Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Functional characterization of two plant type I MADS-box genes in *Arabidopsis thaliana* – *AGL40* and *AGL62*

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

Ryohei Kaji

2008

Abstract

MADS-box transcription factors (TF) are a family of evolutionary conserved genes found across various eukaryotic species. Characterized by the conserved DNA binding MADS-box domain, MADS-box TF has been shown to play various roles in developmental processes. MADS-box genes can be based on MADS-box structural motifs divided into type I and type II lineages. In plants very limited functional characterization have been achieved with type I genes MADS-box genes.

In this project we attempted to functionally characterize 2 closely related members of the type I lineage MADS-box genes *AGL40* and *AGL62* and give further support to the hypothesis that plant type I MADS-box genes are also crucial to normal plant development. Based on our expression domain characterization assay using AGL62: GUS fusion construct, we have shown expression of *AGL62* in various tissues but especially strong in developing seeds, pollen and seedling roots and shoots. The web based microarray data suggesting that AGL62 may have a function in seed, pollen and seedling development backed up this result.

Interestingly when we carried out PCR based genotyping with segregating population of heterozygous *AGL62* T-DNA insertion lines (*agl62/+*) to identify the homozygous T-DNA insertion lines we detected no homozygous T-DNA insertion line indicating loss-of-function of AGL62 may be lethal to plant.

With reference to the *AGL62* expression in pollen, seed and seedling root and shoot, we carried out phenotypic assay on each of these tissues in *agl62/+* background to investigate whether there was any phenotypic defect observed. Significant reduction in number of seeds was observed in *agl62/+* indicating possible role of AGL62 in seed development. Our microscopic observation of seeds from *agl62/+* plants showed defective embryos and confirmed that AGL62 plays a role in seed development.

Our data on AGL62 is the first report that confirms AGL62's involvement in plant development and can be a ground work for further works on functional characterization of other members of plant type I MADS-box genes.

Acknowledgements

First and foremost I would like to thank my supervisor Dr Barbara Ambrose for her help, support and encouragements. I especially appreciate your patience with me for the times when we encountered a problem. Without your support I would not have been able to finish this thesis. I have learnt so many things from your advices both in and out of the laboratory. I will always be grateful for the times I spent in MADS-house laboratory.

I would also like to say thanks to the people in MADS-house laboratory both present and past. Thank you very much Kalika for your skills, knowledge and your dedication. You have taught me many things including discipline. Thank you Arti for being very nice to me and being my friend in the lab. I enjoyed working with you. I thank Xiuwen for her technical supports. You have been very nice teacher and taught me many skills that I would have otherwise not learned. Thank you very much Emilio for taking care of my plants all the time. Your support greatly helped finish my experimental part of the project. I thank all of you again for making the times I spent at MADS-house such an exciting and enjoyable moments.

To my family, mom, dad and my brother, thanks for all the patience and encouragement that you gave me. I really appreciate your support. Finally to my brothers and sisters in church, thanks for all the love, support and discipline you gave me. You guys have been really close friends for all these times and I am looking forward to spending more time with you guys. My special thanks to Gary for disciplining me all these times and leading me. I really enjoy having time together with you.

Table of contents

Chapter 1: Introduction

Chapter 2: Materials and Methods

Chapter 3: AGL62::GUS temporal & spatial expression pattern

Abbreviations

AGL	AGAMOUS-LIKE X
Agrobacterium	*Agrobacterium tumefaciens*
Arabidopsis	*Arabidopsis thaliana*
bp	base pair
°C	degrees Celsius
cDNA	complementary DNA
DAF	day(s) after flowering
DAP	days(s) after plating
Da	Dalton
DNA	deoxyribonucleic acid
dNTP	deoxy-nucleotide-triphosphate
E. coli	*Eschrchia coli*
EDTA	ethylene diamine tetra acetate
GFP	green florescent protein
GUS	β-glucuronidase
Kb	kilo base
L	liter
LB	Luria-Bertani
MOPS	3-(N-morpholino) propanesulfonic acid
MQ	Milli-Q
MS	Murashige & Skoog
mg	milligram
min	minute
ml	milliliter
mmol	millimol
ng	nanogram
nmol	nanomol
PCR	polymerase chain reaction
pmol	picomol

RNA	ribonucleic acid
Rpm	revolutions per minute
RT-PCR	reverse transcriptase PCR
T1	first generation after transformation
T-DNA	transfer DNA
TE	Tris-EDTA
Tris	tris (hydroxymethyl) aminomethane
µg	microgram
µl	microliter
µm	micrometer
µmol	micromole
UV	ultraviolet
V	volts
WT	wild type
X-gluc	5-bromo 4-chloro 3-indolyl glucuronide

