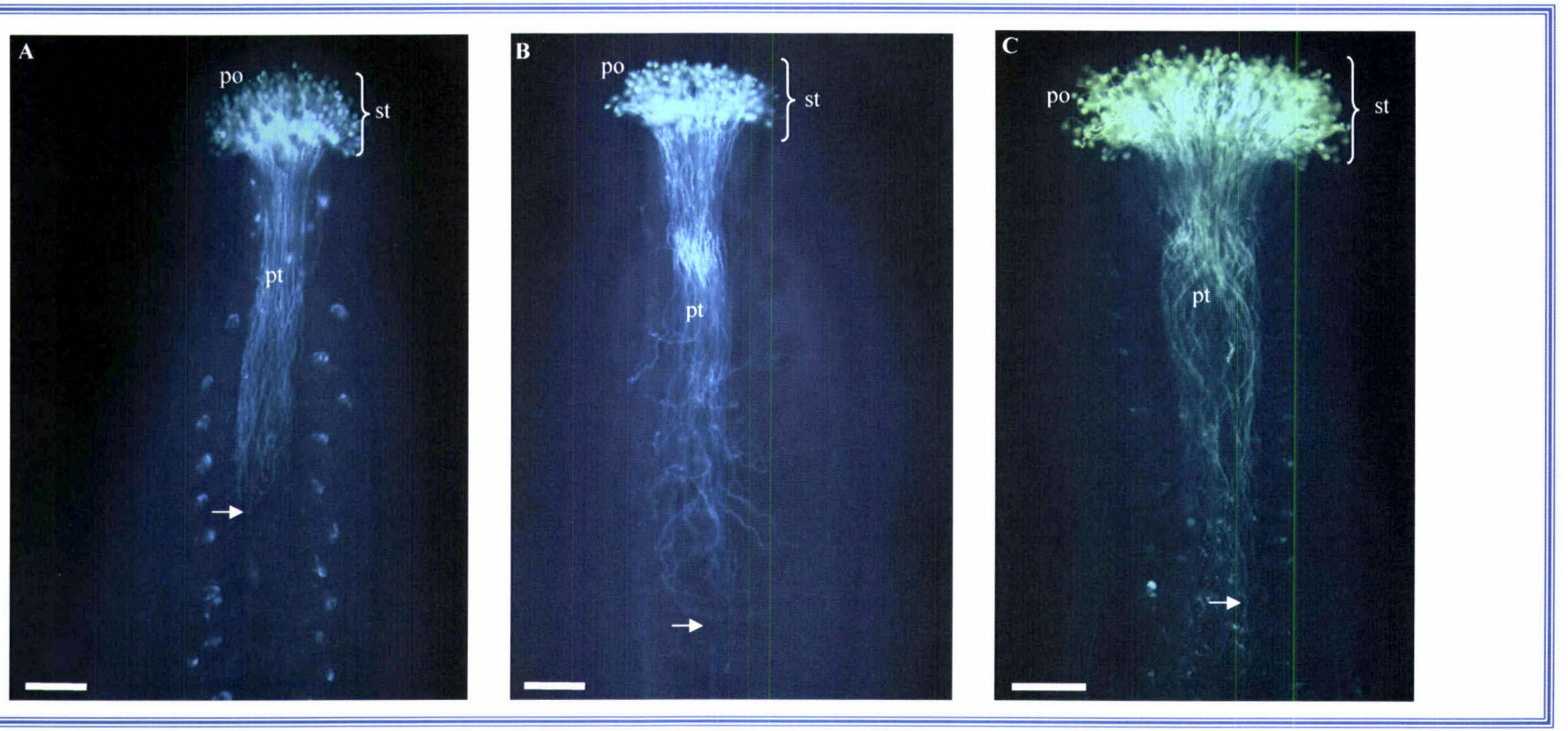

Figure 4.8 Pollen germinated *in planta* for 4hrs

A: wildtype; **B** 35S::*AGL104*; **C:** 35S::35S::*AGL104*.

Flowers of respective mutants were emasculated and hand pollinated the following day using pollen from the same plant. Pollen was allowed to germinate on the plant. After 4 hrs, the flowers were picked, fixed and cleared and stained with Analine Blue for 3-5hrs in the dark at rt. Flowers were then examined under the fluorescence microscope (10X) using the DAPI filter. Pollen tubes that had germinated along the style were examined and the length of the longest visible tube was recorded. A number of flowers were examined for each mutant and an average length of the pollen tubes was calculated.

po- pollen grain; **ov**- ovules; **pt**- pollen tubes; **st**- stigma; **white arrow**- marks the end of the longest visible pollen tube

Scale bar = 100μm

35S::*AGL104* pollen tubes after 4 hrs of germination compared to wt pollen grown for 4 hrs. The 35S::35S::*AGL104* lines, on the other hand, exhibited an increase of only 1.3%.

4.3 Discussion

4.3.1 Generation of multiple mutants of *AGL104*, *AGL66* and *AGL67*

RT-PCR results using total RNA from the leaf tissues (where these genes are not endogenously expressed) of mutants confirmed the generation of overexpression lines of *AGL104*, *AGL66* and *AGL67*. The difference in the expression level (brightness and thickness of bands) of the respective genes among the individual plants seen in Figure 4.4 may have been due to differences in the number of insertions of the overexpressing construct incorporated into the genome of these plants or due to the differences in the amount of total RNA used for the PCR reactions. Since our purpose was to determine whether an individual plant overexpressed the gene of interest or not, the quantity of total RNA used in each PCR reaction was not equalized.

4.3.2 *AGL104* activity in the sporophytic tissues/whole plants

Since *AGL104* was constitutively expressed outside its normal domain of expression, it raised the possibility that the cellular fate of tissues overexpressing these genes would be altered. Morphological analyses of the whole plants of the overexpressing *AGL104* mutants revealed that the phenotype of sporophytic tissues such as leaves, stems and siliques were not altered in the overexpressing lines. In addition, the phenotype of the inflorescence such as its phyllotaxy and spacing distance of individual flowers in the inflorescence were not affected. Looking at the organs of individual flowers, it was observed that the size and shape of floral organs i.e. sepals, petals, stamens and carpel, were also not altered in the mutant lines compared to wt plants. These results implied that *AGL104* did not play a role in the sporophytic tissues of the *Arabidopsis* plant.

Multiple mutants were also generation considering the possibility that *AGL66* and *AGL67* acted in a combinatorial manner with *AGL104* to enable *AGL104* to carry out its role. However, the lack of a mutant phenotype in the sporophytic tissues of the various multiple mutant lines implied that these genes did not act in a combinatorial manner. It may be possible

that *AGL66*, *AGL67* and *AGL104* have acquired new or partially overlapping functions after duplication (section 1.4). Recently, the results of Verelst et al. (2007) discovered that *AGL66* and *AGL104* acted redundantly, and not in a combinatorial manner. In addition, data published on the interaction of MADS-box genes following the commencement of our study showed that *AGL104* interacted with *AGL65* and *AGL97* while *AGL66* interacted with *AGL65*, *AGL82*, *AGL90* and *AGL96* (de Folter et al., 2005). This study did not detect interaction of *AGL67* with any of the MADS-box genes investigated. However, *AGL67* has recently been implicated to play a role in embryo development (de Folter et al., 2004). Considering this, the lack of a mutant phenotype in the sporophytic tissues of the overexpressing *AGL66/104*, *AGL67/104* and *AGL66/67/104* mutant lines is thus not surprising. Further experiments overexpressing the gene pairs that interact to form functional dimers (de Folter et al., 2005; Verelst et a., 2007) can help decipher the roles of *AGL66*, *AGL67* and *AGL104*. In addition, similar to the study carried by Verelst et al (2007), knocking out a combination of genes would reveal the genes that act redundantly.

4.3.3 *AGL104* activity in the male gametophyte (pollen)

Various methodologies such as *in situ* hybridization, Macro- and Micro-analyses, and Northern Blot analyses had established a strong expression of *AGL104* in the mature pollen (Becker et al., 2003; Honys and Twell, 2004; Kofuji et al., 2003; Parenicova et al., 2003; Pina et al., 2005). The results of our *AGL104::GUS* expression analysis (Chapter 3) confirmed these observations. The mature pollen, that is the male gametophyte, has two primary functions: production of two sperm cells (SC) and their transport within the pollen tube and through the tissues of the style and ovary into the embryo sac of the ovule, where they participate in double fertilization (Mascarenhas, 1993). When a pollen grain is released from the anther it exists as a free organism until it lands on the stigma of a pistil, where it then begins another phase of its life. If the pistil is compatible, the pollen germinates and extrudes a tube within which the sperm cells are transported to the embryo sac.

The growth of the pollen tube is controlled by the vegetative cell (VC) (Mascarenhas, 1993) (Figure 1.5). As the pollen tube elongates, the VC and the two SC exits the pollen grain and travels at the tip of the growing pollen tube (McCormick, 2004). In elongating pollen tubes, the vegetative nucleus (VN) was observed to have a high level of activity whereas the SC had a very low level of general activity (Derksen et al., 2002; Wagner et al., 1990). Linking the

LAT52 promoter of tomato, which is a late pollen-expressed gene, to the GUS protein gene revealed that the LAT52 promoter was active specifically in the vegetative nucleus. This proved that late pollen-expressed genes were activated in the VC (Tanaka, 1997; Twell, 1992). As mentioned in Chapter 3, late pollen-expressed genes have been implicated in pollen germination and tube elongation (Mascarenhas, 1990).

Our results established that *AGL104* continued to be expressed during the initial pollen germination and tube growth stages and also at later stages of tube elongation since GUS activity continued to be detected at the growing tips of the pollen tube that had been elongating for 24 hrs (Figure 3.4). Since *AGL104* is a late pollen-expressed gene, these results implied that *AGL104* played a role during pollen germination and pollen tube elongation. Similar expression patterns have been detected for other late pollen-expressed MADS-box genes in other species, such as *ZmMADS2* in maize and *DEFH125* in *Antirrhinum* (Heuer et al., 2000; Zachgo et al., 1997), both of which were proposed to function during pollen germination and tube elongation. To investigate the involvement of *AGL104* during pollen germination and tube elongation, we performed *in vitro* and *in vivo* pollen germination experiments for the mutants generated.

4.3.3.1 AGL104 activity during pollen germination

In vitro pollen germination performed in our study indicated that overexpression of *AGL104* led to an increased rate of pollen germination in the overexpressing mutants compared to the wt (Table 4.1). Pollen germination frequency of wt pollen was 38.5% while the frequency of other mutants ranged from 50.9% to 68.3%. Tricellular pollen (such as pollen from *Arabidopsis*) are known to be more difficult to germinate *in vitro* (Taylor and Hepler, 1997), which is thought to be due to factors such as a low density of secretory vesicles at the tip; preferential location of mitochondria at the wall; and moderate velocity of the organelle population (Derksen et al., 2002). The observed high speed of growth of pollen *in vivo*, therefore, it thought to relate to an efficient supply of nutrients and to unknown growth factors from the female sporophytic tissue(Derksen et al., 2002).

The double 35S mutant line (35S::35S::*AGL104*) also exhibited a higher germination rate compared to wt but its germination rate was lower (52.0%) compared to the single 35S::*AGL104* line (68.3%). We had generated the 35S::35S::*AGL104* lines in order to express

AGL104 at a higher level compared to the 35S::*AGL104* lines. The lower germination frequency in the 35S::35S::*AGL104* lines suggests that either it did not express *AGL104* at a higher level or higher expression levels of *AGL104* does not increase germination frequency. To differentiate between these possibilities, it would be necessary to assess the expression level of *AGL104* in the 35S::*AGL104* and 35S::35S::*AGL104* lines either by Northern Blot analysis or quantitative RT-PCR. This information can then be used to determine if the germination-promoting phenotype is dosage-dependent, that is, if it is affected by the level of *AGL104* expression.

The increased germination frequency of the overexpressing *AGL104* lines (single and double 35S), combined with the strong expression of *AGL104* in the mature pollen (Honys and Twell, 2004; Kofuji et al., 2003; Parenicova et al., 2003; Pina et al., 2005) and in the germinating pollen (Chapter 3), implied that overexpression of *AGL104* resulted in the synthesis of germination-promoting proteins either **before** the pollen was shed or **during** the germination process itself. To elucidate the dependence of *Arabidopsis* pollen germination on active transcription, Honys and Twell (2004) used Actinomycin D to show that blocking transcription had moderate effects on the frequency of pollen germination and rate of tube elongation. On the other hand, use of a translation inhibitor drastically decreased pollen germination and tube elongation (Honys and Twell, 2004). Thus, Honys and Twell (2004) concluded that pollen germination and tube elongation in *Arabidopsis* occurred relatively independently of transcription but was dependent on translation. Considering this, the increase in pollen germination frequency of the overexpressing *AGL104* lines (Table 4.1) thus suggests that most likely *AGL104* acted as either:

- an **activator** of germination-**promoting** genes **before** mature pollen was shed
- a **repressor** of germination-**inhibiting** genes **during** the germination process. In other words, *AGL104* was a "**repressor of a repressor**".

In the latter circumstance, overexpressing *AGL104* would lead to a higher degree of repression of the inhibiting genes, thus rendering them non-operational, which would then allow pollen germination to proceed. In both scenarios above, blocking transcription would not result in an observable phenotype in the frequency of germinated pollen. To differentiate whether *AGL104* functioned prior to pollen germination or regulated gene expression during the process, it is

necessary to determine the downstream genes of *AGL104*. Examining the promoter region of genes that are upregulated and downregulated when *AGL104* is expressed may reveal this information. Examining the changes (if any) in the expression level of these genes both before and during pollen germination would then determine when *AGL104* regulates these genes. In addition, localization of *AGL104* into the nuclei of the vegetative cell (section 3.3.3.2) during pollen germination would also be indicative of active regulation during the process.

4.3.3.1.1 Interaction of AGL104 with AGL66 and AGL 67 during pollen germination

Considering pollen germination frequencies, Table 4.1 shows that the 35S::*AGL104* lines led to the highest increase in pollen germination (68.3%) compared to wt (38.5%). Even though the double (35S::*AGL104/66* and 35S::*AGL104/67*) and triple (35S::*AGL104/66/67*) mutants led to an increase of 50.9%, 55.3% and 52.6%, respectively, the i*n vitro* results of our study indicated that overexpression of *AGL66* and *AGL67* did not have a cumulative effect on the germination rates of the mutant pollen (Table 4.1). This suggested that either *AGL66* and *AGL67* did not function during pollen germination or, if they did, they did not interact physically with *AGL104* (section 4.3.2) to effect germination. In other words, *AGL104* alone was sufficient to promote pollen germination and did not require interaction with *AGL66* and *AGL67* to initiate pollen germination.

Transcriptome and expression analyses of pollen, nevertheless, suggested a role for *AGL66* in late pollen development (Honys and Twell, 2004; Kofuji et al., 2003; Pina et al., 2005). The lower germination frequency in the double (35S::*AGL66/104*) and triple (35S::*AGL66/67/104*) overexpressing mutants compared to the single 35S::*AGL104* line (Table 4.1) could be explained by the existence of competition for the formation of functional heterodimers to control gene expression. Verelst et al. (2007) showed that, during pollen germination, *AGL104* interacted with *AGL65* and *AGL30* while *AGL66* interacted with *AGL65*, *AGL30* and *AGL94*. Formation of correct heterodimers was sufficient for their translocation to the nuclei to effect gene regulation (Verelst et al., 2007). When we overexpressed *AGL104* alone (35S::*AGL104* line), the percentage germination rate was 68.3%. However, when we overexpressed both *AGL104* and *AGL66* (35S::*AGL66/104*), the germination percentage dropped to 50.9%. The slight reduction in the germination rate of the 35S::*AGL66/104* line compared to the 35S::*AGL104* line could have been due to competition of *AGL65* and *AGL30* to bind to either *AGL66* or *AGL104* in order to form functional heterodimers: *AGL30/66*, *AGL65/66*,

AGL30/104 and *AGL65/104* (Verelst et al., 2007). Since *AGL30* and *AGL65* were expressed at endogenous levels only in our mutants, they would have been able to form a certain number of functional heterodimers before being depleted while the overexpressed *AGL104* and *AGL66* proteins would have been present in excess quantities.

Future analysis of the pollen germination frequency in the 35S::*AAGL66* lines (generated in our study but not analysed phenotypically) and its comparison with the pollen germination frequency of 35S::*AGL104* lines will enable us to determine if these two genes have a comparable effect on pollen germination or if one gene enhances pollen germination more than the other. In addition, future experiments designed to simultaneously overexpress pairs of genes involved in the formation of functional heterodimers (Verelst et al., 2007) would help determine how these MIKC* type genes interacted to promote pollen germination and what other roles they played during pollen development. These overexpressing mutants can also be analysed for *in vivo* pollen germination to determine if the appropriate functional heterodimer-forming genes have an effect on the rate of pollen tube elongation in addition to their effect in pollen germination.

4.3.3.2 AGL104 activity during pollen tube elongation

To determine if pollen tube elongation was affected by overexpressing *AGL104*, *in vivo* pollen germination was performed for the various single and multiple mutant lines. As indicated in Table 4.2 and Figure 4.7, the length of the pollen tube of the mutants compared to wt ranged from a decrease of 5.8% to an increase of 9.8% when measured 2 hrs after pollination (hap). In contrast to the *in vitro* germination results (Table 4.1), the triple overexpressing mutants (35S::*AGL104/66/67*) had the highest increase in the length of its pollen tubes while the single overexpressing line, 35S::*AGL104*, showed an increase of only 5.8%. At first glance, it appeared that interaction of *AGL104* with *AGL66* and *AGL67* had a cumulative effect on the length of the pollen tubes in contrast to their effect on the germination frequency (Table 4.1 and section 4.3.2.1.1). However, considering the tube lengths of the double overexpressing mutants, 35S::*AGL104/66* and 35S::*AGL104/67*, this conclusion became invalid since these two lines exhibited shorter tube lengths compared to the 35S::*AGL104* line. Therefore, we could not reach a definitive conclusion regarding a cumulative effect (or lack of it) based on the available data.

It is also noted that the *in vivo* germination experiment may have led to a discrepancy in the total amount of time pollen on each individual pistil were allowed to germinate before being fixed and examined. Since pollination of emasculated pistil required searching for fresh, viable pollen (section 2.7.4), there was a time delay of 45min to an hour between the first and the last pistil pollinated on a particular day. This would have resulted in some pollen tubes having elongated for up to 5 hrs while others had elongated for only 4 hrs. This difference of an hour would have had a significant effect on the final length of the pollen tubes at the time of measurement. This discrepancy can be eliminated by pollinating and examining a fewer number of pistils each day. In the current study, we examined a maximum of fifteen pistils per day. Reducing this to five pistils per day would reduce the time difference between pollinations. In addition, it is also noted that the pollen tubes did not grow perfectly straight down the style (Figure 4.7 and 4.8). Some pollen tubes coiled a few times as they elongated. Since we measured the tube length from the top of the stigma to the tip of the longest pollen tube (and averaged the data) (section 2.9.2), this did not allow an accurate measure of its true length. In order to negate the discrepancies in time and growth pattern (coiled or bent) of pollen tubes, a larger sample, i.e. more pollinations per mutant line is required in order to get an accurate data on the length of pollen tubes.

Since after 2 hrs the pollen tubes had lengthened to just below the stigma, it was likely that this represented the initial elongation, the raw materials (mRNAs and proteins etc.) for which is known to be present in mature pollen before it is shed (Honys et al., 2000; Kim et al., 2002; Mascarenhas, 1989; Muschietti et al., 1994; Muschietti et al., 1998; Schrauwen et al., 1990). Therefore, a changes in the length of pollen tubes caused by the action of regulatory genes was likely to become apparent after the stored raw materials were used up and the pollen initiated new protein synthesis (Honys and Twell, 2004). Assessing the data on pollen tube length collected at 4hap of mutants overexpressing *AGL104* revealed a pattern (Table 4.2). The single overexpressing line, 35S::*AGL104*, displayed a significant increase of 18.8% in its tube length at 4hap compared to an increase of 5.8% at 2hap. It is possible that overexpression of *AGL104* affected the processes active **during** tube elongation and led to an increased length of the pollen tubes. The negligible increase (1.3%) in pollen tube length of the 35S::35S::*AGL104* line compared to wt may reflect a lower expression level of *AGL104* in this mutant compared to its expression level in the 35S::*AGL104* line (section 4.3.3.1). Evaluation of the tube lengths of the other multiple (double and triple) mutant lines at 4hap (not yet analysed) may also reveal a pattern.