List of figures

List of tables

Chapter 1 Introduction

1.1 MADS-box genes

MADS-box genes are an evolutionary conserved gene family found in a range of organisms from yeast to humans to plants and have a role in regulation of various developments (Baum, 1998; Vergara-Silva et al., 2000; Kofuji et al., 2003). MADS-box genes code for transcription factors (TF) and named so after an acronym from *MINICHROMOSOME MAINTENANCE 1* (*MCM1*), *AGAMOUS* (*AG*), *DEFICIENS* (*DEF*) and *SERUM RESPONSE FACTOR* (*SRF*) which are the first genes identified in the family (De Bodt et al., 2003b; De Bodt et al., 2003a). Changes in expression pattern of TF genes usually result in changes of expression pattern of downstream genes and may cause global gene switch in some cases. MADS-box genes are shown to have important roles in terms of evolution of body plans as well as various other developmental processes (Becker and Theissen, 2003).

MADS-box genes can be divided into 2 types – type I and type II – based on their structural motifs of the conserved DNA binding MADS-box domain (Figure 1.1). The conserved MADS-domain characteristic of MADS-box genes is known to bind DNA via the CarG box sequence (CC(A/T)6GG) (Johansen et al., 2002; De Bodt et al., 2003b). In type II, the approximately 58 amino acid long MADS-box domains have a structural motif similar to that of *MEF2-like* genes in animals (Figure 1.1). In plants, type II MADS-box genes have been extensively studied which showed they play various important roles in plant development (De Bodt et al., 2003a). Type I MADS-box genes have the SRF-like MADS-box domain at their N-terminus (Figure 1.1). In plants, little is known about the role type I MADS-box genes play in development compared to well studied type II lineage (De Bodt et al., 2003a). Both type I and II MADS-box genes are present in fungal, animal and plant kingdom and phylogenetic analysis of MADS-box genes in all the 3 kingdoms showed that MADS-box genes were already present in last common ancestor of animal, fungal and plant species (Martinez-Castilla and Alvarez-Buylla, 2004). Animal/Fungal MADS-box genes also have either SRF-like MADS-box domain or MEF2-like MADS-box domain respectively and have been reported to play important roles in developmental processes (De Bodt et al., 2003a). (Figure 1.1).

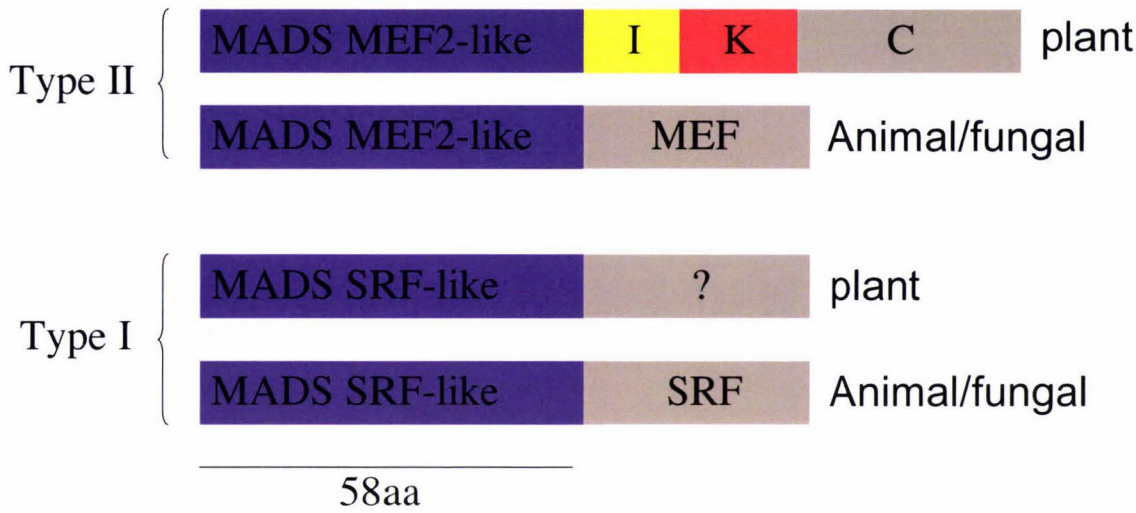


Figure 1.1 MADS-box genes in fungal, animal and plant species. Both type I and type II lineages were identified in all 3 kingdoms. Type I lineages are defined by their conserved SRF like MADS-box DNA binding domain while type II lineages are defined by their MEF-2 like MADS-box domain. All MADS-box genes also have less conserved C-terminal domain. In plants, function of C-terminal domain of type I MADS-box genes are yet unidentified and therefore denoted here as? There are extra I and K domains between the MADS-box domain and C-terminal domain in plant type II MADS-box genes.

1.2 *Arabidopsis thaliana* as plant model organism

Arabidopsis thaliana is a small flowering plant that belongs to a Brassicaceae (mustard) family (Figure 1.2), which includes many of the cultivated species such as cabbage. Although *Arabidopsis* has no major agronomic value, it has been used as a plant model organism for more than a decade for several reasons. *Arabidopsis* has a small size and relatively short life cycle of approximately 6weeks making it easy for researchers to culture large number of them in restricted space (http://www.arabidopsis.org/). It is also amenable to transformation by *Agrobacterium* resulting in production of many mutant lines (Clough and Bent, 1998). In 2000, whole *Arabidopsis* genome has been sequenced leading to identification of many new genes (Arabidopsis Genome, 2000). Currently, it is thought 125Mb *Arabidopsis* genome contains approximately 25000 genes, of which 5% encodes a transcription factor (TF) (Arabidopsis Genome, 2000). There are 11 TF family present in *Arabidopsis* and one of the family MADS-box family is thought to be homologous to animal Homeo box genes in terms of their functional roles in development (De Bodt et al., 2003a). Since the genome sequencing in 2000, 108 MADS-

box genes have been identified in *Arabidopsis* (Martinez-Castilla and Alvarez-Buylla, 2004). Examples of MADS-box genes include *AGAMOUS* (*AG*) and *APETALLA1* (*AP1*) which are the floral organ identity genes (Becker and Theissen, 2003). In *Arabidopsis*, all functionally uncharacterized MADS-box genes are named *AGAMOUS-LIKE* (*AGL*) X after *AG* –the first plant MADS-box genes to have its function characterized (De Bodt et al., 2003a). Since there are much less MADS-box genes present in animal and fungal species, much of the information about function of MADS-box genes so far came from studies in *Arabidopsis* (De Bodt et al., 2003a). Coupled with its genome sequence data, availability to many mutant lines and amenability to many biochemical and molecular biological assays, *Arabidopsis thaliana* therefore works as an ideal plant organism to be studying function of MADS-box genes.

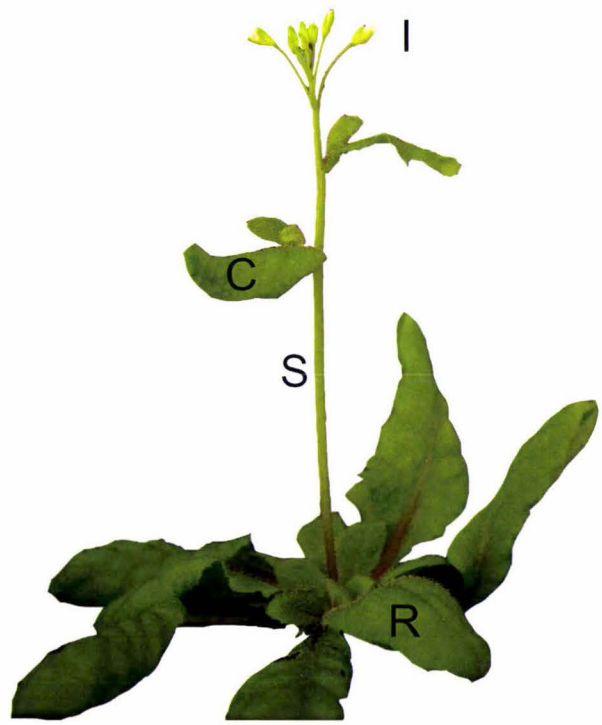


Figure 1.2 Basic body plan of *Arabidopsis thaliana*. 3 weeks old *Arabidopsis* Columbia Wild Type (WT). Arabidopsis plant is a small plant but consists of all major plant organs including root (not shown), rosetta leaves (R), stem (S), cauline leaves (C), and inflorescence (I).

1.3 Type II plant MADS-box genes

Plant type II MADS-box genes have been extensively studied and found to be involved in various aspects of plant development. (Johansen et al., 2002; De Bodt et al., 2003b). In plant type II MADS-box genes, there are additional C-terminal domains including K, I, and C–domains following the conserved N-terminal MADS-box domain (fig 1.1). The K (keratin-like) domain is well conserved between plant type II MADS-box genes and consists of approximately 80 amino acids that form a hydrophobic coiled-coil structure involved in protein – protein interaction to form dimer or multimer with other TFs (Yang et al., 2003; Yang and Jack, 2004). The I (intervening) domain is less conserved compared to MADS or K domain and is involved in specification of protein-protein interaction (Yang and Jack, 2004). The least conserved C-terminal domain is thought to be involved in Transactivation of target genes similar to other clades of MADS-box genes (Yang and Jack, 2004). Because of these MADS-box, I, K and C domains, plant type II MADS-box genes are also known as MIKC type MADS-box genes (Martinez-Castillo and Alvarez-Buylla, 2004). The I, K and C domains are only present in plant type II MADS-box genes indicating I, K and C domains have developed specifically in plant type II lineages after diversion of type I and type II lineages in plants (Figure 1.1) (Martinez-Castilla and Alvarez-Buylla, 2004).