It has been suggested that pollen tube elongation is strictly dependent on **active** translation (Honys and Twell, 2004) and the increase in the tube length of mutants overexpressing *AGL104* (35S::*AGL104*) implied that *AGL104* promoted translation. Given that *AGL104* is a transcription factor and not a protein that forms part of the translation apparatus, the implication that it promoted translation was intriguing in view of the fact that it most probably did not do so by simply activating/transcribing genes encoding proteins used during translation (Honys and Twell, 2004). Therefore, the mechanism used by *AGL104* to promote tube elongation was curiosity-provoking.

Upon relating the transcriptional independence of pollen tube elongation to the *in vivo* germination results showing that *AGL104* promoted tube elongation (Table 4.2), it seemed likely that during tube elongation *AGL104* resulted in:

- **repression** of genes that **inhibited** pollen tube elongation, thus allowing the tubes to increase in length.

In agreement with the results of (Honys and Twell, 2004), blocking transcription would not have had an observable effect on tube elongation since in this scenario active transcription was not the determining factor during tube elongation. Existence of genes that inhibited pollen tube elongation would seem to be necessary in nature since pollen tube growth ceases once it reaches an ovule (Huck et al., 2003) (Chapter 3), thus indicating a need for a regulatory mechanism. It is possible that *AGL104* is part of this regulatory mechanism where it suppresses genes that inhibit tube elongation while the pollen tube grows towards an ovule and once the ovule is reached, *AGL104* becomes non-functional and thus leads to these genes being expressed resulting in a cessation of tube growth.

It is also noted that since *AGL104* is expressed in the embryo sac (Chapter 3), it may be involved in pollen tube guidance. Due to time constraints, we were unable to examine the phenotype of the female gametophyte in the overexpressing mutants generated for gain-of-function analysis. It would not be surprising to discover that pollen tube guidance by the embryo sac also affected the rate of pollen tube elongation, thus influencing the *in vitro* and *in vivo* pollen germination assays performed in our study. Overexpressing *AGL104* may result in the mutant embryo sac eliciting a faster growth of pollen tubes. Therefore, to accurately determine the effect of *AGL104* in pollen tube elongation, the phenotype of the embryo sac in

the overexpressing mutants and their effects on the rate of tube elongation needs to be established.

However, from our current set of *in vitro* and *in vivo* pollen germination results, we can conclude that:

- *AGL104* promotes pollen germination by acting as an activator of germination-promoting genes before the mature pollen is shed or by acting as a repressor of germination-inhibiting genes during the germination process.
- *AGL104* does not affect germination by interacting with *AGL66* and *AGL67* in a combinatorial manner. However, it most likely acts redundantly with *AGL66*.
- *AGL104* promotes pollen tube elongation by repressing genes that inhibit pollen tube elongation.

The challenge then is to determine what these downstream genes are, their functions during pollen germination and tube elongation and how they are regulated by *AGL104*.

4.3.3.3 Processes involved during pollen germination and tube elongation

Even though a number of transcription factors have been identified in pollen of various species (Chapter 3), concrete evidence on the type or classes of genes regulated by these transcription factors and the mechanisms and regulatory pathways utilized during pollen germination and tube growth are still lacking in the current literature. So far, the putative genes activated by the pollen-expressed transcription factors have been deduced from promoter analyses of genes showing similar expression patterns as the transcription factors themselves.

The expression of novel genes during pollen germination and tube elongation (compared to the genes expressed before the pollen is shed) is a good indicator of the existence of a unique genetic program in the pollen. Despite some successful attempts (Guyon et al., 2000), overall it has been difficult to identify genes that are transcribed only upon pollen germination or to identify mRNAs that are essentially translated only after pollen germination (Wittink et al., 2000). In the plant species, *Tradescantia*, which has active transcription occurring during pollen germination, analysis suggested that the mRNAs present in the non-germinated pollen

grain of *Tradescantia* were the same as the mRNAs that were newly synthesized during germination (Mascarenhas and Mermeisten, 1981). However, it was noted that the method used in their study did not have the capability to detect rare or moderately rare mRNAs, thus the possibility that new genes were expressed during pollen germination was maintained.

Recently, methods such as Microarray or pollen transcriptome analyses have shed light on the genes that are transcribed in the male gametophyte. Using the Affymetrix ATH1 arrays, which covers approximately 80% of the *Arabidopsis* genome, recent studies showed that the variety of genes expressed in pollen were approximately three-folds fewer than the genes expressed in vegetative tissues (Pina et al., 2005). This was due to either the pollen transcripts originating from a single cell type (as compared to vegetative tissues that have a variety of cell types) or the transcriptome reflecting the specialization of pollen towards specific activities. Analysis of the putative functions of the expressed genes in the smaller but unique transcriptome of the pollen led to the discovery that genes with proposed functions in signal transduction, cell wall biosynthesis, vesicle transport, ion dynamics, and cytoskeletal dynamics were either selectively (exclusively) expressed in pollen or were pollen-enriched (transcripts were more abundant in pollen compared to other tissues) (Becker et al., 2003; da Costa-Nunes and Grossniklaus, 2003; Feijo and Moreno, 2004; Hepler, 2001; Honys and Twell, 2003; Honys and Twell, 2004; Pina et al., 2005). The involvement of the above processes (signal transduction, cell wall biosynthesis, vesicle transport, ion dynamics, cytoskeletal dynamics) during pollen germination and tube elongation is outlined.

Earlier microscopy studies of growing pollen tubes revealed a characteristic zonation in which the apical region of the tube possessed a clear cap with more granular elements behind it. These granular internal components exhibited a vigorous "reverse fountain" cytoplasmic streaming (Taylor and Hepler, 1997). Recently, Derksen et al. (2002) germinated *Arabidopsis* pollen *in vitro* and looked at its growth; cellular organization and ultrastructure; cytoskeletal organization; organelle movement; deposition and structure of the wall; and the occurrence of coated pits, all elements assumed to be relevant for tip growth. Their study revealed a growing tip where large organelles were absent; wall secretion was facilitated by a single type of secretory vesicle (Golgi); and the presence of a layered primary cell wall that disappeared after secretion of the secondary, callosic wall. This research group also established that individual organelles moved in distinct lanes aligned parallel to the long axis of the tube except in the tip region, where they were more randomly distributed.

During pollen tube elongation, growth takes places exclusively at the apex of the tube (Mascarenhas, 1993). In trying to elucidate the mechanisms involved in pollen tube elongation, scientists initially assumed that the molecular mechanisms that shaped and controlled growth existed at the tip. Given the large number of individual events involved and the speed at which pollen tubes elongated, it was also assumed that the nature of the regulatory events was governed by biophysical phenomena or fast chemical/biochemical reactions (Feijo et al., 2001). This implied that the regulatory mechanisms involved only the organelles and cytoplasmic structures present in the tip area and these structures functioned as a dynamic condition maintained by a localized process (Feijo et al., 2001). According to these assumptions, therefore, pollen tube growth could **not** be regulated by direct gene expression or enzyme induction since the apparatus for these processes were not present in the immediate growth zone (Derksen et al., 2002; Feijo et al., 2001). As mentioned previously, Honys and Twell (2004) recently showed that germination and tube growth were independent of transcription in *Arabidopsis*, thus lending support to the above assumptions.

Among the structures that are situated most closely to the locus of growth are the Golgi vesicles, ion channels and a few randomly organized actin- and micro-tubules (Derksen et al., 2002; Feijo et al., 2001), all of which act together to bring about the extension/growth of the pollen tube. The Golgi-derived vesicles contain precursors of the pollen tube wall and the fusion of these vesicles with the plasma membrane and the release of its contents lead to the extension of the pollen tube (Becker et al., 2003; Feijo et al., 2001). Fusion of the Golgi vesicles with the plasma membrane is regulated by the transport of inorganic ions such as calcium (Ca^{2+}) and potassium (K^{+}) across the plasma membrane (Fan et al., 2001). There is evidence that intracellular Ca^{2+} gradients and pollen tube growth exhibit oscillatory behaviour as indicated by the periodic depositions of cell wall material. Immuno-cytochemistry illustrated that arabinogalactans and pectins that form the pollen tube cell wall were deposited in a ring-like pattern with remarkable periodicity along the length of tobacco and petunia pollen tubes (Feijo et al., 2001). In addition, the frequency of these rings was roughly correlated with the periodic changes in the growth rate, which in turn was correlated to the concentration of cytosolic Ca^{2+} ($[Ca^{2+}]$). In the growing pollen tubes, elevation of cytosolic $[Ca^{2+}]$ resulted in a faster growth rate whereas depression of cytosolic $[Ca^{2+}]$ led to a slower growth rate (Feijo et al., 2001; Messerli and Robinson, 1997; Pierson et al., 1996).

Potassium ions (K^+) were also shown to be important for both pollen germination and tube growth, since a decrease in the external concentration of K^+ ($[K^+]$) inhibited pollen germination and tube growth whereas an increase in external $[K^+]$ stimulated pollen tube growth but inhibited pollen germination slightly (Fan et al., 2001). It was not surprising, therefore, that pollen transcriptome analyses found transcripts of ion transporters to be selectively expressed in pollen (Becker et al., 2003). In addition, recent expression studies and functional analyses discovered a gene encoding a K^+ channel, *Shaker Pollen Inward K^+ channel* (*SPIK*), to be specifically expressed in pollen and pollen tubes of *Arabidopsis* (Mouline et al., 2002). Consistent with the results of Fan et al. (2001), the *spik-1* mutant displayed a slight reduction in pollen germination but a significant reduction in pollen tube length due to its inability to take up enough K^+ to sustain tube elongation (Mouline et al., 2002).

It is well established that the Golgi vesicles delivering the raw materials to the tube apex are transported by an acto-myosin cytoskeletal system (Becker et al., 2003). However, the exact direction of the vesicles to the point of fusion and the regulation of docking and exocytosis events were, until recently, unknown. Initial studies were not able to detect actin in the region where Golgi vesicles fused, which was further compounded by the fact that phase contrast (DIC) microscopy revealed that the organelle movement in the tip zone exhibited chaotic movement rather than movement in organized linear lanes observed in regions further away from the tip (Feijo et al., 1995; Mascarenhas, 1993). This led to the proposal that physio-chemical forces present in the apical cytoplasm were instrumental in orchestrating the organization of the vesicles and in regulating their fusion to a specific region in the plasma membrane.

Large polarized ionic currents that could act as an electric field to promote movement of charged particles according to the electric attracting and repulsive forces of the vesicle membrane proteins were found to exist in cells (Feijo et al., 2001; Jaffe et al., 1974). In addition, it was observed that isolated vesicles moved on an electric field (Heslop-Harrison and Heslop-Harrison, 1982). Since ion channels were present in the tip region, the polarised spatial distribution or activation of ion carriers and pumps was proposed to be the elements directing Golgi vesicles to the plasma membrane and was thus the basis of the regulatory mechanism controlling pollen tube growth (Feijo et al., 2001). It was concluded that these ion channels would ensure that reaction-diffusion patterns were easily transmitted spatially, both through signalling cascades and other reaction-diffusion mechanisms.

Recently, however, studies using fluorescence staining and immunohistochemistry confirmed the presence and a highly random and irregular distribution of microtubules and actin filaments in the pollen tip (Derksen et al., 2002), thus opening the possibility that these structures, together with the ion channels, were after all involved in Golgi vesicle guidance to the point of fusion on the pollen tube plasma membrane. The randomness of the cytoskeleton at the tip of the pollen tube could be purposely designed to enable release and free movement thereon of the Golgi vesicles to the plasma membrane for fusion.

4.3.3.4 Types of genes regulated during pollen germination and tube elongation

Given that the mechanism controlling pollen tip growth most likely exists in the tip region and involve only the structures present there (Feijo et al., 2001), it is possible that *AGL104* plays a role in pollen tube elongation by influencing the synthesis or activity of these structures. It may be possible that *AGL104* controls the synthesis of microtubules or actin filaments since similar to *AGL104*, genes involved in cytoskeleton synthesis is known to be selectively expressed and enriched in the pollen transcriptome (Becker et al., 2003; Honys and Twell, 2004; Pina et al., 2005). Recently, genes encoding small GTPases of the RAB, ARF and Rho/RAC families have been implicated in vesicle trafficking, cytoskeleton synthesis and signalling in the pollen tubes (reviewed by Samaj et al., 2006). It was reported that these genes, together with their regulatory proteins, were essential for spatio-temporal regulation of the above processes.

Studies have firmly established that pharmacological and genetic perturbation of actin assembly ultimately resulted in inhibition of polar growth in pollen tubes (Samaj, 2004). The RAB proteins, which are molecular markers associated with endomembranes that recruit phosphoinositide kinases, were localized to the tip region of the elongating pollen tubes. These RAB proteins played a role in endocytosis, which is essential for sustained tip growth. It was shown that localization and thus function of RAB was dependent on an intact actin cytoskeleton (de Graaf, 2005). In addition, the apical plasma membrane contained Rho/RAC GTPases that performed a variety of functions such as organization of actin cytoskeleton, generation of reactive oxygen species (ROS) and calcium-dependent signalling (Samaj et al., 2006). Regulators of Rho proteins, such as *RhoGDI*, are known to spatially control polar growth in pollen tubes (Klahre et al., 2006). Another Rho regulatory protein from tobacco pollen tubes is responsible for recycling RAC5 from the flanks to the apex of the growing tip (Klahre et al., 2006). RAC proteins such as RAC10 are involved in endocytosis (Bloch et al.,

2005) and therefore their recycling is important during pollen tube growth. Moreover, the ARF proteins have also been implicated to play a role in vesicle trafficking during root hair and pollen tube development (Song et al., 2006). Since all these genes are necessary for tip growth, promoter analysis of these genes will enable us to determine if *AGL104* regulated one or more of these genes.

The ion channels are also an integral part of the pollen tube elongation mechanism. Thus, ensuring synthesis of proteins of the ion channels itself or proteins required for the activation and proper functioning of these ion channels are also a possible role of *AGL104*. As mentioned above, the Rho/RAC GTPases generate ROS. These act as activators of calcium channels and thus trigger calcium signaling (Foreman et al., 2003). Activation and signal transduction during the opening and closing of ion channels usually involve a cascade. The transient and sequential expression of seven zinc-finger transcription factors at different stages of early pollen development (Kobayashi et al., 1998) had earlier led to the proposal that these transcription factors may constitute a regulatory cascade with each having specific target genes, including the next-acting transcription factor in the cascade. It is possible that the AGL104 protein is part of a similar cascade, especially considering the specific expression and enrichment of other members of the MIKC* clade transcription factors in pollen (Honys and Twell, 2004; Kofuji et al., 2003; Pina et al., 2005). As mentioned previously (Chapter 1 and 3), genes of the MIKC* type were exclusively expressed and enriched in the pollen transcriptome (Pina et al., 2005) and Kofuji et al. (2003) detected similar expression patterns for the four (*AGL30*, *AGL65*, *AGL66*, and *AGL104*) MIKC* type genes using Macroarrays and Northern Blotting. It remains to be determined if these other MIKC* proteins are also translocated to the pollen tip where they regulate the necessary genes. It would also be interesting to determine if the MIKC* genes act in a cascade-like manner to control processes associated with pollen germination and tube elongation. This is most probable since it was recently reported that one of the genes downregulated in a *agl66/104* double mutant was *MYB97* (Pina et al., 2005; Verelst et al., 2007). *MYB97* is a transcription factor and its regulation by MIKC* genes indicate the existence of a complex regulatory network in pollen.

Considering the importance of the above processes in tube elongation, it would be expected that the GTPase genes would be up-regulated. Since *AGL104* most probably does not activate genes during pollen germination and tube growth (Honys and Twell, 2004), it is likely that it acts as a repressor of genes that inhibit the synthesis or activity of these GTPases i.e. indirect

regulation. Verelst et al. (2007) reported the existence of genes possessing CArG boxes that were bound specifically by certain MIKC* complexes. However, these genes were not overrepresented in tri-cellular and mature pollen, implying that they may be indirect targets.

Alternatively, *AGL104* could directly regulate the GTPases. It was reported recently that genes with functions related to vesicle transport, cytoskeleton, cell wall and signal transduction were up-regulated following the appearance of MIKC* complexes during the tri-cellular stage of pollen development (Pina et al., 2005; Verelst et al., 2007). These up-regulated genes possessed MEF2 motifs that are bound by MIKC* complexes. Promoter analysis of the aforementioned GTPases would help determine if they are also directly regulated by MADS-box transcription factors and if yes, what is the mode of regulation i.e. are the GTPase genes activated or repressed? Overexpression of a constitutive active form of RAC10 resulted in aberrations of actin organization that correlated with inhibition of endocytosis in root hairs and thus, disruption in growth (Bloch et al., 2005). Direct repression of RAC10 would be required in order to facilitate tube elongation. *AGL104*, being a transcription factor, would be suitable for such a role.

Becker et al. (2003) reported that certain genes were specifically down-regulated in the pollen transcriptome compared to the vegetative tissues. These genes had putative functions in metabolism, protein biosynthesis, stress response and water channels (Becker et al., 2003). It is possible that *AGL104* is specifically involved in down-regulating such genes. Utilization of transcription factors for this purpose seems logical since the pollen has to conserve its resources and use it for the specific purposes of germination and tube elongation. In this scenario, however, *AGL104* would be involved in regulating genes that do not directly affect pollen germination or tube elongation but need to be regulated in order to allow germination and tube growth to occur.

Looking at the above arguments on pollen gemination and pollen tube elongation, it is clear that to accurately determine the effect of *AGL104* on these processes, further studies are required since gain-of-function analysis is not absolutely informative in itself. One of the disadvantages of gain-of-function analysis is that in some instances, it is not clear whether a phenotype reflects the true function of a gene or whether it is simply due to the interference of unrelated processes. This was demonstrated in the gene *ANTHOCYANINLESS2* (*ANL2*), where the overexpression phenotype did not correlate with its loss-of-function phenotype (Kubo et

al., 1999; Weigel et al., 2000). Generation of loss-of-function mutants are important as it will compliment the gain-of-function *in vitro* and *in vivo* germination results and help us determine if the above phenotypes observed for the overexpressing mutants (Table 4.1 and 4.2) are due to the true function of *AGL104* in pollen or due to other unknown factors. The phenotype of loss-of-function mutants would also enable us to determine whether *AGL104* is a repressor or activator. Chapter 5 discusses the phenotype of putative loss-of-function mutants generated in our study.

4.4 Chapter Summary

Due to the existence of a high degree of redundancy among MADS-box genes, knocking out *AGL104* would most probably have led to its redundant partners compensating for loss-of-function of *AGL104*. Therefore, we generated gain-of-function mutants in order to express *AGL104* outside of its domain of expression, i.e. overexpress *AGL104* in cells where it is not endogenously expressed. This technique has been used extensively to decipher functions of a number of genes (section 4.1). The concept behind this method is that cells overexpressing the gene of interest would most probably acquire the cellular fate specified by this gene, thus indicating its function or the cells would multiple the effect of this gene and result in a phenotype magnified several times compared to its effect in wt. The gain-of-function method can be designed so that the gene is either constitutively active, in which case it will be active at all times and in all tissues of the mutant or conditionally active i.e. active in certain tissues only. To fulfill our purpose, we opted for constitutively active gain-of-function mutants.