1.3.1 Plant type II MADS-box genes and flower development

Type II MADS-box genes are particularly important in flower development although not restricted to it (Irish, 2003). Currently, there are over 250,000 known angiosperm species (flowering plants) present all over the world and the flowers come in a variety of sizes, shapes, and colors (Irish, 2003). However, despite these diversities in floral appearance, their basic organization remains the same in many flowers (Irish, 2003). Flowers arise from a floral meristem that gives rise to 4 floral organs – sepal, petal, stamen (male reproductive organ) and carpel (female reproductive organ). These floral organs arise in an orderly manner around the floral meristem axis (Lawton-Rauh et al., 2000). Sepals are formed first followed by petals and then stamens then carpels (Lawton-Rauh et al., 2000). Analysis of homeotic floral mutants which had floral organ identity miss-specification phenotype led to the identification of several type II plant MADS-box genes that are classified into 3 subclasses – A, B, and C class genes. *APETALA1* (*AP1*) represents A

class gene, B class genes include *PISTILLATA (PI)* and *APETALA3 (AP3)* and *AGAMOUS* (*AG*) represents C class gene. This in turn led to the formulation of a genetic model termed the ABC model that explains how a combination of A, B, and C classes of type II MADS-box genes are involved in specification of identity of 4 whorl organs of the flower – sepal, petal, stamen, and carpel (Figure 1.3) (Pelaz et al., 2001; Irish, 2003). In the ABC model, A class genes alone specify sepals, A and B class genes together specify petals, B and C class genes act together to specify stamens and C class genes alone specify carpels (fig. 1.3) (Purugganan et al., 1995; Lawton-Rauh et al., 2000; Pelaz et al., 2001). These ABC class genes are mutually exclusive and are restricting expression of each other in wild type (WT) to regulate their expression domains (Purugganan et al., 1995). Therefore, when one class of genes are knocked out in floral homeotic mutants, the other class of genes are over-expressed and result in conversion of one organ to another (homeotic mutation) (Figure 1.3).

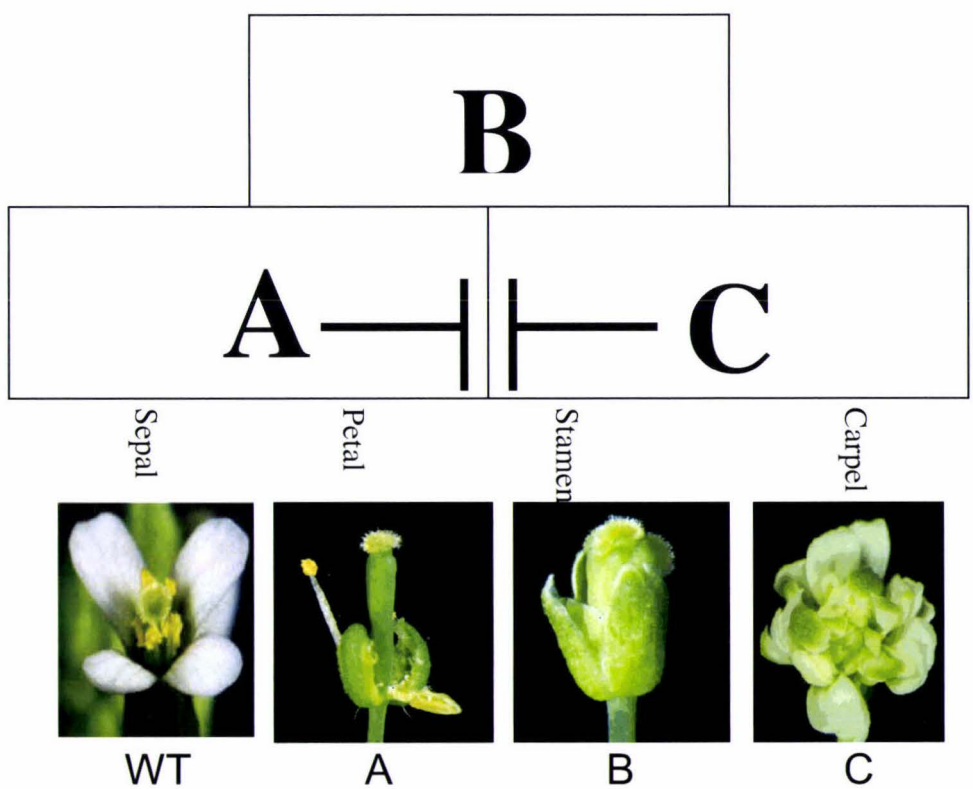


Figure 1.3 ABC floral organ identity model. Photos taken from www.Arabidopsis.org/ In ABC floral organ identity model, A class gene alone specify sepal, A and B class genes together specify petal, B and C class genes specify stamen and C class gene alone specify carpel (Top diagram). WT flowers have all 4 floral organs –petal, sepal, stamen and carpel (Bottom Figure WT). On the other hand, A class mutants lack sepal and petal (Figure A), B class mutant lack petal and stamen (Figure B) and C class mutant lacks stamen and carpel (Figure C)

1.3.2 Plant type II MADS-box genes outside the flower

In addition to floral organ identity, type II MADS-box genes have regulatory roles in flowering time of plants (Rounsley et al., 1995). The flowering pathway in plant requires quite a complex network of gene interactions. Several plant types II MADS-box genes have been isolated whose mutations resulted in altered flowering time. These genes include *SOC (SUPRESSOR OF CONSTANS), AGL24 (AGAMOUS-LIKE 24)* and *FLC (FLOWERING LOCUS C)* (Michaels et al., 2003). Isolation of such genes gave strong evidence that type II MADS-box genes are involved in control of flowering time as well.

The importance of type II MADS-box genes in plant development is not restricted to floral organ specification (Rounsley et al., 1995; Alvarez-Buylla, 2001; Ng and Yanofsky, 2001). As well as there regulatory role in reproductive organ development, MADS-box genes play important roles in development of vegetative tissues such as roots, stems, and leaves (Alvarez-Buylla, 2001). Hence, the MADS-box gene family has diverse roles throughout plant development. For example, a PCR based cloning approach resulted in the identification of *AGL16 (AGAMOUS-LIKE 16), AGL18 (AGAMOUS-LIKE 18)*, and *AGL19 (AGAMOUS-LIKE19)* and yeast-2-hybrid system identified *AGL27 (AGAMOUS-LIKE 27)* and *AGL31 (AGAMOUS-LIKE 31)* that are closely related to *FLC* (Alvarez-Buylla, 2001). RT-PCR showed that *AGL16* is highly expressed in rosette leaves and moderately expressed in roots and stems (Alvarez-Buylla, 2001). Further, *in situ* hybridization showed that *AGL16* is expressed in guard cells and trichomes in both abaxial and adaxial epidermis of rosette leaves (Alvarez-Buylla, 2001). These experiments indicated that *AGL16* may have a role in the regulation of stomata development (Alvarez-Buylla, 2001). Evolution of stomata was one of a key events during the early evolution of land plants which is consistent with the idea that MADS-box genes played an important role in the evolution of body plans in plants (Baum, 1998; Kellogg, 2004). *AGL19* is specifically expressed in root and not in any other tissues indicating a role of *AGL19* in root development (Alvarez-Buylla, 2001). Within root, this gene is expressed in columella, lateral root cap, and epidermal cells of the meristematic regions (Alvarez-Buylla, 2001). The root cap plays a central role in perception of environmental cues such as gravity and is important for gravitropic responses in plant.

AGL27 and *AGL31* are similar in sequence and expression pattern to *FLC* and are expressed in most of the plant tissues including root, leaf, stem, flower and in silique (Alvarez-Buylla, 2001). Similarity of *AGL27* and *AGL31* in sequence and expression pattern to *FLC* suggest that theses genes may have redundant activities to each other indicating *AGL27* and *AGL31* may have a function in controlling flowering time as well. However, *FLC* also seems to have at least some independent roles because single *flc* loss-of-function mutant shows a clear early flowering phenotype which is not observed in *agl27* nor *agl31* single mutants (Michaels et al., 2003). Other examples of MADS-box genes that have function in development of plant outside the flower include *SHATTERPROOF1 (SHP1), SHATTERPROOF2 (SHP2), FRUITFUL (FUL)*, and *TRANSPARENT TESTA16 (TT16)* that are involved in seed development (Rounsley et al., 1995; Alvarez-Buylla, 2001).