Since MADS-box genes act in a combinatorial manner and overexpressing genes outside of its normal domain results in a phenotype only when all the interacting partners are also present in the overexpressing cells, we generated multiple gain-of-function mutants using *AGL104*, *AGL66* and *AGL67* (section 4.1.1). *AGL66* and *AGL67* were chosen due to their close phylogenetic relationship to *AGL104* which increased the possibility that they associated in a combinatorial manner. Thus, to enable characterization of *AGL104* function, we hoped that overexpressing its putative combinatorial partners would result in a mutant phenotype representative of the function of *AGL104*. GUS expression analysis and various published results had established the expression of *AGL104* in mature pollen and also in the embryo sac (Chapter 3). Due to time constraints, we were able to analyze the phenotype of the

overexpressing pollen only and not the embryo sac. Since *AGL104* was expressed at very high levels in mature pollen, it was suggested that it played a role during pollen germination and pollen tube elongation. Thus, we germinated the mutant pollen *in vitro* and *in vivo* and looked at features such as pollen germination frequency and length of the pollen tubes and compared these features with that of wt pollen.

We examined the phenotype of structures such as leaves, stems, siliques, inflorescence patterning and floral organs to determine if overexpression of *AGL104* (and *AGL66* and *AGL67*) revealed a mutant phenotype. Consistent with *AGL104* being a gametophyte-specific gene (Chapter 3), no observable mutant phenotype was detected in the various overexpressing mutants (section 4.2.3.1) since the structures examined above were sporophytic in origin. We then scored the pollen germination frequency in the various mutant lines to determine if *AGL104* played a role during this process. There was a significant increase in the frequency of germinated pollen in all the overexpressing mutant lines compared to wt (section 4.2.3.2). The 35S::*AGL104* mutant exhibited the highest increase in pollen germination frequency while the double and triple mutants exhibited significantly higher germination compared to wt but the increase was not as high as that of the 35S::*AGL104* line. We had also generated a double 35S *AGL104* mutant line (35S::35S::*AGL104*) in order to overexpress *AGL104* at a higher level than the single 35S line (35S::*AGL104*). Due to time constraints, we were unable to determine the copy number or the level of expression of *AGL104* in these two mutant types. Therefore, we could not conclude if the mutant phenotype was dependent on the "dosage" of *AGL104*. However, it was apparent that *AGL104* promoted pollen germination since overexpressing this gene led to an increased frequency of germinated pollen. Since recent studies had established that pollen germination occurred independently of transcription, it most proposed that *AGL104* promoted pollen germination by acting as an activator of germination-promoting genes before the pollen was shed or as a repressor of germination-inhibiting gene during the germination process.

We had generated multiple mutants using *AGL66* and *AGL67* assuming that these genes acted in a combinatorial manner with *AGL104*. The triple overexpression mutant overexpressing all three genes did not exhibit a cumulative effect on the pollen germination frequency, thus indicating that these genes may not act in a combinatorial manner. Recent publications reported that *AGL104* formed functional heterodimers with other MIKC* type genes but not with *AGL66* and *AGL67* (section 4.3.2.1.1). These studies also showed that *AGL104* acted

redundantly with *AGL66* instead of in a combinatorial manner. This explained the slightly lower frequency of germination in the multiple overexpressing mutants compared to the single 35S::*AGL104* line. Competition between *AGL104* and *AGL66* to form functional heterodimers with the other factors was suggested to be the cause of the difference in germination frequency observed between the single 35S::*AGL104* line and the double 35S::*AGL66/104* mutant. Further experiments incorporating all the genes involved in forming germination promoting functional heterodimers were suggested.

In vivo germination results after 2 hrs of germination showed a general increase in the length of the pollen tube compared to wt. However, a clear pattern regarding the contributions of each of the genes in the multiple mutants was not seen. A repeat of the *in vivo* experiment was suggested since there were discrepancies in the total amount of time pollen was allowed to germinate on different pistils. In addition, there were differences in the number of pistils of each mutant line used to determine the average tube length.

Since in 2 hrs the pollen tubes had elongated to just below the stigma, it was likely that this growth represented the initial elongation, the raw materials for which is known to be present in the mature pollen. It was postulated that a difference in the tube length may become apparent after a longer period of germination once the stored raw materials were used up and the pollen initiated new protein synthesis. Thus, we germinated pollen for 4 hrs *in planta* using the single 35S::*AGL104* mutant line. A significant increase in the length of the pollen tube was observed after 4 hrs. Since pollen germination occurred independently of transcription, it was most likely that *AGL104* promoted pollen tube elongation by acting as a repressor of genes inhibiting tube elongation.

A number of different processes are known to act together to effect pollen tube growth. In addition, pollen transcriptome had reported up-regulation of genes with functions in signal transduction, cell wall biosynthesis, vesicle transport, ion dynamics, and cytoskeletal dynamic. It was likely that *AGL104* regulated genes involved in one or more of these processes. Promoter analysis of genes exhibiting alteration in expression level during the time of *AGL104* expression was suggested to determine the type of genes regulated by *AGL104*. In addition, a recent report had shown that complexes formed by MIKC* proteins both up-regulated and down-regulated genes during pollen germination. This was logical since pollen germination and tube elongation are results of a number of different processes acting together and therefore,

a complex regulatory network was likely to exist. Thus, it was proposed that perhaps *AGL104*, together with other genes, acted as both a repressor and activator of genes during pollen germination and tube elongation. Further studies, specifically the generation of loss-of-function mutants, were suggested to accurately elucidate the function of *AGL104*.

CHAPTER 5

LOSS-OF-FUNCTION ANALYSIS

5.1 Background

Historically, genes were functionally and genetically characterized through the "forward genetics" method whereby genes were first defined by the mutant phenotype and then isolated. Recently, the "reverse genetics" method has gained preference among scientists (Galbiati et al., 2000; Marsch-Martinez et al., 2002; Samson et al., 2002). In this method, mutants or disrupted sequences are first isolated and then their functions are characterized. One of the popular techniques used for reverse genetics is T-DNA- or transposon-insertional mutagenesis where a variety of tactics are used to generate large populations of tagged mutants, which are then screened for disruptions in specific genes (Bingham et al., 1981; Parinov et al., 1999; Speulman et al., 1999; Tissier et al., 1999). T-DNA insertion mutagenesis has been successful in elucidating functions of a variety of genes (Liljegren et al., 2000; Pelaz et al., 2000; Ratcliffe et al., 2000). However, as with all techniques, this method has its pros and cons. The disadvantages of the T-DNA insertion technique are that insertions in genes of interest may not be found and the alleles containing the insertion may not be a null allele, i.e. expression of the gene is not fully abolished.

To overcome these disadvantages, researchers developed directed techniques for reverse genetics. For example, RNA interference (RNAi) was designed when it was discovered that small RNAs could silence endogenous transcripts. This method eliminates gene expression without disrupting the gene itself (Bosher and Labouesse, 2000). In RNAi, the sense and antisense RNA strand of the gene of interest is expressed simultaneously, which are cleaved into short double-stranded RNA (dsRNA) pieces (Waterhouse et al., 1998). The antisense strands then bind to the specific endogenous mRNAs and result in either the cleavage of the endogenous RNA (through the action of the DICER enzyme in the RNA-Induced Silencing Complex (RISC)) or its translational suppression (Baulcombe, 2004). Even though RNAi has led to isolation of loss-of-function mutants of genes such as *AG*, *API* and *PERIPANTHIA*

(*PAN*) (Chuang and Meyerowitz, 2000), it was discovered that it did not always abolish gene expression completely.

We, therefore, used both the *AGL104* T-DNA insertion line available from the *Arabidopsis* Biological Resource Center (ABRC) and generated *AGL104* dsRNAi lines in the hope of obtaining an observable loss-of-function phenotype of *AGL104*. Our results are presented below.

5.2 Results

5.2.1 Loss-of-function mutants of *AGL104*

5.2.1.1 Generating AGL104 dsRNAi lines

Figure 5.1 demonstrates the strategy used for making the *AGL104* dsRNAi construct. The IKC region of *AGL104* (*AGL104* gene without the MADS-box region) was cloned using PCR and subcloned into the pRNA69 vector (kind gift from J.L Bowman) in the sense direction. The full length fragment of *AGL104* was then cloned into the same vector (already containing the IKC region) in the antisense direction. The segment containing the IKC and full length *AGL104* fragments was then subcloned into the binary vector, pART27. The binary vector enables transformation of the construct into *Agrobacterium* (see section 2.6.4 for details on generating the construct). During the generation of the above construct, the incorporation of the *IKC* and *AGL104* fragment in the correct direction in each vector was verified repeatedly using restriction digests. The restriction enzymes used and the fragments obtained through each restriction digest are presented in Figure 5.1.

Arabidopsis wt plants were transformed with *Agrobacterium* containing the dsRNAi construct. The seeds collected from the putative transformed plants were selected on antibiotic plates to obtain resistant/mutant seedlings. These seedlings were then transplanted into soil and allowed to self-fertilize. Approximately twenty independent lines were generated out of which we used five lines of the T_2 generation for detailed phenotypic analyses. We designated these lines *agl104-2*, *agl104-3*, *agl104-4*, *agl104-5* and *agl104-6*.

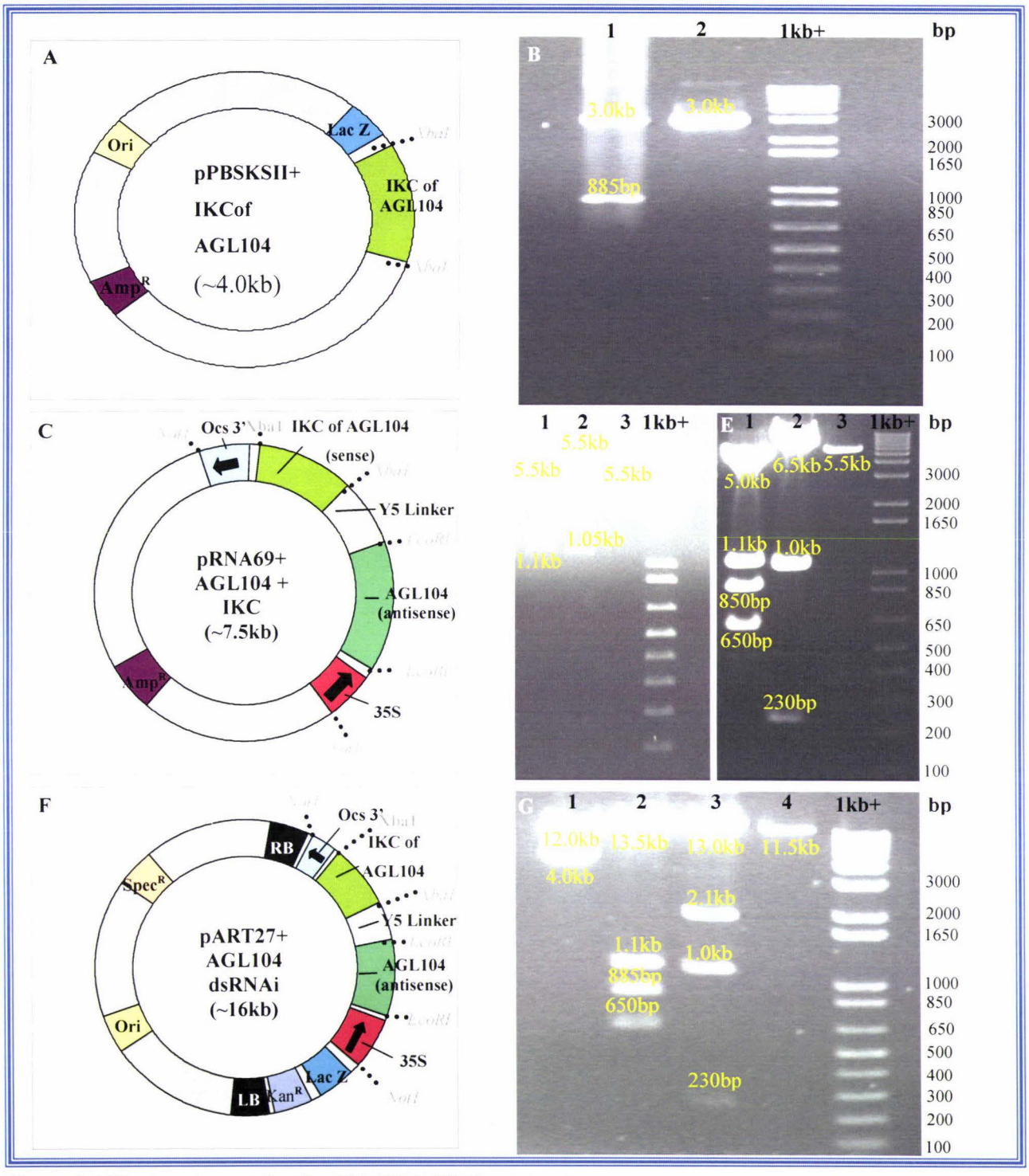

Figure 5.1 Cloning strategy for the *AGL104* dsRNAi construct.

A,C,F: plasmid maps showing the sense and antisense strand of *AGL104* in various vectors. The endonuclease restriction sites shown (grey prints) were used to check the construct or excise and ligate the sense and antisense *AGL104* fragments into these vectors.

B, D, E, G: restriction digests confirming the construction of A, C and F, respectively. Each lane represents the respective construct digested with various enzymes. Fragment sizes are marked. Enzymes used for digestion is listed below. **B:** lane 1- XbaI; lane 2- pBSKSII (control). **D:** lane 1- EcoRI; lane 2- BamHI; lane 3- pRNA69 (control). **E:** lane 1- EcoRI and XbaI; lane 2- BamHI; lane 3- pRNA69 (control). **G:** lane 1- NotI; lane 2-EcoRI and XbaI; lane 3- BamHI; lane 4- pART27 (control).

1kb+- DNA ladder; **Amp^R**- ampicillin resistance; **bp**- base pairs; **35S**- constitutive promoter from Cauliflower mosaic virus; **Kan^R**- kanamycin resistance; **kb**- kilo bases; **Lac Z**-gene used for blue-white selection of transformed *E.coli*; **LB**- left border; **Ori**- origin of replication; **RB**- right border; **Spec^R**- spectinomycin resistance.

5.2.1.2 Obtaining AGL104 T-DNA insertion homozygous lines

The *AGL104* T-DNA insertion line (SALK_013898) was obtained from the ABRC. The plants were allowed to self-fertilize until homozygous individuals (designated *agl104-1/agl104-1*) were obtained (Figure 5.2B). All *AGL104* T-DNA insertion individuals selected for phenotypic analysis had the same genotype. As shown in Figure 5.2A, the T-DNA insertion is in the 9th exon on the 3` end of *AGL104*.

5.2.2 Phenotypic Analysis

5.2.2.1 Gross morphological analysis of the loss-of-function lines

The homozygous *AGL104* T-DNA line (*agl104-1/agl104-1*) and the five *AGL104* dsRNAi lines (*agl104-2*, *agl104-3*, *agl104-4*, *agl104-5* and *agl104-6*) of the T_2 generation were phenotypically analyzed. Gross morphological analyses of these putative loss-of-function lines showed that there were no obvious phenotypic variation between these mutants and the wt (Figure 5.3). Panels A to C show whole plant, inflorescence and a flower of a wt plant; panels D to F show the respective structures of a representative *AGL104* dsRNAi line while panels G to I show the structures of the *AGL104* T-DNA line used for analysis. In the whole plants, the phenotype of sporophytic structures such as leaves, stems, siliques and inflorescences of the mutants were compared to that of wt plants. Features such as shape, size, and number of organs (for inflorescence) were examined. As shown in Figure 5.3 panels A, D and G, there were no differences in the phenotype of these structures between wt and the mutants. In addition, there were no differences in the phyllotaxy or spacing of flowers in the inflorescence between the wt and the mutants (panels B, E, H). The structure of the individual flowers also did not exhibit any differences in phenotype (panels C, F, I) i.e. the number of sepals, petals, stamens and carpel and their shape and size in the mutants and wt flowers were the same. In addition, it was noted that the putative loss-of-function *AGL104* T-DNA individual and *AGL104* dsRNAi lines were fertile since seeds appeared normal and germinated well and there was no indication of decreased fertility.

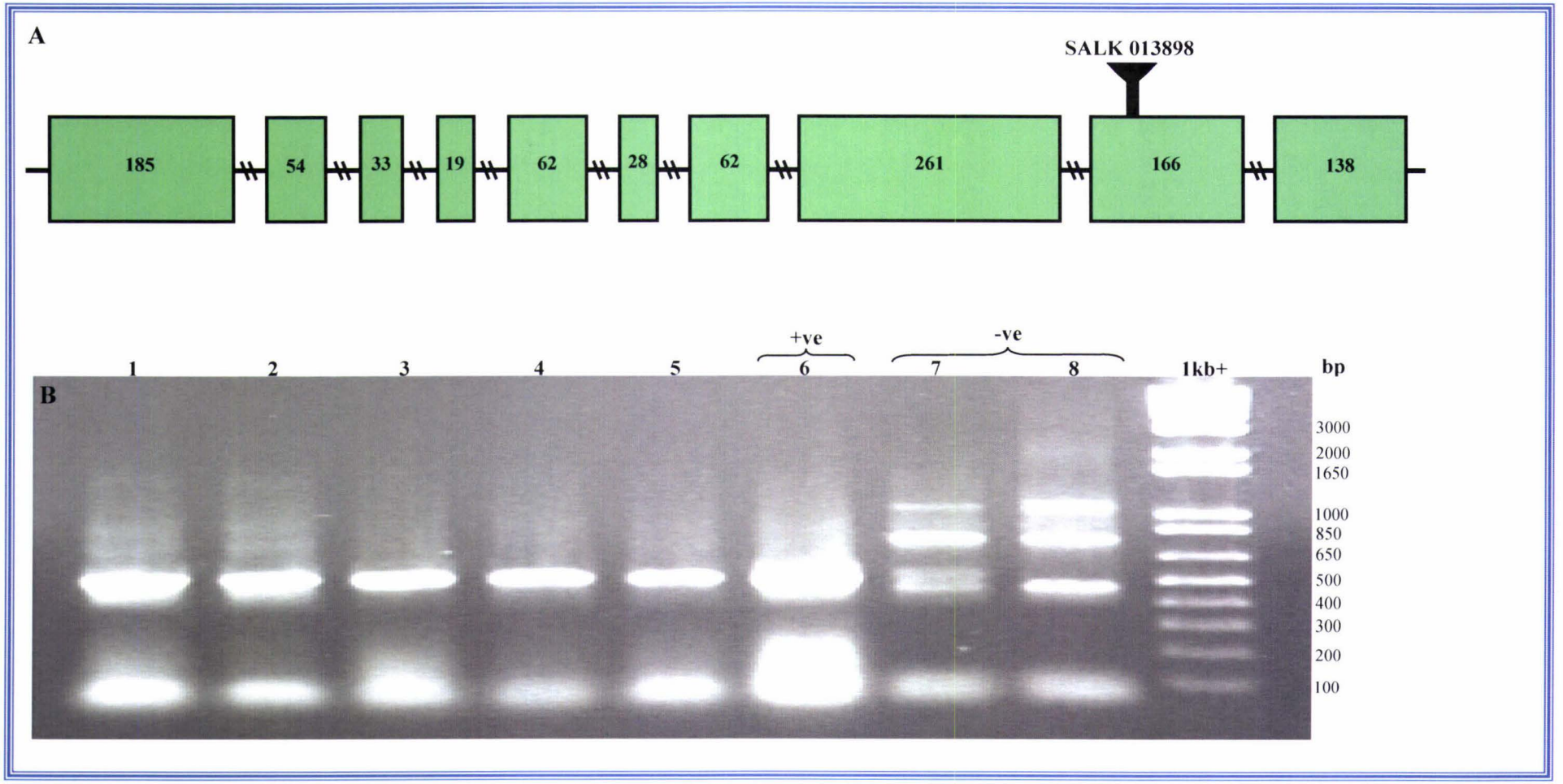

Figure 5.2 Diagram showing T-DNA insertion in *AGL104* and RT-PCR confirming generation of homozygous lines.

A: *AGL104* with its 10 exons represented by the green boxes. A T-DNA insertion, SALK 013898, is found on the 3` end of the gene. **B:** RT-PCR confirming the generation of homozygous individuals. Lanes 1-5: *AGL104* T-DNA homozygous individuals; lane 6- positive control, known homozygous individual; lane 7 and 8- negative control, known heterozygous individuals.

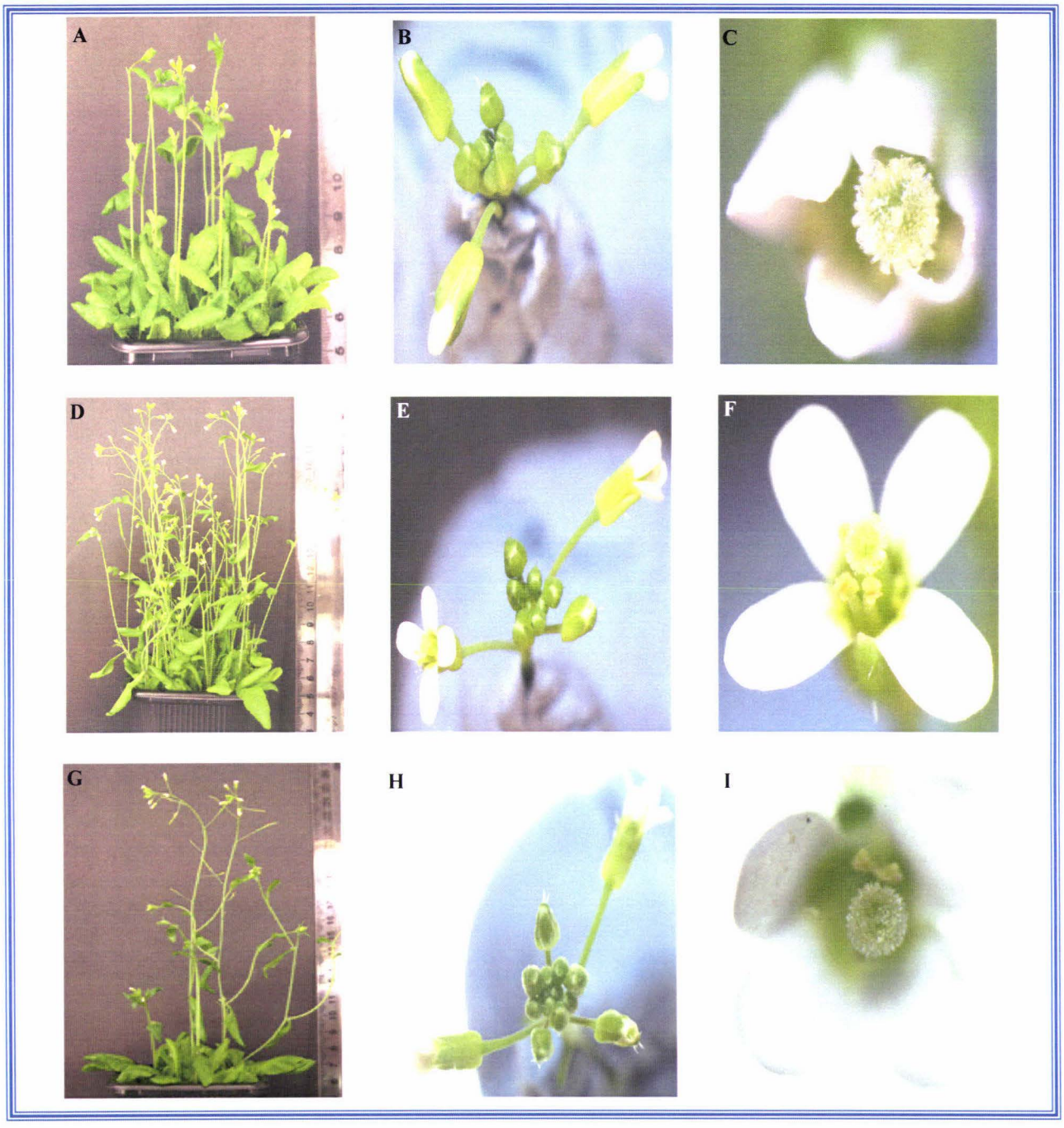

Figure 5.3 Gross morphological phenotype of the putative loss-of-function mutants

A-C- wt; **D-F**- *AGL104* dsRNAi line; **G-I**- *AGL104* T-DNA insertion line.

Gross morphological phenotype (whole plants, inflorescence and individual flowers) of the dsRNAi and T-DNAmutants were examined. There were no observable changes in the morphology of the mutants compared to the wildtype. In addition, all mutant lines were fertile.

5.2.2.2 In vitro pollen germination

As shown in Table 4.1, all the putative loss-of-function *AGL104* T-DNA (*agl104-1/agl104-1*) and the five *AGL104* dsRNAi lines (*agl104-2*, *agl104-3*, *agl104-4*, *agl104-5* and *agl104-6*) showed a comparable frequency of pollen germination to that of wt (section 2.9.1 describes the method used to record this data). The pollen germination frequency of wt pollen was 38.5% whereas the frequencies of the *AGL104* T-DNA and *AGL104* dsRNAi lines were 37.8% and 38.2%, respectively. In addition, the morphology of the germinated pollen from the above two mutants was compared with germinated wt pollen (Figure 5.4). We examined features such as the number of pollen tubes germinating out of each pollen grain and the difference in diameter (width) of the pollen tube. No difference in the phenotype of the elongating pollen tube between the wt and the *AGL104* T-DNA and *AGL104* dsRNAi pollen was observed.

5.2.2.3 In vivo pollen germination

Pollen was germinated *in vivo* to determine differences (if any) in the length of the pollen tube between the wt and the *AGL104* T-DNA (*agl104-1/agl104-1*) and *AGL104* dsRNAi lines (*agl104-2*, *agl104-3*, *agl104-4*, *agl104-5* and *agl104-6*). We germinated the pollen for 2 hrs and measured the length of the pollen tubes (see section 2.9.2 for *in vivo* method and measuring). Germination of 2 hrs allowed clear visualization of the pollen tubes as the growing tubes reached the style after 2 hrs (Figure 5.5). As shown in Table 4.2, these putative loss-of-function lines showed an increase in the length of the pollen tubes compared to wt. The *AGL104* dsRNAi lines showed an increase of 28.8% in the length of its pollen tube compared to wt while the *AGL104* T-DNA insertion individuals exhibited a slight (not significant) increase of 5.8%.

5.3 Discussion

The purpose of generating *AGL104* loss-of-function lines was to determine the function of this gene by examining the effects of its non-expression on the phenotype of *Arabidopsis*. We had intended to check the expression level of *AGL104* in the *AGL104* T-DNA homozygous individual using RT-PCR (data not shown). However, despite several attempts, we were unsuccessful in achieving a conclusive result in time for the thesis writing since we were

unable to equalize the amount of cDNA used in the PCR, which resulted in bands of different intensity for control samples. Northern Blots are currently being performed by other members of the laboratory to conclusively determine whether *AGL104* expression level is decreased, abolished or unchanged in the *AGL104* T-DNA and the *AGL104* dsRNAi lines. We opted to phenotypically analyze these putative loss-of-function lines while awaiting confirmation of disruption or non-disruption of *AGL104* expression.

5.3.1 *AGL104* activity in the sporophytic tissues/whole plants

Analyses of the gross morphology of the putative loss-of-function *AGL104* T-DNA and *AGL104* dsRNAi lines and wt did not reveal any differences in the sporophyte generation (Figure 5.3). Since we were unable to establish that *AGL104* expression was knocked out in these two mutants, the lack of difference between the wt and mutant plants could have been due to the expression of *AGL104* not being abolished. However, the phenotype observed in the putative loss-of-function lines could also have been due to *AGL104* not being a sporophytic gene (as discussed in Chapter 4) and thus not having an effect in the sporophytic structures even when its expression is knocked out. This was confirmed by the results of Verelst et al. (2007), who also did not observe a mutant phenotype in the sporophytic tissues of their *AGL104* loss-of-function line. This group used an *AGL104* T-DNA insertion line that led to a strong reduction in the expression level of *AGL104*. Therefore, we conclude that *AGL104* does not affect the morphology of the sporophytic features such as shape and sizes of leaves, stems, inflorescences and flowers of the *Arabidopsis* plants. This is consistent with *AGL104* being a gametophytic-specific gene (Chapter 3) (expressed in the embryo sac of ovules and pollen) and thus does not function in the sporophyte generation.

In addition, the putative loss-of-function *AGL104* T-DNA and *AGL104* dsRNAi mutant plants exhibited normal fertility and seed set was equivalent to wt. Based on the expression of *AGL104* in the unicellular microspore, the bi- and tri-cellular pollen, we had proposed (section 3.2.2.2) that this gene was involved in the development of pollen. If the expression level of *AGL104* is shown (future Northern Blots) to be significantly decreased or abolished in the *AGL104* T-DNA and *AGL104* dsRNAi lines, our current set of results would indicate that knocking out the expression of *AGL104* did not alter pollen viability in these loss-of-function lines, thereby implying that *AGL104* is not involved in pollen development but accumulates in pollen to regulate late occurring processes such as pollen germination and tube elongation

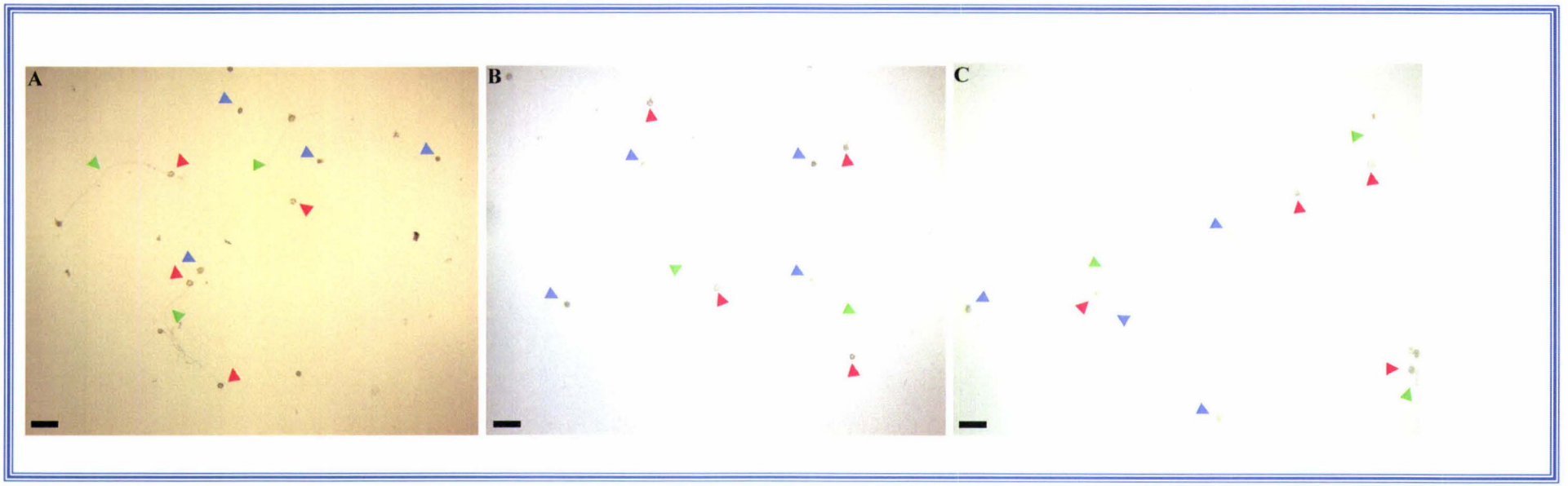

Figure 5.4: Phenotype of *in vitro* germinated pollen of putative loss-of-function lines.

A: wt; **B:** *AGL104* dsRNAi; **C:** *AGL104* T-DNA.

Phenotype of pollen germinated *in vitro* for 24 hrs (10X). There were no observable changes in the phenotype such as width of the pollen tube or the number of pollen tubes germinating out of each pollen grain. Pollen with visible protrusions or an elongated tube were considered "germinated". A few of the germinated and ungerminated pollen and elongated pollen tubes are marked.

red arrow head- germinated pollen; **blue arrow head**- ungerminated pollen; **green arrow head**- fully elongated pollen tubes.

Scale bar= 100µm

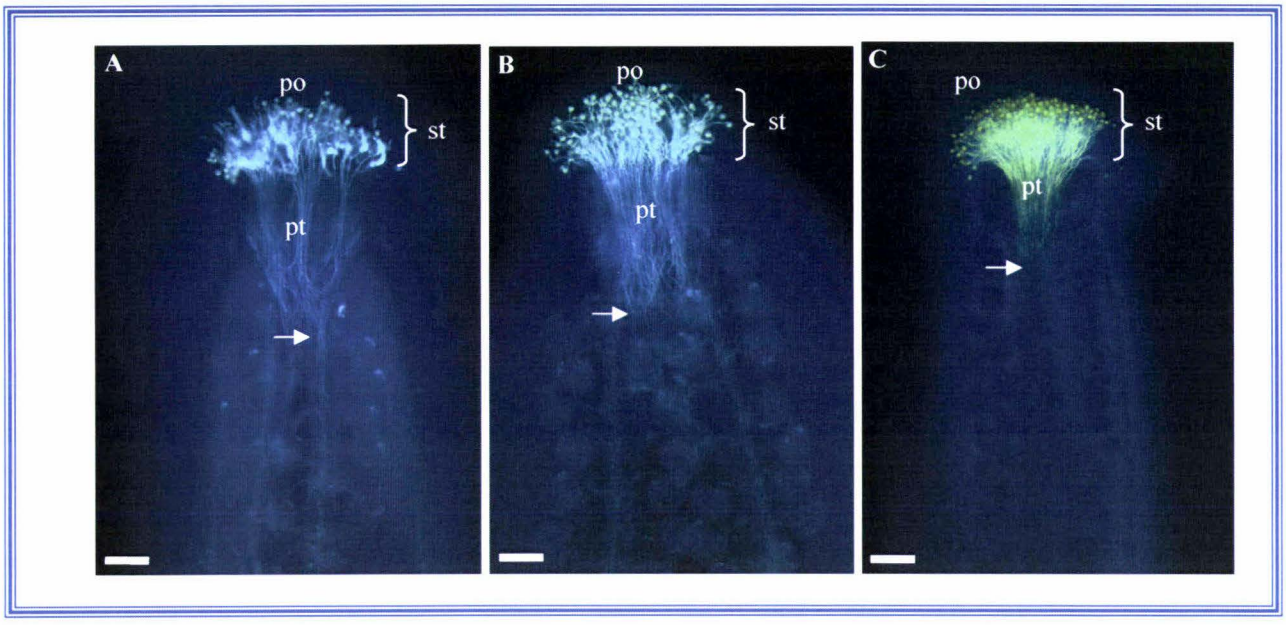

Figure 5.5 Pollen of putative loss-of-function lines germinated *in planta* for 2hrs

A: wildtype **B:** *AGL104* dsRNAi; **C:** *AGL104* T-DNA. Pollen tubes are seen germinating down the style.

Flowers of respective mutants were emasculated and hand pollinated the following day using pollen from the same plant. Pollen was allowed to germinate on the plant. After 2 hrs, the flowers were picked, fixed and cleared and stained with Analine Blue for 3-5hrs in the dark at rt. Flowers were then examined under the fluorescence microscope (10X) using the DAPI filter. Pollen tubes that had germinated along the style were examined and the length of the longest visible tube was recorded. Multiple flowers were examined for each mutant and an average length of the pollen tubes was calculated.

po- pollen grain; **pt**- pollen tubes; **st**- stigma; **white arrow**- marks the end of the longest visible pollen tube

Scale bar = 100μm

only. The results obtained from loss-of-function analysis carried by Verelst et al. (2007), which was published **during the writing of this thesis**, adds credibility to the above scenario. This group observed that knocking out the expression of *AGL104* and *AGL66* in the double mutant, *agl66/104*, led to a virtual inability of the *agl66*/104 pollen to germinate *in vitro*. Thus, it was expected that this double mutant would be male sterile. However the *agl66/104* mutant plants did not exhibit impaired fertility *in vivo* and therefore it was concluded that *AGL104* and *AGL66* did not play a role in determining pollen fertility (Verelst et al., 2007). They also scored pollen viability in the various loss-of-function mutants of the MIKC* type genes and even though one or more of the genes were knocked out, pollen viability was not affected. It was suggested that the *agl66/104* mutant phenotype (and phenotypes of other gene mutants) might be more subtle *in vivo*, perhaps affecting speed and efficiency of pollen germination and tube growth (Verelst et al., 2007).

5.3.2 *AGL104* activity during pollen germination

Scoring the frequency of germinated pollen showed a lack of difference between the wt and the putative loss-of-function *AGL104* T-DNA and *AGL104* dsRNAi mutants *in vitro* (Table 4.1). The overexpression of *AGL104* had resulted in an increased frequency of pollen germination (Table 4.1), thereby establishing that *AGL104* played a role in pollen germination. Thus, we had expected an opposite effect, i.e. a decrease in the frequency of pollen germination, in the putative loss-of-function lines. The comparable frequencies of pollen germination in the wt, *AGL104* T-DNA and *AGL104* dsRNAi lines (Table 4.1), therefore, could be due to:

- *AGL104* expression not being knocked out
- *AGL104* being redundant with other genes

Considering the *AGL104* T-DNA line used and the effect of a dsRNAi construct in mature pollen, there is a high probability that the expression was not knocked out in these loss-of-function lines. This is due to the fact that in the *AGL104* T-DNA homozygous line, the transposon (SALK 013898) is inserted in the 3` region ($9^{th}$ exon). Depending on the importance of this 3` region in the regulation of *AGL104*, there is a possibility that *AGL104* expression will not be completely knocked out, thus not affecting pollen germination frequency (as shown in Table 4.1). Therefore, (future) Northern Blot results showing endogenous levels of *AGL104* expression in the T-DNA mutant lines is not unpredicted. In the

event the (future) Northern Blot results revealed a decrease in the level of expression of *AGL104* in the T-DNA mutant line compared to wt, it would imply that the 3` exon region where the transposon was inserted is important in the regulation of *AGL104* expression.

Taking into account the dsRNAi lines used, expression analyses of fifteen genes involved in the small RNA pathway in *Arabidopsis* showed that the transcripts of these genes, with the exception of ARGONAUTE 1, were absent from the mature pollen transcriptome (Pina et al., 2005). This implied that small RNA pathways that are used by the RNAi method to "operate" the dsRNAi construct in order to knock out gene expression may be inactivated in the mature pollen grains, thus resulting in the expression level of *AGL104* not being decreased in pollen as expected. This scenario would result in pollen germination frequency not being affected (as shown in Table 4.1). (It is noted, however, that the downregulation of the small RNA pathway genes were reported in mature pollen only (Pina et al., 2005). Therefore, the dsRNAi construct would be functional in other types of cells (leaves, stems, siliques etc) in the plant and lead to a decreased expression of *AGL104*).

On the other hand, the implication of redundancy between *AGL104* and other MADS-box genes is not surprising. As discussed previously (section 4.1.1), we suspected that *AGL66* and *AGL67* acted redundantly with *AGL104*. However, we could not generate mutants with multiple gene knock outs due to time constraints and thus could not draw a conclusion regarding redundancy of *AGL104* with other genes using our current data set. Nevertheless, the results of Verelst et al. (2007) published during the writing of this thesis established that *AGL104* was involved in pollen germination and most probably acted redundantly with *AGL66*. A strong reduction of *AGL104* expression in the single *agl104* mutant line and the complete loss of function of *AGL66* in the single *agl66* line did not affect pollen germination when these genes were knocked out alone. However, pollen of the double *agl66/104* mutant was virtually unable to germinate *in vitro*, thus implying redundancy between *AGL66* and *AGL104*.