1.3.3 Plant type II MADS-box genes in other plant species

Completion of genome sequencing in *Arabidopsis thaliana* in 2000 identified approximately 26000 genes in this plant model organism of which just over 100 of the identified genes are MADS-box genes (Kofuji et al., 2003; Parenicova et al., 2003; Martinez-Castilla and Alvarez-Buylla, 2004). Because of a high level of conservation in the DNA binding MADS-box domain, isolation of MADS-box genes from other distantly related plant species such as maize, rice, and orchid plants were possible by using degenerate primers that bind to the conserved MADS-box domain at N-terminal (Mena et al., 1995; Becker et al., 2000; Jia et al., 2000b). Identification of MADS-box genes in these monocot plants indicated that MADS-box genes were already present before the divergence of monocots and dicots (Jia et al., 2000a). Fact that the ABC model of floral organ identity is conserved between monocots and dicots supports the above statement (Ambrose et al., 2000). In line with the ancestral presence of MADS-box genes in plants, MADS-box genes have also been identified in lower plant species such as moss (*P. patens*), fern, and gymnosperms (*G. gnemon*) (Winter et al., 1999; Henschel et al., 2002; Theissen and Becker, 2004). MADS-box genes are also functional in monocot plant development. This is of an agricultural importance as well as the scientific importance because important cereals including rice plant belong to the monocot (Jia et al., 2000b). Rice (*Oryza sativa)* genome sequencing has been completed recently and this enabled the

comparative genomics approach to compare MADS-box genes from dicots (*Arabidopsis thaliana*) to that of monocots (*Oryza sativa*) (Jia et al., 2000b). This resulted in the identification of 71 MADS-box genes in *Oryza sativa* but the actual number of MADS-box genes in rice is thought to be higher as genome annotation in rice is far from completion at this stage (Jia et al., 2000a). Examples of MADS-box genes identified in rice plant include *FDRMADS6* and *FDRMADS7* which showed high homology with *AP1* of *Arabidopsis thaliana* and *AP1* orthologues from other plant species (Jia et al., 2000a). FDRMADS6 protein showed 62% identity with AP1 in its first 154 amino acids while FDMADS7 protein showed 52% identity with AP1 in its first 231 amino acids (Jia et al., 2000a). Because *FSMADS6* and *FSMADS7* have high a level of similarity with *AP1*, an A class MADS-box gene involved in specification of floral organs, they are likely to have a function in flower development as well. Expression pattern analysis by RT PCR showed *FSMADS6* is exclusively expressed in inflorescence and no signal was detected in the vegetative tissues including root, leaf or stem (Jia et al., 2000a). On the other hand, *FSMADS7* is expressed mainly in inflorescence but weaker expression was also observed in root and shoot tissues as well (Jia et al., 2000a). Because MADS-box genes tend to be expressed in a cell where they function, this result suggested that MADS-box genes may have functions outside the flower as well in monocots such as *Oryza sativa (Jia et al., 2000b)*.

1.4 Animal/Fungal type I MADS-box genes

In plants, only little is accomplished in terms of functional characterization of type I MADS-box gene. However, in animal and fungal kingdoms, type I MADS-box have been shown to play various important roles in development (Treisman and Ammerer, 1992). Examples of animal and fungal type I MADS-box genes include *SRF* and *MCM*1 (Treisman and Ammerer, 1992).

1.4.1 Serum Response Factor

SERUM RESPONSE FACTOR (*SRF*), a type I MADS-box genes in mammals, is a transcription factor (TF) that regulate cell cycle phase transitions and muscle development (Treisman and Ammerer, 1992). SRF is a ubiquitously expressed TF protein that consists of 508 amino acids and has a molecular weight of 64Kda (Arsenian et al.,

1998). It regulates expression of many cellular immediate early genes such as c-fos in response to serum growth factor (Treisman, 1992). In response to serum, SRF recognizes a specific DNA sequence termed Serum Response Element (SRE) through its interaction with AT rich CarG-box sequence (CC(A/T)6GG) present at the DNA element to attach itself to the target genes and regulate expression of the targets (Treisman, 1992). *SRF* gene has been cloned and X-ray crystal structure of SRF MADS-box domain bound to SRE has been solved (Treisman, 1992; Treisman and Ammerer, 1992). Together, these data were used to map DNA binding, dimerization and transactivation domains in SRF. These mapping experiments mapped DNA binding and dimerization domain to the N-terminal MADS-domain respectively at positions 133 –222 and 168 – 222 and transactivation domain at C-terminal (Wynne and Treisman, 1992). GAL4-SRF chimerical transcription factor protein reporter assay experiment surprisingly showed that GAL4-full SRF construct could not activate expression of reporter gene (Wynne and Treisman, 1992). However, the same experiment showed when SRF is deleted at N-terminal up to position 203, the construct constitutively activated expression of reporter gene even in absence of the serum growth factor (Wynne and Treisman, 1992). This indicated presence of repressor domain at the N-terminal of SRF which over-lap with the MADS-box DNA binding and dimerization domain (Wynne and Treisman, 1992). Although MADS domain and repressor domain over-lap, DNA binding and dimerization domain were not responsible for the inhibition of SRF activity since mutations that affect DNA binding or dimerization did not affect inhibitory activity (Wynne and Treisman, 1992). Repressor domain function to inhibit SRF dependent transactivation of target genes when SRF is not bound to SRE allowing repression of the gene expression when transactivation is unnecessary (Davis et al., 2002). This discovery opened a way for application of the repressor domain to further study the function of SRF (Davis et al., 2002). Deletion of SRF trans-activation domain at C-terminal resulted in expression of truncated SRF that is capable of competing with WT SRF for binding to the target DNA elements but lacks the ability to upregulate transcription (Davis et al., 2002). Hence the truncated SRF functioned as a dominant negative mutant of SRF and was used in many studies to show the importance of SRF in various muscle-related gene expression (Davis et al., 2002).