5.3.3 *AGL104* activity during pollen tube elongation

As indicated in Table 4.2, the length of the pollen tubes in the putative loss-of-function *AGL104* T-DNA and *AGL104* dsRNAi mutants was longer compared to the wt. The *AGL104* dsRNAi mutant lines showed a significant increase of 28.8% while the *AGL104* T-DNA

insertion line showed an increase of 5.8%. Assuming that *AGL104* expression in the T-DNA and the dsRNAi mutants was knocked out, this difference in the length of the pollen tube between these lines could have been due to different degrees of knock outs i.e. different levels of *AGL104* expression. It would be logical to assume that the insertion in the 3` region of the *AGL104* T-DNA mutant would not result in complete abolishment of *AGL104* expression and therefore some amount of *AGL104* would be expressed, leading to only a slight difference in the length of *AGL104* T-DNA pollen tubes compared to wt. On the other hand, if the *AGL104* dsRNAi construct was able to function normally (i.e. small RNA pathway was active in the mature pollen) it was likely to lead to a greater degree of decrease in the expression level of *AGL104* compared to the *AGL104* T-DNA individuals.

The *AGL104* overexpression lines (Figure 4.4), led to an increase in the length of pollen tubes (except in the 35S::*AGL66/104* line) compared to wt. This result implied that *AGL104* acted as an **activator** of genes **promoting tube elongation**. Considering this phenotype and assuming that *AGL104* expression is knocked out in the putative loss-of-function mutants, the *AGL104* T-DNA and *AGL104* dsRNAi lines were therefore expected to yield an opposite effect, i.e. a decrease in pollen tube length. However, we observe a much more significant increase in pollen tube lengths in the *AGL104* dsRNAi lines (and a comparable increase in pollen tube lengths of *AGL104* T-DNA) compared to the overexpression lines (Table 4.2), implying that *AGL104* acted as a **repressor** of genes **promoting tube elongation**. In this scenario, knocking out the expression of the repressor would thus lead to a higher degree of elongation of pollen tubes as shown in Table 4.2.

This difference in the expected phenotype in the overexpression and putative loss-of-function lines can be explained if we assume that *AGL104* acts as both a repressor and an activator during pollen tube elongation. Reliable support for this assumption was provided by Verelst et al. (2007) who quantified the expression level of certain downstream genes in the double *agl66/104* mutant. Their results showed that certain genes were up-regulated while some were down-regulated. All these genes examined were putative targets of *AGL104* and *AGL66* as predicted by their *in silico* technique. It would be informative to examine the expression level of these target genes in our overexpression and putative loss-of-function lines to determine how these genes are affected.

As described earlier (section 4.3.2.3), pollen germination and tube elongation is a result of a number of processes acting together. The expression and suppression of *AGL104* in the overexpression and putative loss-of-function mutants, respectively, may have varying effects on pollen germination and tube elongation and therefore, the phenotypic observations recorded in Table 4.1, Table 4.2 and the results of Verelst et al. (2007) may represent a collective end result of various pathways active during pollen germination and tube elongation. This contrasting roles of *AGL104* (and *AGL66*), i.e. its role as an activator as well as a repressor (Verelst et al., 2007), is not surprising considering the number of different protein complexes formed between all the MIKC* type genes expressed in pollen. The co-expression of five of the six MIKC* genes during the late stages of pollen development led to the proposal that they might constitute a transcription factor network in mature pollen where the different protein complexes acted together to regulate pollen germination and tube elongation (Verelst et al., 2007). As suggested earlier (section 4.3.2.2), it is necessary to regulate pollen tube growth since after a period of rapid growth, as soon as a pollen tube reaches an ovule, it is mandatory to cease growth in order to release the sperm cells (Huck et al., 2003). It is possible that the inhibition of further growth of pollen tubes is through a feed-back mechanism involving the different complexes formed by *AGL104* and *AGL66* and other MIKC* genes expressed in pollen (Verelst et al., 2007).

Alternatively, it is noted that we measured the tube lengths of only 11 pistils of the 5 different *AGL104* dsRNAi lines and averaged the lengths. It is possible that the significant increase in tube length of the *AGL104* dsRNAi lines examined may have been due to positional effects instead of solely due to knock out of *AGL104* expression. A repeat of the experiment using more *AGL104* dsRNAi mutant lines and a larger number of pistils from each line would yield a clearer picture. Since the number of pistils used to determine the average tube lengths in the different mutants ranged from 4 to 27 (Table 4.2), it may be rational to repeat the experiment using more pistils of all the mutant lines. As mentioned in Chapter 3, *AGL104* may be involved in pollen tube guidance due to its expression in the embryo sac. Due to time constraints, we did not analyze the phenotype of the female gametophyte (embryo sac) in the gain-of-function (Chapter 4) or in above putative loss-of-function mutants. The phenotype of the embryo sac in the loss-of-function (and gain-of-function) mutants may affect the rate of pollen tube elongation. Therefore, to elucidate the function of *AGL104* in pollen tube elongation, the effect of knocking out *AGL104* expression on all the possible contributors (including embryo sac) needs to be examined. In addition, further analysis of germination and tube elongation of

pollen from gain-of-function and loss-of-function mutants overexpressing or lacking, respectively, the genes that form functional heterodimer complexes (Verelst et al., 2007) will enable us to broaden our understanding of how *AGL104* and other MIKC* type genes regulate these processes.

5.4 Chapter Summary

Morphological analyses of the sporophytic tissues of the putative loss-of-function *AGL104* T-DNA and *AGL104* dsRNAi mutants did not reveal any mutant phenotype. Since, due to time constraints, we were not able to verify whether *AGL104* expression was knocked out in *AGL104* T-DNA individuals and *AGL104* dsRNAi lines before the writing of the thesis, the lack of a sporophytic mutant phenotype may have been due *AGL104* expression not being decreased. However, considering that overexpression of *AGL104*, *AGL66* and *AGL67* (Chapter 4) also did not yield a mutant phenotype in the sporophytic tissues, it was concluded that *AGL104* does not play a role in the sporophyte. This was not surprising since *AGL104* expression is gametophyte-specific in angiosperms (Chapter 3) and thus it was not envisioned to play a role in leaves, stems, siliques, floral organs or processes such as phyllotaxy and inflorescence spacing, all of which are sporophytic in origin.

Considering the frequency of pollen germination in the putative loss-of-function *AGL104* T-DNA individuals and *AGL104* dsRNAi lines, it appeared that pollen from these mutant lines germinated at the same frequency as that of wt i.e. the (assumed) disruption in the expression level of *AGL104* in the *AGL104* T-DNA and *AGL104* dsRNAi lines did not affect pollen germination frequency. Apart from a lack of change in the *AGL104* expression level, the lack of an observable phenotype in pollen germination was attributed to the *AGL104* T-DNA insertion being in a region which did not regulate *AGL104* expression, thus resulting in a non loss-of-function allele. The lack of a change in pollen germination frequency in the putative *AGL104* dsRNAi lines was assigned to disruptions in the small RNA pathway in mature pollen, which led to the *AGL104* dsRNAi construct not being "activated" and thus deemed ineffective in knocking out the expression of *AGL104* in these lines. The importance of Northern Blots to confirm changes (or lack of it) in the putative loss-of-function *AGL104* T-DNA and *AGL104* dsRNAi lines in order to reach accurate conclusions regarding the observed phenotype was emphasized.

Pollen germination frequency of the *AGL104* T-DNA and *AGL104* dsRNAi lines being comparable to wt also implied existence of redundant genes as we had suspected (Chapter 1 and 4). However, since we did not generate multiple loss-of-function mutants in our study, we could not draw conclusion regarding redundancy of *AGL104* with other MADS-box genes. Nevertheless, the results of Verelst et al. (2007) had recently unambiguously established redundancy between *AGL104* and *AGL66* during the process of pollen germination.

Considering pollen tube elongation in the putative loss-of-function *AGL104* T-DNA and *AGL104* dsRNAi lines, a significant increase compared to wt was observed for the *AGL104* dsRNAi lines while a slight increase was observed for the *AGL104* T-DNA individuals. Assuming that these were indeed loss-of-function lines, the difference in tube length was attributed to a likely difference in the ability of these two mutants to knock out *AGL104* expression. Since the T-DNA insertion was in the 3` exon, which do not affect gene regulation dramatically, this mutant was not likely to result in a significant change (if any) in *AGL104* expression. In other words, *AGL104* would be expected to be expressed in small quantities in the *AGL104* T-DNA individuals even if the insertion resulted in a change in expression level, thus accounting for a smaller degree of increase in the pollen tube length compared to the *AGL104* dsRNAi lines. The *AGL104* dsRNAi line, on the other hand, had the capability of strongly decreasing *AGL104* expression upon activation of the dsRNAi construct.
Since both the overexpression mutants (Chapter 4) and the putative loss-of-function mutants described above resulted in an increase in tube lengths, it was hypothesized that *AGL104* acted as both an activator and repressor of genes involved in tube elongation. Given that pollen tube elongation utilizes a large variety of different processes to effect tube growth (Chapter 4), it was logical to assume that *AGL104* acted to activate some genes and repress others. Recent studies confirmed both activation and repression roles of MIKC* type genes (including *AGL104*) during pollen germination (Verelst et al., 2007).

Finally, a repeat of the *in vivo* germination experiment was recommended to obtain a consistent result since only a few pistils were used in each of the gain-of-function and loss-of-function mutant lines. In addition, due to its involvement in pollen tube guidance (and perhaps an influence in pollen tube elongation) and the observed expression of *AGL104* in the embryo sac (specifically the egg apparatus), a closer scrutiny of the phenotype of the embryo sac in the putative loss-of-function mutants was suggested in order to obtain an unambiguous data on the influence of *AGL104* on tube elongation. In addition, since *AGL104* forms various complexes

with other MIKC* genes (Verelst et al., 2007), generation of gain-of-function and loss-of-function mutants, which overexpress or knockout the genes involved in each of the different complexes, was suggested in order to fully elucidate the role of *AGL104* in the *Arabidopsis* gametophyte.

CHAPTER 6

GENERAL DISCUSSION AND FUTURE DIRECTIONS

The purpose of our project was to functionally characterize *AGL104*, a MIKC* type MADS-box transcription factor. The expression of the MIKC* genes in the gametophyte generation, which is the sexual/reproductive generation, of both mosses (Henschel et al., 2002; Krogan and Ashton, 2000) and angiosperms ((Honys and Twell, 2004; Kofuji et al., 2003; Pina et al., 2005; Verelst et al., 2007) has raised questions whether the developmental program of the reproductive structures in both these taxa are fundamentally similar even though the structures themselves differ greatly in their phenotype. In addition, the exclusive expression of MIKC* type genes, including *AGL104*, in the gametophyte generation of angiosperms (Honys and Twell, 2004; Kofuji et al., 2003; Pina et al., 2005; Verelst et al., 2007) provided a means to fathom the development of its gametophyte, an area of scientific interest still lacking crucial details in the current literature. Since late pollen development is thought to occur independently of transcription, the predominant expression of *AGL104* and other MIKC* type genes, which are transcription factors, in pollen (Honys and Twell, 2004; Kofuji et al., 2003; Pina et al., 2005; Verelst et al., 2007) raised questions regarding the dynamics of pollen development, especially the processes of pollen germination and pollen tube elongation. Thus, functional characterization of *AGL104* was expected to shed light on its role in the development of the male and female gametophyte in angiosperms and also set the groundwork for comparison of their evolutionary roles in the basal and higher plants.

We proceeded to characterize *AGL104* by using a combination of expression analysis and phenotypic analysis. We utilized the proven method of GUS assay for determining the expression of *AGL104*. For phenotypic analysis, we opted to first employ the gain-of-function (overexpression) method. We generated *AGL104* overexpression mutants in order to observe the effects of overexpressing *AGL104* outside of its normal domain of expression in *Arabidopsis*. We also generated multiple overexpression mutants using *AGL104*, *AGL66* and *AGL67* based on the likelihood that these genes acted in a combinatorial manner due to their close phylogenetic relationship. Despite its success, one of the disadvantages of overexpression analysis is that, in some instances, it is not clear whether a phenotype reflects the true function

of a gene or whether it is simply due to the interference of unrelated processes. This was demonstrated in the gene *ANTHOCYANINLESS2* (*ANL2*), where the overexpression phenotype did not correlate with its loss-of-function phenotype (Kubo et al., 1999; Weigel et al., 2000). Hence, we generated *AGL104* loss-of-function mutants with the expectation that the phenotype of these mutants would compliment the phenotype of the overexpression mutants. However, it was noted that one of the drawbacks of loss-of-function analysis is that it is difficult to functionally characterize genes that act redundantly with other genes. Examples of such genes are *SEP1/2/3/4* and *SHP1/2*, where knockouts in individual genes did not result in observable phenotypes (Pelaz et al., 2000; Ditta et al., 2004; Gu et al., 1998; Liljegren et al., 2000). Thus, by using a combination of gain-of-function and loss-of-function analyses, we hoped to obtain data that were complementary to each other and thus enable us to accurately interpret the phenotypic data to determine the likely role of *AGL104* in *Arabidopsis*.

GUS expression analysis revealed *AGL104* expression in the embryo sac, developing anthers and mature pollen. *AGL104* was seen to be first expressed weakly at the uninucleate stage and increase in expression as pollen development progressed, with the highest level of expression in the mature pollen. Our results are the **first** known data showing **spatial** expression of *AGL104* in the different developmental stages of pollen even though temporal expression in these stages was reported (Honys and Twell, 2004). In addition, our results revealed that *AGL104* expression was maintained in germinating pollen and in the elongating pollen tubes. This information was also **novel** since, to date, no research group had reported spatial expression of *AGL104* in the pollen tubes. Similar to suggestions by various published literature (Honys and Twell, 2004; Kofuji et al., 2003; Pina et al., 2005; Verelst et al., 2007), our data implied active regulation of genes by *AGL104* during the processes of pollen germination and tube elongation.

Our gain-of-function analysis results showed that *AGL104* promoted pollen germination. We could not reach a conclusion on the influence of *AGL104* on pollen tube elongation after 2 hours of growth since the data set did not reveal a significant pattern. However, there appeared to be a significant increase in the length of pollen tubes of the *AGL104* gain-of-function mutants when measured 4 hours after pollination. Loss-of-function mutants were expected to compliment the gain-of-function results. Even though we had generated loss-of-function *AGL104* T-DNA and *AGL104* dsRNAi lines, due to difficulties in optimizing the Northern blot conditions, we were unable to verify whether the expression level of *AGL104* was decreased,

diminished or unchanged in these lines. However, looking at the results of a recently published study on the effect of MIKC* genes on pollen germination, it appears that loss of *AGL104* function leads to a drastic decrease in pollen germination (Honys and Twell, 2004; Kofuji et al., 2003; Pina et al., 2005; Verelst et al., 2007). *AGL104* was also expressed in the egg apparatus (contains the egg cell and synergid cells) of the embryo sac. This expression pattern implied roles such as pollen tube guidance, pollen tube capture and release of gametes, fertilization or initiation of seed development. The role of *AGL104* in the female gametophyte was not investigated further due to time constraints.

In conclusion, our preliminary characterization of *AGL104* has revealed some novel information previously not reported. However, in doing so, a lot of new questions have emerged. Completion and improvement of the current on-going experiments in our laboratory and the design of new experiments to answer specific questions will most likely reveal the precise role of *AGL104* in angiosperms. Thus, future directions for this project include further assessment of the expression pattern of *AGL104*, its influence on pollen germination and pollen tube elongation, its role in the female gametophyte, and its interaction with other MIKC* genes as outlined below.

Our GUS expression analysis revealed a specific expression pattern of *AGL104* in developing anthers, mature pollen, during pollen germination, in the elongating pollen tube and in the embryo sac only. However, together with pollen and ovules, literature published previously had revealed expression in additional inflorescence structures such as sepals, petals and the septum of the carpel (Parenicova et al., 2003). In order to eliminate this discrepancy in the expression pattern of *AGL104*, a larger fragment of the *AGL104* promoter, which may contain more regulatory elements, should to be fused to the GUS protein and the mutants analyzed accordingly. Alternatively, other expression analysis methods such *in situ* hybridization assays may be used. This data, we envision, will clarify the exact domain of expression of *AGL104* and enable scientists to narrow down the structures for future functional analyses.

The expression of *AGL104* in mature pollen had prompted us to examine its expression pattern during pollen germination and tube elongation. During the course of our study, it was realized that since the GUS protein is very stable, its detection through the GUS assay could have been due to either its active synthesis during pollen germination and tube elongation or its persistence from the mature ungerminated pollen, i.e. it was a remnant protein present during

these processes. Since this piece of data has huge implication on the role of transcription factors during pollen germination and tube elongation, we suggest further studies to differentiate between the two possibilities. *In situ* hybridization of the elongated pollen tube would reveal if the *AGL104* transcripts are present in the tube, which would in turn imply its active synthesis through translation. In addition, determining the locale of the AGL104 protein would also shed light on its role in the pollen tube. Immuno-precipitation assay combined with DAPI staining is suggested. Localization of *AGL104* to the nuclei during pollen germination and pollen tube elongation would be a positive indicator of its active participation in gene regulation during these processes, which will lend further support to our functional analysis data.

As mentioned above, gain-of-function analysis revealed that overexpression of *AGL104* promoted pollen germination. We had generated double 35S (35S::35S::*AGL104*) mutants in order to overexpress *AGL104* at a much higher level than the single 35S mutants that was generated initially in this study. However, we were unable to determine the level of expression of *AGL104* in these two mutant types. Future directions should involve determining the expression level of *AGL104* in these two mutant types, which would then help establish whether the germination-promoting phenotype of *AGL104* is dosage-dependent, that is, whether a higher level of *AGL104* expression leads to a much higher frequency of pollen germination.

When drawing conclusion on the role of *AGL104* in pollen tube elongation, we realized that discrepancies in the experimental design may have led to an uninformative data set. Since we pollinated at least fifteen pistils per day, there was a time difference of 45 min to an hour between the first and last pollination. We suggest a repeat of the *in vivo* experiment with either a reduction in the number of pollinations per day or simultaneous pollination by a group of people. This will reduce the difference in time allowed for the tubes to elongate so that a reliable data set is obtained. In addition, it was noted that the tubes did not grow straight but coiled several times. To eliminate discrepancy in the length of the measured tubes caused by its growth pattern (coiling), we also suggest a larger sample size for each mutant line. Furthermore, we noted that a clearer pattern was observable when pollen was allowed to elongate for at least 4 hours. Thus, future experiments investigating the length of pollen tubes grown for at least 4 hours in all the overexpressing mutant types generated in this study is suggested to help decipher the effect *AGL104* has on pollen tube elongation.

Data obtained through the analyses of our loss-of-function mutants was expected to compliment the gain-of-function mutant analysis data in order to establish the role of *AGL104*. However, since we were unable to optimize the Northern Blot conditions to determine if the expression level of *AGL104* was changed in the loss-of-function mutants, our future direction should foremost involve optimizing the conditions in order to obtain this information. A thorough phenotypic analysis of the lines showing decreased or diminished *AGL104* expression will then help determine its role.

It was also noted that pollen germination and tube elongation was probably an end result of a large number of processes acting together. Thus, examining the other genes that are upregulated or downregulated during the time of *AGL104* expression will help determine the downstream genes involved. Transcriptome analysis can provide this information. Promoter analyses of these downstream genes will then help determine if they are regulated by *AGL104* and determining the putative roles of these genes will reveal the processes which are controlled by *AGL104*. In addition, analysis of the promoter region of *AGL104* itself will reveal putative regulatory motifs and will provide information on which genes regulate *AGL104* expression and the networks and pathways of which *AGL104* is a member, therefore give clues to its role in pollen.

In a wider context, since most MIKC* type genes have been reported to be highly expressed in the mature pollen (Honys and Twell, 2004; Kofuji et al., 2003; Pina et al., 2005; Verelst et al., 2007), it would be interesting to investigate if other MIKC* genes are also involved during pollen germination and tube elongation similar to *AGL104*. In our study, we had generated multiple overexpression mutants using *AGL104*, *AGL66* and *AGL67* due to their close phylogenetic relationship which led us to suspect that they may act in a combinatorial manner. However, publications following our study implicated only *AGL104* and *AGL66* in pollen germination and tube elongation (Honys and Twell, 2004; Pina et al., 2005). Since a recent study implied redundancy between these two genes (Verelst et al., 2007), future experiments could involve determining the frequency of pollen germination in the single mutants of *AGL104* and *AGL66*, that is, 35S::*AGL104* and 35S::*AGL66*, in order to establish if both genes had the same influence on the frequency of pollen germination (i.e. completely redundant) or if one gene promoted pollen germination more than the other.

To thoroughly investigate the influence of the other MIKC* genes (*AGL66*, *AGL65*, *AGL30*, *AGL94*) in pollen, it would be informative to determine if their expression patterns were similar to that of *AGL104*, that is, whether these genes were also expressed during pollen germination and tube elongation. Generation and analysis of GUS fusion markers will provide this information. Once the expression pattern is established, this can be followed by phenotypic analysis similar to the experiments carried out for *AGL104* characterization. It is noted that the results of Verelst et al (2007) implied the action of certain functional heterodimers (formed by pairs of MIKC* genes) in pollen germination. It is also noted that Verelst et al (2007) looked at only the pollen germination phenotype (*in vitro*) of their loss-of-function mutants. Therefore, future experiments should involve, firstly, obtaining the loss-of-function mutants generated by Verelst et al (2007) and performing *in vivo* assays to determine the influence of these genes during pollen tube elongation, and secondly, generation and analyses (both *in vivo* and *in vitro*) of gain-of-function mutants expressing the precise MIKC* gene pairs involved in the formation of these functional heterodimers. These experiments will provide a complementary set of data to that obtained by Verelst et al (2007) and thus comparison of the mutant phenotype of all the possible heterodimer-forming MIKC* genes pairs will help establish the correct function of the MIKC* genes during pollen germination and tube elongation.

It is also crucial to characterize the function of *AGL104* in the female gametophyte. Firstly, we have to determine whether *AGL104* was expressed in the egg, synergid or both cell types of the egg apparatus. This can be achieved by determining the position of the nuclei of the synergid and egg cell. The nucleus of the egg cell is located on the chalazal pole of the egg apparatus whereas that of the synergid cells is located at the micropylar pole (Christensen et al., 1997). DAPI staining superimposed on GUS stained embryo sac can easily determine which of the cell types exhibit *AGL104* expression. Once the expression pattern of *AGL104* is established, the phenotype of the embryo sac in the gain-of-function and loss-of-function mutants can be examined. Since the synergid cells and egg cell have been implicated in pollen tube guidance, pollen tube reception, fertilization, embryogenesis, and seed formation (Higashiyama et al., 2001), these phenotypes can be examined carefully to determine if an overexpression or loss-of-function disrupts these processes. Proper pollen tube guidance can be examined using Analine blue staining; pollen tube reception and fertilization can be examined by staining the penetrated synergid cell with DAPI to determine the release of sperm cells and its fusion with the egg and central nuclei; and embryogenesis and seed formation can be examined by dissecting the ovules and seeds, respectively. In addition, since the synergid cells have been

implicated in pollen tube guidance, it may be possible that they also influence the rate of tube elongation. Thus, the effect of the synergid cells on the rate of tube elongation also needs to be examined. *AGL104* gain-of-function and loss-of-function pistils can be pollinated with wt pollen, *AGL104* gain-of-function pollen and *AGL104* loss-of-function pollen separately and the length of their pollen tubes measured to obtain this data.

Overall, the underlying processes involved in pollen germination and pollen tube elongation remain to be elucidated, but it is likely the result of multiple factors. Pollen germination and tube growth is crucial to ensure fertilization since the survival of a species is dependent on it. It is, therefore, not surprising that the plant would utilize many different means. Similar to flowering, where many different pathways exist, it may be possible that there are many pathways to ensure pollen germination and tube elongation due to its crucial role. It appears that *AGL104* and *AGL66* are high up in this network of pathways since their abolishment leads to a drastic decrease in pollen germination frequency (Verelst et al., 2007). In addition, the fact that another transcription factor was downregulated in the *agl66/104* double mutant confirms the existence of a complex network (Verelst et al., 2007). The complexity of a transcription factor network may also be present in the female gametophyte. This being the case, functional characterization of *AGL104* in pollen and the embryo sac thus requires much more stringent phenotypic examination, including the construction of multiple mutant combinations, transcript profiling, and perhaps proteomic and metabolic analyses. The experiments suggested above will not only determine the function of *AGL104* in the male and female gametophyte of angiosperms, it will also provide specific (yet unknown) information on the different types of processes involved during the development of the gametophytes. In addition, establishing the function of *AGL104*, a MIKC* gene, will lay the groundwork for comparison of the roles of MIKC* genes in mosses and angiosperms and help evolutionary biologist to decipher the changes that have occurred in land plant body form during evolution(Alonso, 2003; Alvarez-Buylla et al., 2000).

Appendix 1

List of Primers

Application	Primer Name	Sequence
Cloning *AGL104* cDNA	104FcDNA	CTCTTCCTGCTTCGCCAATATG
	OASR002	CTTCATCGTCTTCTCATGGC
Cloning promoter/upstream sequence of *AGL104*	001AGL104	GACACCAATATAGCAACCG
	002AGL104	CGTAACTTGTCGATTCGTTG
Cloning IKC region of *AGL104*	003AGL104Xba5	CGTCTAGACAAGAACGAGAGAGTGCTC
	004AGL104Xba3	CGTCTAGACTTCATCGTCTTCTCATGGC
Cloning T-DNA fragment in *AGL104*	LP013898	TCTTCATCGTCTTCTCATGGCA
	RP013898	TGAAAAAGCCCTGCAGCTATCA
	LBb1	GCGTGGACCGCTTGCTGCAACT
Cloning *AGL66* cDNA	66FcDNA	GGTTGCATTGATCGGATATAAG
	66RcDNA	CCTTGAGATTATTGGCTACTGAGC
Cloning *AGL67* cDNA	67FcDNA	GCAGTGACCATCGGATATGG
	67RcDNA	GAGGTAAAGATTAGCTAGCAG
Sequencing	M13F	TTGTAAAACGACGGCCAG
	M13R	CAGGAAACAGCTATGACCATG

Appendix 2

Recipe for buffers and solutions

STET solution

The STET solution contained 8% sucrose, 5% triton X-100, 50mM EDTA and 50mM Tris at pH 8.0.

Infiltration solution (for resuspending *Agrobacterium* during transformation of *Arabidopsis*) (Bent, 1998)

The infiltration solution contained 6.6g of MS salt, 0.75g of MES salt and 75g of sucrose in 1.5l of autoclaved water. The pH was adjusted to 5.6-5.8 using 1M Potassium Hydroxide (KOH). Then 15µl of BAP (1mg/ml) and 300ul Silwet 77 (Osi Specialties, Inc., Danbury, CT, USA) was added. This media was used to resuspend *A.tumefaciens* containing the desired plasmids before dipping the plants.

TNE buffer (10X)

10X TNE solution contained 100mM Tris, 10mM EDTA and 1M NaCl. The buffer had its pH adjusted to 7.4 with HCl. 10X TNE was diluted 10-fold with MQ water to make a 1X TNE solution and 10µl of 1mg/ml Hoechst Dye (Amersham Biosciences, NJ USA) was added to 100ml 1X TNE buffer which was used for DNA quantification with fluorometer.

Ligation Buffer (2X)

Ligation buffer used in lab was purchased from PROMEGA (PROMEGA, Madison, WI USA). The 2X ligation buffer contained 60mM Tris-HCl (pH 7.8), 20mM Magnesium Chloride, 20mM DTT, 2mM ATP and 10% PEG.

TE buffer

TE buffer contained 10mM Tris at pH8.0 and 1mM EDTA

Loading Buffer (10X)

10X gel loading buffer contained 1% SDS, 50% glycerol, and 0.05% bromophenol blue. 10X buffer was then diluted 10-fold with DNA solution and run on gel.

TBE buffer (5X)

5X TBE buffer contained 54g Tris base, 27.5g boric acid and 20ml 0.5M EDTA (pH8.0) per 1L of buffer. 5X buffer was then diluted 10-fold to 0.5X TBE buffer which was used in running nucleic acid samples on gel

MOPS (10X)

10X MOPS buffer contained 0.2M MOPS (pH7.0), 20mM Sodium acetate, and 10mM EDTA (pH8.0)

DNA Extraction Buffer (Edwards and Thomson, 1991)

The extraction buffer contained 20ml 1Mtris-HCl (pH 7.5), 5ml M NaCl, 5ml 0.5M EDTA (pH 8), 5ml 10% SDS and 65ml sterile water.

DAPI buffer

The buffer used in DAPI staining contained 5.77ml of 1M NaH_2PO_4, 4.23ml 1M Na_2HPO_4, 0.2ml 500mM EDTA and 0.33ml 30% Triton X. The total volume was made up to 100ml with sterile water.

GUS staining solution

The GUS staining solution contained 0.2% triton X-100, 50mM sodium phosphate buffer (pH 7.2), 2mM ferrocyanide, 2mM ferricyanide and 2mM X-Gluc (in DMF).

FAA

FAA contained 50% ethanol, 5% formaldehyde (using a 37% stock) and 10% glacial acetic acid.

1 kb Plus DNA Ladder

The working solution (final concentration 50ng/µl) of the 1kb Plus DNA ladder (Invitrogen Carlsbad, CA USA) contained 50µl of the stock ladder (1mg/ml), 100µl of 10X loading buffer and 850ul of TE (section 2.2.5). The ladder was kept in aliquots of 200µl.

Gycerol Stock of plasmids

Glycerol stocks were prepared by adding 1:1 volume of the overnight bacterial culture and 30% glycerol. These glycerol stocks were then "snap"-frozen in liquid nitrogen and stored at -80^{o}C. Glycerol stocks of 1ml were kept.

REFERENCES

Alvarez-Buylla, E. R., Liljegren, S. J., Pelaz, S., Gold, S. E., Burgeff, C., Ditta, G. S., Vergara-Silva, F., and Yanofsky, M. F. (2000a). MADS-box gene evolution beyond flowers: expression in pollen, endosperm, guard cells, roots and trichomes. Plant Journal *24*, 457-466.

Alvarez-Buylla, E. R., Pelaz, S., Liljegren, S. J., Gold, S. E., Burgeff, C., Ditta, G. S., de Pouplana, L. R., Martinez-Castilla, L., and Yanofsky, M. F. (2000b). An ancestral MADS-box gene duplication occurred before the divergence of plants and animals. Proceedings of the National Academy of Sciences of the United States of America *97*, 5328-5333.

Ambrose, B. A., Lerner, D. R., Ciceri, P., Padilla, C. M., Yanofsky, M. F., and Schmidt, R. J. (2000). Molecular and genetic analyses of the silky1 gene reveal conservation in floral organ specification. Molecular Cell *5*, 569-579.

Baltz, R., Domon, C., Pillay, D. T., and Steinmetz, A. (1992). Characterization of a pollen-specific cDNA from sunflower encoding a zinc finger protein. Plant J *2*, 713-721.

Bate, N., and Twell, D. (1998). Functional architecture of a late pollen promoter: pollen-specific transcription is developmentally regulated by multiple stage-specific and co-dependent activator elements. Plant Mol Biol *37*, 859-869.

Baulcombe, D. (2004). RNA silencing in plants. Nature *431*, 356-363.

Becker, A., and Theissen, G. (2003). The major clades of MADS-box genes and their role in the development and evolution of flowering plants. Molecular Phylogenetics and Evolution *29*, 464-489.

Becker, J. D., Boavida, L. C., Carneiro, J., Haury, M., and Feijo, J. A. (2003). Transcriptional profiling of Arabidopsis tissues reveals the unique characteristics of the pollen transcriptome. Plant Physiology *133*, 713-725.

Bent, S. J. C. a. A. F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant journal *16*, 735-743.

Besser, K., Frank, A. C., Johnson, M. A., and Preuss, D. (2006). *Arabidopsis HPA2* (*GCS1*) is a sperm-specific gene required for pollen tube guidance and fertilization. Development *133*, 4761-4769.

Bingham, P. M., Levis, R., and Rubin, G. M. (1981). Cloning of DNA sequences from the white locus of D. melanogaster by a novel and general method. Cell *25*, 693-704.

Blackwell, W. H. (2003). Two theories of origin of the land-plant sporophyte: which is left standing? The botanical review *69*, 125-148.

Blanc, G., Hokamp, K., and Wolfe, K. H. (2003). A recent polyploidy superimposed on older large-scale duplications in the Arabidopsis genome. Genome Research *13*, 137-144.

Blanc, G., and Wolfe, K. H. (2004). Functional divergence of duplicated genes formed by polyploidy during Arabidopsis evolution. Plant Cell *16*, 1679-1691.

Bloch, D., Lavy, M., Efrat, Y., Efroni, I., Bracha-Drori, K., Abu-Abied, M., Sadot, E., and Yalovsky, S. (2005). Ectopic expression of an activated RAC in Arabidopsis disrupts membrane cycling. Molecular Biology of Cell *16*, 1913-1927.

Bosher, J. M., and Labouesse, M. (2000). RNA interference: genetic wand and genetic watchdog. Nat Cell Biol *2*, E31-36.

Bottino, P.J. (2005). Gus gene assay in transformed tissues. Viewed 24 May 2005. <http://www.inform.umd.edu/genetics/GUS.html>

Bowers, J. E., Chapman, B. A., Rong, J. K., and Paterson, A. H. (2003). Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. Nature *422*, 433-438.

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. (1991). Genetic interactions among floral homeotic genes of *Arabidopsis*. Development *112*, 1-20.

Bradley, D., Carpenter, R., Sommer, H., Hartley, N., and Coen, E. (1993). Complementary floral homeotic phenotypes result from opposite orientations of a transposon at the plana locus of antirrhinum. Cell *72*, 85-95.

Byzova, M. V., Franken, J., Aarts, M. G. M., de Almeida-Engler, J., Engler, G., Mariani, C., Campagne, M. M. V., and Angenent, G. C. (1999). Arabidopsis STERILE APETALA, a multifunctional gene regulating inflorescence, flower, and ovule development. Genes & Development *13*, 1002-1014.

Chaudhury, A. M. (1993). Nuclear Genes Controlling Male Fertility. Plant Cell *5*, 1277-1283.

Chaudhury, A. M., Ming, L., Miller, C., Craig, S., Dennis, E. S., and Peacock, W. J. (1997). Fertilization-independent seed development in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences USA *94*.

Chen, Y. C., and McCormick, S. (1996). sidecar pollen, an Arabidopsis thaliana male gametophytic mutant with aberrant cell divisions during pollen development. Development *122*, 3243-3253.

Christensen, C. A., King, E. J., Jordan, J. R., and Drews, G. N. (1997). Megagametogenesis in Arabidopsis wild type and the Gf mutant. Sexual Plant Reproduction *10*, 49-64.

Chuang, C. F., and Meyerowitz, E. M. (2000). Specific and heritable genetic interference by double-stranded RNA in Arabidopsis thaliana. Proc Natl Acad Sci U S A *97*, 4985-4990.

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

Colombo, L., Franken, J., Koetje, E., Vanwent, J., Dons, H. J. M., Angenent, G. C., and Vantunen, A. J. (1995). The Petunia Mads Box Gene Fbp11 Determines Ovule Identity. Plant Cell *7*, 1859-1868.

da Costa-Nunes, J. A., and Grossniklaus, U. (2003). Unveiling the gene-expression profile of pollen. Genome Biology *5*, 205.

De Bodt, S., Raes, J., Florquin, K., Rombauts, S., Rouze, P., Theissen, G., and Van de Peer, Y. (2003a). Genomewide structural annotation and evolutionary analysis of the type I MADS-box genes in plants. Journal of Molecular Evolution *56*, 573-586.

De Bodt, S., Raes, J., Van de Peer, Y. V., and Theissen, G. (2003b). And then there were many: MADS goes genomic. Trends in Plant Science *8*, 475-483.

de Folter, S., Busscher, J., Colombo, L., Losa, A., and Angenent, G. C. (2004). Transcript profiling of transcription factor genes during silique development in Arabidopsis. Plant Molecular Biology *56*, 351-366.

de Folter, S., Immink, R. G. H., Kieffer, M., Parenicova, L., Henz, S. R., Weigel, D., Busscher, M., Kooiker, M., Colombo, L., Kater, M. M., *et al.* (2005). Comprehensive interaction map of the Arabidopsis MADS box transcription factors. Plant Cell *17*, 1424-1433.

de Graaf, B. H. e. a. (2005). Rab11 GTPAse-regulated membrane trafficking is crucial for tip-focused pollen tube growth in tobacco. Plant Cell *17*, 2564-2579.

Delwiche, C. F., Karol, K. G., Cimino, M. T., and Sytsma, K. J. (2002). Phylogeny of the genus Coleochaete (Coleochaetales, Charophyta) and related taxa inferred by analysis of the chloroplast gene rbcL. Journal of Phycology *38*, 394-403.

Derksen, J., Knuiman, B., Hoedemaekers, K., Guyon, A., Bonhomme, S., and Pierson, E. S. (2002). Growth and cellular organization of Arabidopsis pollen tubes in vitro. Sexual Plant Reproduction *15*, 133-139.

Diethard Mattanovich, F. R., da Camara Machado, A., Laimer, M., Regner, F., Steinkellner, H., Himmler, G., and Katinger, H. (1989). Efficient transformation of Agrobacterium spp. by electroporation. Nucleic Acids Research *17*, 6747.

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. Current Biology *14*, 1935-1940.

Doebley, J., and Lukens, L. (1998). Transcriptional regulators and the evolution of plant form. Plant Cell *10*, 1075-1082.

Eady, C., Lindsey, K., and Twell, D. (1995). The Significance of Microspore Division and Division Symmetry for Vegetative Cell-Specific Transcription and Generative Cell-Differentiation. Plant Cell *7*, 65-74.

Edwards, K. J., and Thompson, C. (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. NUCLEIC ACIDS RESEARCH *19*, 1349-1349.

Egea-Cortines, M., Saedler, H., and Sommer, H. (1999). Ternary complex formation between the MADS-box proteins SQUAMOSA, DEFICIENS and GLOBOSA is involved in the control of floral architecture in Antirrhinum majus. Embo Journal *18*, 5370-5379.

Eyal, Y., Curie, C., and McCormick, S. (1995). Pollen specificity elements reside in 30 bp of the proximal promoters of two pollen-expressed genes. Plant Cell *7*, 373-384.

Fan, L. M., Wang, Y. F., Wang, H., and Wu, W. H. (2001). In vitro Arabidopsis pollen germination and characterization of the inward potassium currents in Arabidopsis pollen grain protoplasts. Journal of Experimental Botany *52*, 1603-1614.

Feijo, J. A., Malho, R., and Obermeyer, G. (1995). Ion dynamics and its possible role during *in vitro* pollen germination and tube growth. Protoplasma *187*, 155-167.

Feijo, J. A., and Moreno, N. (2004). Imaging plant cells by two-photon excitation. Protoplasma *223*, 1-32.

Feijo, J. A., Sainhas, J., Holdaway-Clarke, T., Cordeiro, M. S., Kunkel, J. G., and Hepler, P. K. (2001). Cellular oscillations and the regulation of growth: the pollen tube paradigm. Bioessays *23*, 86-94.

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*, 725-734.

Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. L., and Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. Genetics *151*, 1531-1545.

Foreman, J., Demidchik, V., Bothwell, J. H. F., and Mylona, P. (2003). Reactive oxygen species produced by NADPH oxidase regulate plant cell growth. Nature *422*, 422-446.

Franke, W. W., Herth, W., Woude, W. J. V. d., and Moore, D. J. (1972). Tubular filamentous structures in pollen tubes: Possible involvement as guide elements in protoplasmic streaming and vectorial migration of secretory vesicles. Planta *105*, 317-341.