It has been shown that SRF interacts with various accessory proteins via its MADS dimerization domain including homeo domain proteins phox1/Mhox and Nkx2.5, NF-kB, ATF6, myogenic bHLH factors, and HMG-I family of non-histone nuclear proteins (Marais et al., 1992; Marais et al., 1993). *In vitro* experiment replacing dimerization domain of SRF with that of MCM1 resulted in a chimerical SRF that can recruit MCM1 accessory protein STE12 (Wynne and Treisman, 1992). Depending these accessory proteins, SRF can function as both activator or repressor of its target genes (Wynne and Treisman, 1992). Extensive studies have been carried out with one interesting accessory protein family of SRF called Ets domain accessory family whose members have been shown to form ternary complex with SRF at SRE (Marais et al., 1993). Hence they are also called ternary complex factors (TCF) (Marais et al., 1993). In SRE region, there is a motif called Ets motif (GGA(A/T)) adjacent to the CarG box sequences and Ets domain proteins SAP-1 and Elk-1 have been shown to bind to the DNA through interaction with the Ets motif (Marais et al., 1993). SAP-1 and Elk-1 possess N-terminal Ets domain that bind to the Ets motif (Marais et al., 1993). These proteins also contain conserved 21amino acid regions called B-box domain located 50 residues C-terminal to the Ets domain and this B-box domain mediate ternary complex formation with SRF (Marais et al., 1993). There have been reports though, that some SREs do not contain the Ets motifs which indicate possibility that different SRF accessory proteins exist such as bHLH, NF-kB and so on (Marais et al., 1993).

Phosphorylation may also play a role in regulating SRF mediated gene expression as reported (Marais et al., 1992; Xi and Kersh, 2002). SRF itself is a phospho-protein and contains at least 4 phosphorylation sites that are targeted by Casein Kinase II (CKII) (Marais et al., 1992). It was shown that SRF have 2 phosphorylated forms – nascent non-phosphorylated form with molecular weight of 64Kda and fully phosphorylated 67Kda form (Marais et al., 1992). Upon phosphorylation, SRF is activated and carries out its function of regulating the gene expression (Marais et al., 1992). SRF have been shown to act both as activator or repressor of expression depending on its phosphorylation state (Marais et al., 1992). However, role of phosphorylation in regulation of SRF-mediated transactivation is somewhat obscure because recent research by other groups has shown that mutations that disrupt these phosphorylatioin sites did not affect transactivation ability of SRF (Iyer et al., 2003). Though phosphorylation may not directly affect activity

of SRF, it may have a role in indirectly regulating SRF-mediated transactivation by phosphorylating the accessory proteins of SRF (Bebien et al., 2003). For instance, Ets domain family proteins SAP-1 and Elk-1 possess MAP kinase target site at their C-terminal and when phosphorylated at the C-terminal, interaction between these proteins and SRF are induced to form a ternary complex which is capable of turning on the expression of target genes (Marais et al., 1993).

In vivo mechanism of how SRF activate transcription is yet to be discovered but SRF has been reported to interact with general transcription factor TFIIF, homeodomain pHOX1, and chromatin remodeling protein CBP (Arsenian et al., 1998). These 3 proteins could be interacting with SRF at SRE in response to serum to up-regulate the expression of target genes (Arsenian et al., 1998). Alternatively there could be yet unidentified protein factors that interact with SRF to induce expression in response to serum (Xi and Kersh, 2002).

There are many cases where SRF is shown to play role in developmental processes (Simon et al., 1997; Escalante and Sastre, 1998). For example, mice *srf* null mutant have shown that SRF is required for mesoderm formation in mice embryo and that disruption of this gene in mice results in embryonic lethal phenotype or dies within couple days of birth due to severe defect in muscle development (Arsenian et al., 1998). In *Dictyosterium*, SRFA have been shown to play role in final step of spore differentiation and in Drosophila, dSRF plays a role in development of trachea (Simon et al., 1997; Escalante and Sastre, 1998) .