Fridborg, I., Kuusk, S., Moritz, T., and Sundberg, E. (1999). The Arabidopsis dwarf mutant shi exhibits reduced gibberellin responses conferred by overexpression of a new putative zinc finger protein. Plant Cell *11*, 1019-1032.

Galbiati, M., Moreno, M. A., Nadzan, G., Zourelidou, M., and Dellaporta, S. L. (2000). Large-scale T-DNA mutagenesis in Arabidopsis for functional genomic analysis. Funct Integr Genomics *1*, 25-34.

Gilbert, S. F., Opitz, J. M., and Raff, R. A. (1996). Resynthesizing evolutionary and developmental biology. Dev Biol *173*, 357-372.

Gleave, A. P. (1992). A versatile binary vector system with a T-DNA organisational structure conducive to efficient integration of cloned DNA into the plant genome. Plant Mol Biol *20*, 1203-1207.

Goto, K., and Meyerowitz, E. M. (1994). Function and regulation of the Arabidopsis floral homeotic gene PISTILLATA. Genes Dev *8*, 1548-1560.

Graham, L. E., and Wilcox, L. W. (1983). The Occurrence and Phylogenetic Significance of Putative Placental-Transfer Cells in the Green-Alga Coleochaete. American Journal of Botany *70*, 113-120.

Grossniklaus, U., and Schneitz, K. (1998). The molecular and genetic basis of ovule and megagametophyte development. Seminars in Cell and Developmental Biology *9*, 227-238.

Gu, Q., Ferrandiz, C., Yanofsky, M. F., and Martienssen, R. (1998). The FRUITFULL MADS-box gene mediates cell differentiation during Arabidopsis fruit development. Development *125*, 1509-1517.

Gu, Z. L., Nicolae, D., Lu, H. H. S., and Li, W. H. (2002). Rapid divergence in expression between duplicate genes inferred from microarray data. Trends in Genetics *18*, 609-613.

Guyon, V. N., Astwood, J. D., Garner, E. C., Dunker, A. K., and Taylor, L. P. (2000). Isolation and characterization of cDNAs expressed in the early stages of flavonol-induced pollen germination in petunia. Plant Physiol *123*, 699-710.

Haig, D., and Wilczek, A. (2006). Sexual conflict and the alternation of haploid and diploid generations. Philosophical Transactions of the Royal Society B-Biological Sciences *361*, 335-343.

Hartmann, U., Hohmann, S., Nettesheim, K., Wisman, E., Saedler, H., and Huijser, P. (2000). Molecular cloning of SVP: a negative regulator of the floral transition in Arabidopsis. Plant Journal *21*, 351-360.

Hasebe, M., Wen, C. K., Kato, M., and Banks, J. A. (1998). Characterization of MADS homeotic genes in the fern Ceratopteris richardii. Proceedings of the National Academy of Sciences of the United States of America *95*, 6222-6227.

Heard, J., Caspi, M., and Dunn, K. (1997). Evolutionary diversity of symbiotically induced nodule MADS box genes: Characterization of nmhC5, a member of a novel subfamily. Molecular Plant-Microbe Interactions *10*, 665-676.

Henschel, K., Kofuji, R., Hasebe, M., Saedler, H., Munster, T., and Theissen, G. (2002). Two ancient classes of MIKC-type MADS-box genes are present in the moss Physcomitrella patens. Molecular Biology and Evolution *19*, 801-814.

Hepler, C. D. (2001). Regulating for outcomes as a systems response to the problem of drug-related morbidity. J Am Pharm Assoc (Wash) *41*, 108-115.

Heslop-Harrison, J., and Heslop-Harrison, Y. (1982). The growth of grass pollen tube: Characteristics of the polysaccharide particles ("p-particles") associated with apical growth. Protoplasma *112*, 71-80.

Heuer, S., Lorz, H., and Dresselhaus, T. (2000). The MADS box gene ZmMADS2 is specifically expressed in maize pollen and during maize pollen tube growth. Sex Plant Reproduction *13*, 21-27.

Higashiyama, T., Inatsugi, R., Sakamoto, S., Sasaki, N., Mori, T., Kuroiwa, H., Nakada, T., Nozaki, H., Kuroiwa, T., and Nakano, A. (2006). Species preferentiality of the pollen tube attractant derived from the synergid cell of *Torenia fournieri*. Plant Physiology *142*, 481-491.

Higashiyama, T., Yabe, S., Sasaki, N., Nishimura, Y., Miyagishima, S., Kuroiwa, H., and Kuroiwa, T. (2001). Pollen tube attraction by the synergid cell. Science *293*, 1480-1483.

Hoekstra, F. A., and Bruinsma, J. (1979). Protein synthesis of binucleate and trinucleate pollen and its relationship to tube emergence and growth. Planta *146*, 559-566.

Honma, T., and Goto, K. (2001). Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. Nature *409*, 525-529.

Honys, D., Combe, J. P., Twell, D., and Capkova, V. (2000). The translationally repressed pollen-specific ntp303 mRNA is stored in non-polysomal mRNPs during pollen maturation. Sexual Plant Reproduction *13*, 135-144.

Honys, D., and Twell, D. (2003). Comparative analysis of the Arabidopsis pollen transcriptome. Plant Physiology *132*, 640-652.

Honys, D., and Twell, D. (2004). Transcriptome analysis of haploid male gametophyte development in Arabidopsis. Genome Biology *5*, R85.

Huang, B. Q., and Russell, S. D. (1992). Female Germ Unit - Organization, Isolation, and Function. International Review of Cytology-a Survey of Cell Biology *140*, 233-293.

Huang, B. Q., and Russell, S. D. (1994). Fertilization in *Nicotiana tabacum*: cytoskeletal modifications in the embryo sac during synergid degeneration. Planta *194*, 200-214.

Huck, N., Moore, J. M., Federer, M., and Grossniklaus, U. (2003). The Arabidopsis mutant feronia disrupts the female gametophytic control of pollen tube reception. Development *130*.

Hughes, A. L. (1994). The Evolution of Functionally Novel Proteins after Gene Duplication. Proceedings of the Royal Society of London Series B-Biological Sciences *256*, 119-124.

Huijser, P., Klein, J., Lonnig, W. E., Meijer, H., Saedler, H., and Sommer, H. (1992). Bracteomania, an inflorescence anomaly, is caused by the loss of function of the MADs-box gene squamosa in Antirrhinum majus. EMBO Journal *11*, 1239-1249.

Hulskamp, M., Kopczak, S. D., Horejsi, T. F., Kihl, B. K., and Pruitt, R. E. (1995). Identification of Genes Required for Pollen-Stigma Recognition in Arabidopsis-Thaliana. Plant Journal *8*, 703-714.

Ichiro Kasajima, Y. I., Ohkama-Ohtsu, N., Hayashi, H., Yoneyama, T., and Fujiwara, T. (2004). A Protocol for Rapid DNA Extraction From Arabidopsis thaliana for PCR Analysis. Plant Molecular Biology Reporter *22*, 49–52.

Immink, R. G. H., Gadella, T. W. J., Ferrario, S., Busscher, M., and Angenent, G. C. (2002). Analysis of MADS box protein-protein interactions in living plant cells. Proceedings of the National Academy of Sciences of the United States of America *19*, 2416-2421.

Irish, V. F., and Litt, A. (2005). Flower development and evolution: gene duplication, diversification and redeployment. Curr Opin Genet Dev *15*, 454-460.

Irish, V. F., and Sussex, I. M. (1990). Function of the *APETALA-1* gene during *Arabidopsis* floral development. Plant Cell *2*, 741-751.

Jack, T., Brockman, L. L., and Meyerowitz, E. M. (1992). The homeotic gene APETALA3 of Arabidopsis thaliana encodes a MADS box and is expressed in petals and stamens. Cell *68*, 683-697.

Jaffe, L. F., Robinson, K. R., and Nuccitelli, R. (1974). Local cation entry and self-electrophoresis as an intracellular localization mechanism. Ann N Y Acad Sci *238*, 372-389.

Kandasamy, M. K., Nasrallah, J. B., and Nasrallah, M. E. (1994). Pollen Pistil Interactions and Developmental Regulation of Pollen-Tube Growth in Arabidopsis. Development *120*, 3405-3418.

Kardailsky, I., Shukla, V. K., Ahn, J. H., Dagenais, N., Christensen, S. K., Nguyen, J. T., Chory, J., Harrison, M. J., and Weigel, D. (1999). Activation tagging of the floral inducer FT. Science *286*, 1962-1965.

Karol, K. G., McCourt, R. M., Cimino, M. T., and Delwiche, C. F. (2001). The closest living relatives of land plants. Science *294*, 2351-2353.

Kasahara, R. D., Portereiko, M. F., Sandaklie-Nikolova, L., Rabiger, D. S., and Drews, G. N. (2005). *MYB98* is required for pollen tube guidance and synergid cell differentiation in *Arabidopsis*. The Plant Cell *17*, 2981-2992.

Kempin, S. A., Savidge, B., and Yanofsky, M. F. (1995). Molecular basis of the cauliflower phenotype in Arabidopsis. Science *267*, 522-525.

Kenrick, P., and Crane, P. R. (1997). The origin and early evolution of plants on land. Nature *389*, 33-39.

Kim, H. U., Cotter, R., Johnson, S., Senda, M., Dodds, P., Kulikauska, R., Tang, W., Ezcura, I., Herzmark, P., and McCormick, S. (2002). New pollen-specific receptor kinases identified in tomato, maize and Arabidopsis: the tomato kinases show overlapping but distinct localization patterns on pollen tubes. Plant Mol Biol *50*, 1-16.

Kirik, V., Schnittger, A., Radchuk, V., Adler, K., Hulskamp, M., and Baumlein, H. (2001). Ectopic expression of the Arabidopsis AtMYB23 gene induces differentiation of trichome cells. Dev Biol *235*, 366-377.

Klahre, U., Becker, C., Schmitt, A. C., and Kost, B. (2006). Nt-RhoGD12 regulates Rac/Rop signaliing and polar cell growth in tobacco pollen tubes. Plant Journal *46*.

Klucher, K. M., Chow, H., Reiser, L., and Fischer, R. L. (1996). The AINTEGUMENTA gene of arabidopsis required for ovule and female gametophyte development is related to the floral homeotic gene APETALA2. Plant Cell *8*, 137-153.

Kobayashi, A., Sakamoto, A., Kubo, K., Rybka, Z., Kanno, Y., and Takatsuji, H. (1998). Seven zinc-finger transcription factors are expressed sequentially during the development of anthers in petunia. Plant J *13*, 571-576.

Kofuji, R., Sumikawa, N., Yamasaki, M., Kondo, K., Ueda, K., Ito, M., and Hasebe, M. (2003). Evolution and divergence of the MADS-box gene family based on genome-wide expression analyses. Molecular Biology and Evolution *20*, 1963-1977.

Krizek, B. A., and Meyerowitz, E. M. (1996). The Arabidopsis homeotic genes APETALA3 and PISTILLATA are sufficient to provide the B class organ identity function. Development *122*, 11-22.

Krogan, N. T., and Ashton, N. W. (2000). Ancestry of plant MADS-box genes revealed by bryophyte (Physcomitrella patens) homologues. New Phytologist *147*, 505-517.

Kubo, H., Peeters, A. J., Aarts, M. G., Pereira, A., and Koornneef, M. (1999). ANTHOCYANINLESS2, a homeobox gene affecting anthocyanin distribution and root development in Arabidopsis. Plant Cell *11*, 1217-1226.

Lamb, R. S., and Irish, V. (2003). Functional divergence within the APETALA3/PISTILLATA floral homeotic gene lineages Proceedings of the National Academy of Sciences of the United States of America *100*, 6558-6563.

Lauri, A., Xing, S., Heidmann, I., Saedler, H., and Zachgo, S. (2006). The pollen-specific DEFH125 promoter from Antirrhinum is bound in vivo by the MADS-box proteins DEFICIENS and GLOBOSA. Planta *224*, 61-71.

Lehti-Shiu, M. D., Adamczyk, B. J., and Fernandez, D. E. (2005). Expression of MADS-box genes during the embryonic phase in Arabidopsis. Plant Molecular Biology *58*, 89-107.

Lennon, K. A., Roy, S., Hepler, P. K., and Lord, E. M. (1998). The structure of the transmitting tissue of *Arabidopsis thaliana* (L.) and the path of pollen tube growth. Sex Plant Reproduction *11*, 49-59.

Liljegren, S. J., Ditta, G. S., Eshed, H. Y., Savidge, B., Bowman, J. L., and Yanofsky, M. F. (2000). SHATTERPROOF MADS-box genes control seed dispersal in Arabidopsis. Nature *404*, 766-770.

Long, M. Y., and Langley, C. H. (1993). Natural-Selection and the Origin of Jingwei, a Chimeric Processed Functional Gene in Drosophila. Science *260*, 91-95.

Lynch, M., and Conery, J. S. (2000). The evolutionary fate and consequences of duplicate genes. Science *290*, 1151-1155.

Lynch, M., and Force, A. (2000). The probability of duplicate gene preservation by subfunctionalization. Genetics *154*, 459-473.

Lynch, M., O'Hely, M., Walsh, B., and Force, A. (2001). The probability of preservation of a newly arisen gene duplicate. Genetics *159*, 1789-1804.

Makova, K. D., and Li, W. H. (2003). Divergence in the spatial pattern of gene expression between human duplicate genes. Genome Research *13*, 1638-1645.

Mandaron, P., Niogret, M., Mache, R., and Monegar, F. (1990). In vitro pollen synthesis in isolated microspores of Zea mays at several stages of development. Theoretical and Applied Genetics *80*, 134-138.

Mandel, M. A., Gustafson-Brown, C., Savidge, B., and Yanofsky, M. F. (1992). Molecular characterization of the Arabidopsis floral homeotic gene APETALA1. Nature *360*, 273-277.

Marsch-Martinez, N., Greco, R., Van Arkel, G., Herrera-Estrella, L., and Pereira, A. (2002). Activation tagging using the En-I maize transposon system in Arabidopsis. Plant Physiol *129*, 1544-1556.

Martinez-Castilla, L. P., and Alvarez-Buylla, E. R. (2003). Adaptive evolution in the Arabidopsis MADS-box gene family inferred from its complete resolved phylogeny.

Proceedings of the National Academy of Sciences of the United States of America *100*, 13407-13412.

Marton, M. L., Cordts, S., Broadhvest, J., and Dresselhaus, T. (2005). Micropylar pollen tube guidance by Egg Appratus 1 of maize. Science *307*.

Mascarenhas, J. (1975a). The male gametophyte of flowering plants. Plant Cell *1*, 657-664.

Mascarenhas, J. P. (1975b). The biochemistry of angiosperm pollen development. . Botanical Review *41*, 259-314.

Mascarenhas, J. (1989). The male gametophyte of flowering plants. Plant Cell *1*, 657-664.

Mascarenhas, J. (1990). Gene activity during pollen development. Annual Review of Plant Physiology and Plant Molecular Biology *41*, 317-338.

Mascarenhas, J. P. (1966). Pollen tube growth and ribonucleic acid synthesis by vegetative and generative nuclei of *Tradescantia*. . American Journal of Botany *53*, 563-569.

Mascarenhas, J. P. (1992). Pollen gene expression: molecular evidence. Int Rev Cytol *140*, 3-18.

Mascarenhas, J. P. (1993). Molecular Mechanisms of Pollen Tube Growth and Differentiation. Plant Cell *5*, 1303-1314.

Mascarenhas, J. P., and Mermeisten, J. (1981). Messenger RNAs: Their utilization and degradation during pollen germination and tube growth. Acta Soc Bot Polon *50*, 13-20.

McCormick, S. (2004). Control of male gametophyte development. Plant Cell *16*, S142-S153.

Messerli, M., and Robinson, K. R. (1997). Tip localized Ca2+ pulses are coincident with peak pulsatile growth rates in pollen tubes of Lilium longiflorum. J Cell Sci *110 (Pt 11)*, 1269-1278.

Messier, W., and Stewart, C. B. (1997). Episodic adaptive evolution of primate lysozymes. Nature *385*, 151-154.

Michaels, S. D., and Amasino, R. M. (1999). The gibberellic acid biosynthesis mutant ga1-3 of Arabidopsis thaliana is responsive to vernalization. Developmental Genetics *25*, 194-198.

Moore, R. C., and Purugganan, M. D. (2005). The evolutionary dynamics of plant duplicate genes. Current Opinion in Plant Biology *8*, 122-128.

Moore, R. C., and Purugganan, M. D. (2003). The early stages of duplicate gene evolution. Proceedings of the National Academy of Sciences of the United States of America *100*, 15682-15687.

Mouline, K., Very, A. A., Gaymard, F., Boucherez, J., Pilot, G., Devic, M., Bouchez, D., Thibaud, J. B., and Sentenac, H. (2002). Pollen tube development and competitive ability are impaired by disruption of a Shaker K+ channel in Arabidopsis. Genes & Development *16*, 339-350.

Mouradov, A., Glassick, T. V., Hamdorf, B. A., Murphy, L. C., Marla, S. S., Yang, Y. M., and Teasdale, R. D. (1998). Family of MADS-box genes expressed early in male and female reproductive structures of monterey pine. Plant Physiology *117*, 55-61.

Mueller, C. G. F., and Nordheim, A. (1991). A Protein Domain Conserved between Yeast Mcm1 and Human Srf Directs Ternary Complex-Formation. Embo Journal *10*, 4219-4229.

Munster, T., Pahnke, J., DiRosa, A., Kim, J. T., Martin, W., Saedler, H., and Theissen, G. (1997). Floral homeotic genes were recruited from homologous MADS-box genes preexisting in the common ancestor of ferns and seed plants. Proceedings of the National Academy of Sciences of the United States of America *94*, 2415-2420.

Munster, T., Faigl, W., Saedler, H., and Theissen, G. (2002). Evolutionary aspects of MADs-box genes in the Eusporangiate fern Ophioglossum. Plant Biol (Stuttg) *4*, 474-483.

Murgia, M., Charzynska, M., Rougier, M., and Cresti, M. (1991). Secretory tapetum of Brassica oleracea L: polarity and ultrastructural features. Sexual Plant Reproduction *4*, 28-35.

Murgia, M., Huang, B. Q., Tucker, S. C., and Musgrave, M. E. (1993). Embryo sac lacking antipodal cells in *Arabidopsis thaliana* (Brassicaceae). American Journal of Botany *80*, 824-838.

Muschietti, J., Dircks, L., Vancanneyt, G., and McCormick, S. (1994). LAT52 protein is essential for tomato pollen development: Pollen expressing antisense LAT52 RNA hydrates and germinates abnormally and cannot achieve fertilization. Plant Journal *6*, 321-338.

Muschietti, J., Eyal, Y., and McCormick, S. (1998). Pollen tube localization implies a role in pollen-pistil interactions for the tomato receptor-like protein kinases LePRK1 and LePRK2. Plant Cell *10*, 319-330.

Nagasawa, N., Miyoshi, M., Sano, Y., Satoh, H., Hirano, H., Sakai, H., and Nagato, Y. (2003). SUPERWOMAN and DROOPING LEAF genes control floral organ identity in rice. Development *130*, 705-718.

Nam, J., Kim, J., Lee, S., An, G. H., Ma, H., and Nei, M. S. (2004). Type I MADS-box genes have experienced faster birth-and-death evolution than type II MADS-box

genes in angiosperms. Proceedings of the National Academy of Sciences of the United States of America *101*, 1910-1915.

Nesi, N., Debeaujon, I., Jond, C., Stewart, A. J., Jenkins, G. I., Caboche, M., and Lepiniec, L. (2002). The TRANSPARENT TESTA16 locus encodes the ARABIDOPSIS BSISTER MADS domain protein and is required for proper development and pigmentation of the seed coat. Plant Cell *14*, 2463-2479.

Ng, M., and Yanofsky, M. F. (2001). Function and evolution of the plant MADS-box gene family. Nature Reviews Genetics *2*, 186-195.

Nickrent, D. L., Parkinson, C. L., Palmer, J. D., and Duff, R. J. (2000). Multigene phylogeny of land plants with special reference to bryophytes and the earliest land plants. Molecular Biology and Evolution *17*, 1885-1895.

Nishiyama, T., and Kato, M. (1999). Molecular phylogenetic analysis among bryophytes and tracheophytes based on combined data of plastid coded genes and the 18S rRNA gene. Molecular Biology and Evolution *16*, 1027-1036.

Norman, C., Runswick, M., Pollock, R., and Treisman, R. (1988). Isolation and Properties of Cdna Clones Encoding Srf, a Transcription Factor That Binds to the C-Fos Serum Response Element. Cell *55*, 989-1003.

Ohad, N., Margossian, L., Hsu, Y. C., Williams, C., Repetti, P., and Fischer, R. L. (1996). A mutation that allows endosperm development without fertilization. Proceedings of the National Academy of Sciences USA *93*, 5319-5324.

Palanivelu, R., Brass, L., Edlund, A. F., and Preuss, D. (2003). Pollen tube growth and guidance is regulated by POP2. an Arabidopsis gene that controls GABA levels. Cell *114*, 47-59.

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*, 1538-1551.

Parinov, S., Sevugan, M., Ye, D., Yang, W. C., Kumaran, M., and Sundaresan, V. (1999). Analysis of flanking sequences from dissociation insertion lines: a database for reverse genetics in Arabidopsis. Plant Cell *11*, 2263-2270.

Park, S. K., Howden, R., and Twell, D. (1998). The Arabidopsis thaliana gametophytic mutation gemini pollen1 disrupts microspore polarity, division asymmetry and pollen cell fate. Development *125*, 3789-3799.

Passmore, S., Maine, G. T., Elble, R., Christ, C., and Tye, B. K. (1988). Saccharomyces-Cerevisiae Protein Involved in Plasmid Maintenance Is Necessary for Mating of Mat-Alpha Cells. Journal of Molecular Biology *204*, 593-606.

Pelaz, S., Ditta, G. S., Baumann, E., Wisman, E., and Yanofsky, M. F. (2000). B and C floral organ identity functions require SEPALLATA MADS-box genes. Nature *405*, 200-203.

Pelaz, S., Tapia-Lopez, R., Alvarez-Buylla, E. R., and Yanofsky, M. F. (2001). Conversion of leaves into petals in Arabidopsis. Current Biology *11*, 182-184.

Perry, S. E., Nichols, K. W., and Fernandez, D. E. (1996). The MADS domain protein AGL15 localizes to the nucleus during early stages of seed development. Plant Cell *8*, 1977-1989.

Pierson, E. S., Miller, D. D., Callaham, D. A., van Aken, J., Hackett, G., and Hepler, P. K. (1996). Tip-localized calcium entry fluctuates during pollen tube growth. Dev Biol *174*, 160-173.

Pina, C., Pinto, F., Feijo, J. A., and Becker, J. D. (2005). Gene family analysis of the Arabidopsis pollen transcriptome reveals biological implications for cell growth, division control, and gene expression regulation. Plant Physiology *138*, 744-756.

Portereiko, M. F., Lloyd, A., Steffen, J. G., Punwani, J. A., Otsuga, D., and Drews, G. N. (2006). AGL80 is required for central cell and endosperm development in Arabidopsis. Plant Cell *18*, 1862-1872.

Purugganan, M. D. (1997). The MADS-box floral homeotic gene lineages predate the origin of seed plants: phylogenetic and molecular clock estimates. J Mol Evol *45*, 392-396.

Purugganan, M. D. (1998). The molecular evolution of development. Bioessays *20*, 700-711.

Quigley, D. S. H. a. M. (1981). A Rapid Boiling method for the preparation of bacterial plasmids. Analytical biochemistry *114*, 193 - 197.

Ratcliffe, O. J., Riechmann, J. L., and Zhang, J. Z. (2000). INTERFASCICULAR FIBERLESS1 is the same gene as REVOLUTA. Plant Cell *12*, 315-317.

Ratcliffe, O. J., G.C.Nadzan, T.L.Reuber, and J.L.Riechmann (2001). Regulation of Flowering in Arabidopsis by an FLC Homologue. Plant Physiology *126*, 122-132.

Reeves, P. H., and Coupland, G. (2000). Response of plant development to environment: control of flowering by daylength and temperature. Current Opinion in Plant Biology *3*, 37-42.

Reiser, L., Modrusan, Z., Margossian, L., Samach, A., Ohad, N., Haughn, G. W., and Fischer, R. L. (1995). The Bell1 Gene Encodes a Homeodomain Protein Involved in Pattern-Formation in the Arabidopsis Ovule Primordium. Cell *83*, 735-742.

Riechmann, J. L., Krizek, B. A., and Meyerowitz, E. M. (1996). Dimerization specificity of Arabidopsis MADS domain homeotic proteins APETALA1,

APETALA3, PISTILLATA, and AGAMOUS. Proceedings of the National Academy of Sciences of the United States of America *93*, 4793-4798.

Riechmann, J. L., and Meyerowitz, E. M. (1997). Determination of floral organ identity by Arabidopsis MADS domain homeotic proteins AP1, AP3, PI, and AG is independent of their DNA-binding specificity. Molecular Biology of the Cell *8*, 1243-1259.

Rotman, N., Rozier, F., Boavida, L., Dumas, C., Berger, F., and Faure, J. E. (2003). Female control of male gamete delivery during fertilization in Arabidopsis thaliana. Current Biology *13*, 432-436.

Rounsley, S. D., Ditta, G. S., and Yanofsky, M. F. (1995). Diverse Roles for Mads Box Genes in Arabidopsis Development. Plant Cell *7*, 1259-1269.

Russell, S. D. (1992). Double fertilization. Internal Review in Cytology *140*, 357-388.

Russell, S. D. (1993). The egg cell: development and role in fertilization and early embryogenesis. Plant Cell *5*.

Samach, A., Onouchi, H., Gold, S. E., Ditta, G. S., Schwarz-Sommer, Z., Yanofsky, M. F., and Coupland, G. (2000). Distinct roles of CONSTANS target genes in reproductive development of Arabidopsis. Science *288*, 1613-1616.

Samaj, J., Muller, J., Beck, M., Bohm, N., and Menzel, D. (2006). Vesicular trafficking cytoskeleton and signalling in root hairs and pollen tubes. Trends in Plant Science *11*.

Samaj, J. e. a. (2004). Endocytosis, actin skeleton, and signalling. Plant Physiology *135*, 1150-1161.

Sambrook J., R. D. W. (2001). Molecular Cloning A LABORATORY MANUAL. New York, Cold Spring Harbor Laboratory Press.

Samson, F., Brunaud, V., Balzergue, S., Dubreucq, B., Lepiniec, L., Pelletier, G., Caboche, M., and Lecharny, A. (2002). FLAGdb/FST: a database of mapped flanking insertion sites (FSTs) of Arabidopsis thaliana T-DNA transformants. Nucleic Acids Res *30*, 94-97.

Sanders, P., Bui, A., and Weterings, K. (1999). Anther developmental defects in Arabidopsis thaliana male sterile mutants. Sexual Plant Reproduction *11*, 297-322.

Schoof, H., Lenhard, M., Haecker, A., Mayer, K., 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*, 635-644.

Schrauwen, J. A. M., Degroot, P. F. M., Vanherpen, M. M. A., Vanderlee, T., Reynen, W. H., Weterings, K. A. P., and Wullems, G. J. (1990). Stage-Related Expression of Messenger-Rnas During Pollen Development in Lily and Tobacco. Planta *182*, 298-304.

Sheldon, C. C., Rouse, D. T., Finnegan, E. J., W.J.Peacock, and Dennis, E. S. (1999). The molecular basis of vernalization: The central role of FLOWERING LOCUS C (FLC). Proceedings of the National Academy of Sciences of the United States of America *97*, 3753-3758.

Shimizu, K. K., and Okada, K. (2000). Attractive and repulsive interactions between female and male gametophytes in *Arabidopsis* pollen tube guidance. Development *127*, 4511-4518.

Shiu, S. H., Shih, M. C., and Li, W. H. (2005). Transcription factor families have much higher expansion rates in plants than in animals. Plant Physiol *139*, 18-26.

Shivanna, K. R., Jaiswal, V. S., and Ram, H. Y. M. (1974). Inhibition of gamete formation by cycloheximide in pollen tubes of *Impatiens balsamina*. Planta *117*, 173-177.

Shore, P., and Sharrocks, A. D. (1995). The MADS-box family of transcription factors. Eur J Biochem *229*, 1-13.

Sieburth, L. E., and Meyerowitz, E. M. (1997). Molecular dissection of the AGAMOUS control region shows that cis elements for spatial regulation are located intragenically. The Plant Cell *19*, 355-365.

Simon, R., Igeno, M. I., and Coupland, G. (1996). Activation of floral meristem identity genes in Arabidopsis. Nature *384*, 59-62.

Singh, K. B. (1998). Transcriptional regulation in plants: the importance of combinatorial control. Plant Physiology *118*, 1111-1120.

Skinner, D. J., Hill, T. A., and Gasser, C. S. (2004). Regulation of ovule development. Plant Cell *16*, S32-S45.

Skipper, M., Pedersen, L. B., Frederiksen, S., and Johansen, B. B. (2006). Cloning and transcription analysis of an AGAMOUS- and SEEDSTICK ortholog in the orchid Dendrobium thyrsiflorum (Reichb. f.). Gene *366*, 266-274.

Smyth, D. R., Bowman, J. L., and Meyerowitz, E. M. (1990). Early Flower Development in Arabidopsis. Plant Cell *2*, 755-767.

Somerville, C., and Somerville, S. (1999). Plant functional genomics. Science *285*, 380-383.

Sommer, H., Beltran, J. P., Huijser, P., Pape, H., Lonnig, W. E., Saedler, H., and Schwarz-Sommer, Z. (1990). Deficiens, a homeotic gene involved in the control of flower morphogenesis in Antirrhinum majus: the protein shows homology to transcription factors. Embo J *9*, 605-613.

Song, X. F., Yang, C. Y., Liu, J., and Yang, W. C. (2006). RPA, a class II ARFGAP protein, activates ARF1 and U5 and plays a role in root hair development in *Arabidopsis*. Plant Physiology *141*, 966-976.

Speulman, E., Metz, P. L., van Arkel, G., te Lintel Hekkert, B., Stiekema, W. J., and Pereira, A. (1999). A two-component enhancer-inhibitor transposon mutagenesis system for functional analysis of the Arabidopsis genome. Plant Cell *11*, 1853-1866.

Sundstrom, J., Carlsbecker, A., Svensson, M. E., Svenson, M., Johanson, U., Theissen, G., and Engstrom, P. (1999). MADS-box genes active in developing pollen cones of Norway spruce (Picea abies) are homologous to the B-class floral homeotic genes in angiosperms. Developmental Genetics *25*, 253-266.

Svensson, M. E., and Engstrom, P. (2002). Closely related MADS-box genes in club moss (Lycopodium) show broad expression patterns and are structurally similar to, but phylogenetically distinct from, typical seed plant MADS-box genes. New Phytologist *154*, 439-450.

Svensson, M. E., Johannesson, H., and Engstrom, P. (2000). The LAMB1 gene from the clubmoss, Lycopodium annotinum, is a divergent MADS-box gene, expressed specifically in sporogenic structures. Gene *253*, 31-43.

Sweetman, J., Spurr, C., Eliasson, A., Gass, N., Steinmetz, A., and Twell, D. (2000). Isolation and characterization of two pollen-specific LIM domain protein cDNAs from *Nicotiana tabacum*. Sexual Plant Reproduction *12*, 339-345.

Tanabe, Y., Hasebe, M., Sekimoto, H., Nishiyama, T., Kitani, M., Henschel, K., Munster, T., Theissen, G., Nozaki, H., and Ito, M. (2005). Characterization of MADS-box genes in charophycean green algae and its implication for the evolution of MADS-box genes. Proceedings of the National Academy of Sciences of the United States of America *102*, 2436-2441.

Tanabe, Y., Uchida, M., Hasebe, M., and Ito, M. (2003). Characterization of the Selaginella remotifolia MADS-box gene. Journal of Plant Research *116*, 71-75.

Tanaka, I. (1997). Differentiation of generative and vegetative cells in angiosperm pollen. Sexual Plant Reproduction *10*, 1-7.

Tandre, K., Albert, V. A., Sundas, A., and Engstrom, P. (1995). Conifer homologues to genes that control floral development in angiosperms. Plant Molecular Biology *27*, 69-78.

Taylor, L. P., and Hepler, P. K. (1997). Pollen germination and tube growth. Annual Review of Plant Physiology and Plant Molecular Biology *48*, 461-491.

Taylor, P. E., Glover, J. A., Lavithis, M., Craig, S., Singh, M. B., Knox, R. B., Dennis, E. S., and Chaudhury, A. M. (1998). Genetic control of male fertility in Arabidopsis thaliana: structural analyses of postmeiotic developmental mutants. Planta *205*, 492-505.

Theissen, G., and Saedler, H. (1995). Mads-Box Genes in Plant Ontogeny and Phylogeny - Haeckels Biogenetic Law Revisited. Current Opinion in Genetics & Development *5*, 628-639.

Theissen, G., and Saedler, H. (2001). Plant biology. Floral quartets. Nature *409*, 469-471.

Theissen, G., Becker, A., Di Rosa, A., Kanno, A., Kim, J. T., Munster, T., Winter, K. U., and Saedler, H. (2000). A short history of MADS-box genes in plants. Plant Molecular Biology *42*, 115-149.

Theissen, G., and Saedler, H. (2001). Plant biology. Floral quartets. Nature *409*, 469-471.

Thiessen, G., Munster, T., and Henschel, K. (2001). Why don't mosses flower? New Phytologist *150*, 1-5.

Tissier, A. F., Marillonnet, S., Klimyuk, V., Patel, K., Torres, M. A., Murphy, G., and Jones, J. D. (1999). Multiple independent defective suppressor-mutator transposon insertions in Arabidopsis: a tool for functional genomics. Plant Cell *11*, 1841-1852.

Trobner, W., Ramirez, L., Motte, P., Hue, I., Huijser, P., Lonnig, W. E., Saedler, H., Sommer, H., and Schwarz-Sommer, Z. (1992). GLOBOSA: a homeotic gene which interacts with DEFICIENS in the control of Antirrhinum floral organogenesis. Embo J *11*, 4693-4704.

Twell, D. (1992). Use of a Nuclear-Targeted Beta-Glucuronidase Fusion Protein to Demonstrate Vegetative Cell-Specific Gene-Expression in Developing Pollen. Plant Journal *2*, 887-892.

Twell, D. (1999). Mechanisms of microspore polarity and differential cell fate determination in developing pollen, in *Fertilization in Higher Plants. Molecular and Cytological Aspects*, Vol 201-215 (Berlin: Springer-Verlag).

Twell, D. (2002). The developmental biology of pollen. Plant Reproduction *6*, 86-153.

Twell, D., and Howden, R. (1998a). Cell polarity, asymmetric division and cell fate determination in developing pollen, in *Reproductive Biology* (Kew: Royal Botanic Gardens).

Twell, D., and Howden, R. (1998b). Mechanisms of asymmetric division and cell fate determination in developing pollen, in *Androgenesis* and *Haploid Plants, 967-1994. In memory of Jean-Pierre Bourgin* (Berlin: INRA-Springer-Verlag).

Tzeng, T.-Y., Chen, H.-Y., and Yang, C.-H. (2002). Ectopic expression of carpel-specific MADS Box genes from Lily and Lisianthus causes similar homeotic conversion of sepal and petal in *Arabidopsis*. Plant Physiology *130*, 1827-1836.

van der Meer, I. M., Stam, M. E., van Tunen, A. J., Mol, J. N., and Stuitje, A. R. (1992). Antisense inhibition of flavonoid biosynthesis in petunia anthers results in male sterility. Plant Cell *4*, 253-262.

Verelst, W., Saedler, H., and Munster, T. (2007). MIKC* MADS-protein complexes bind motifs enriched in the proximal region of late pollen-specific Arabidopsis promoters. Plant Physiology *143*, 447-460.

Wagner, V. T., Cresti, M., Salvatici, P., and Tiezzi, A. (1990). Changes in volume, surface area, and frequency of nuclear pores on the vegetative nucleus of tobacco pollen in fresh, hydrated, and activated conditions. Planta *181*, 304-309.

Wagner, D., Sablowski, R. W., and Meyerowitz, E. M. (1999). Transcriptional activation of APETALA1 by LEAFY. Science *285*, 582-584.

Waterhouse, P. M., Graham, M. W., and Wang, M. B. (1998). Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. Proc Natl Acad Sci U S A *95*, 13959-13964.

Weigel, D., Alvarez, J., Smyth, D. R., Yanofsky, M. F., and Meyerowitz, E. M. (1992). LEAFY controls floral meristem identity in Arabidopsis. Cell *69*, 843-859.

Weigel, D., Ahn, J. H., Blazquez, M. A., Borevitz, J. O., Christensen, S. K., Fankhauser, C., Ferrandiz, C., Kardailsky, I., Malancharuvil, E. J., Neff, M. M., *et al.* (2000). Activation tagging in Arabidopsis. Plant Physiol *122*, 1003-1013.

Weterings, K., Reijnen, W., van Aarssen, R., Kortstee, A., Spijkers, J., van Herpen, M., Schrauwen, J., and Wullems, G. (1992). Characterization of a pollen-specific cDNA clone from Nicotiana tabacum expressed during microgametogenesis and germination. Plant Mol Biol *18*, 1101-1111.

Weterings, K., and Russell, S. D. (2004). Experimental analysis of the fertilization process. Plant Cell *16 (suppl.)*, S107-S118.

Whipple, C. J., Ciceri, P., Padilla, C. M., Ambrose, B. A., Bandong, S. L., and Schmidt, R. J. (2004). Conservation of B-class floral homeotic gene function between maize and *Arabidopsis*. Development *131*, 6083-6091.

Winter, K. U., Becker, A., Munster, T., Kim, J. T., Saedler, H., and Theissen, G. (1999). MADS-box genes reveal that gnetophytes are more closely related to conifers than to flowering plants. Proceedings of the National Academy of Sciences of the United States of America *96*, 7342-7347.

Wittink, F. R. A., Knuiman, B., Derksen, J., Capkova, V., Twell, D., Schrauwen, J. A. M., and Wullems, G. J. (2000). The pollen-specific gene *Ntp303* encodes a 69-kDa glycoprotein associated with the vegetative membranes and the cell wall. sex Plant Reproduction *12*, 276-284.

Yadegari, R., and Drews, G. N. (2004). Female gametophyte development. Plant Cell *16*, S133-S141.

Yang, S., Sweetman, J., and Amirsadeshi, S. (2001). Novel anther-specific myb genes from tobacco as putative regulators of phenylalanine ammonia lyase expression. Plant Physiology *126*, 1738-1753.

Yang, W. C., and Sundaresan, V. (2000). Genetics of gametophyte biogenesis in Arabidopsis. Current Opinion in Plant Biology *3*, 53-57.

Yang, W. C., Ye, D., Xu, J., and Sundaresan, V. (1999). The SPOROCYTELESS gene of Arabidopsis is required for initiation of sporogenesis and encodes a novel nuclear protein. Genes & Development *13*, 2108-2117.

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*, 35-39.

Zachgo, S., Saedler, H., and Schwarz-Sommer, Z. (1997). Pollen-specific expression of DEFH125, a MADS-box transcription factor in Antirrhinum with unusual features. Plant J *11*, 1043-1050.

Zhang, H. M., and Forde, B. G. (1998). An Arabidopsis MADS box gene that controls nutrient-induced changes in root architecture. Science *279*, 407-409.

The Arabidopsis Information Resource <http://www.arabidopsis.org/>

SALK Institute <www.signal.salk.edu/>