1.4.2 Miniature chromosome 1

MCM1 (Miniature chromosome 1) is a yeast homologue of SRF and has been shown to determine mating type differentiation after activation by pheromone induced signal transduction pathway (Treisman and Ammerer, 1992). In yeasts, there are 3 cell types – haploid a, α, and diploid a/α (Treisman and Ammerer, 1992). These cells differ from each other in terms of their mating specificity and ability to form spore (Treisman and Ammerer, 1992). These differences are due to regulation of cell-type specific gene expression by MCM1 (Treisman and Ammerer, 1992). In α-cell, MCM1 forms a hetero-tetramer complex with homeodomain protein MAT α-2 (Treisman and Ammerer, 1992). MCM1-α2 complex binds 31bp partially symmetrical conserved DNA sequence located upstream of transcriptional start site of a-specific genes and repress their expressions

(Treisman and Ammerer, 1992). In contrast, in a-cell, MCM1 interacts with α-1 homeo-domain protein to positively regulate the transcription of a-cell specific genes to promote a-cell fate (Treisman and Ammerer, 1992).

MCM1 shares 72% identity with SRF over the conserved MADS-domain but binds slightly variant consensus sequence CC(C/T)AA(A/T)NNGG instead of the CarG box sequence (Wynne and Treisman, 1992). In addition, accessory proteins of SRF and MCM1 are quite diverse and some are shown to be specific to MCM1 or SRF while some other accessory proteins are capable of binding to both MCM1 and SRF (Wynne and Treisman, 1992). Like SRF, MCM1 is also capable of functioning as activator or repressor of transcription of target genes depending on the interacting partner such as stated in above example of interaction with α1 and α2 (Treisman and Ammerer, 1992).

In summary, type I MADS-box genes are playing various functions in wide range of developmental processes in animal and fungal kingdom (Treisman and Ammerer, 1992). Molecular mechanisms of how type I MADS-box genes are regulating transcription of target genes are not yet identified but their accessory proteins have been shown to play significant role in the regulation such that types of the interacting partner can signal type I MADS-box genes to function either as transctivator or repressor (Treisman and Ammerer, 1992).

1.5 Plant type I MADS-box genes

Recently functional characterization have been achieved with 2 members of plant type I MADS-box clade supporting the hypothesis that type I MADS-box genes have a functional role in plant development (Kohler et al., 2003; Portereiko et al., 2006). These 2 genes are *AGL80* and *AGL37* (*PHERES1*) (Kohler et al., 2003; Portereiko et al., 2006). Both AGL80 and PHERES1 were shown to be important for seed development in plants (Kohler et al., 2003; Portereiko et al., 2006).

1.5.1 *AGL80*

In angiosperms (flowering plants), seed formation initiates with a fusion of male gametes and female gametes (Faure et al., 2002). In Arabidopsis, double fertilization takes place in which 2 sperm nuclei (n) produced from male gametes fertilizes egg cell (n) and central cell (2n) of female gamete to produce diploid embryo (2n) and triploid endosperm

(3n) respectively (Faure et al., 2002). Embryo will eventually give rise to mature plant while endosperm is responsible for transporting maternally derived nutrient to embryo for proper embryo development (Faure et al., 2002). In last decade, number of genes involved in female gametophyte development have been isolated including *FIE, FIS2* and *MEA* (Luo et al., 2000).

In *agl80* mutant, endosperm development was absent even when the mutant was pollinated with WT pollen suggesting role of AGL80 (AGAMOUS-LIKE 80) in female gametophyte development (Portereiko et al., 2006). Consistent with the above statement, further phenotypic analysis showed central cell in *agl80* mutant plant had much smaller vacuole and nucleolus compared to the WT central cell that is characterized by a large vacuole and nucleolus (Portereiko et al., 2006). Further, AGL80::GFP reporter gene system showed expression of AGL80::GFP in central cell and in early developmental stage endosperm 3 days after pollination (Portereiko et al., 2006). In addition, looking at expressions of endosperm expressed genes *DME*, *FIS2* and *DD46* in *agl80* mutant back-ground, it showed expression of *DME* and *DD46* were disrupted in *agl80* back-ground while *FIS2* expression remained intact. In summary AGL80 is required for expression of at least some of the genes essential for endosperm development and knocking out the function of AGL80 results in defective endosperm development.

1.5.2 *PHERES1*

PHERES1 (*PHE1*) is another plant type I MADS-box gene whose role has been suggested in seed development (Kohler et al., 2003). *PHE1* was first identified by microarray assay as an up-regulated gene in *fis* class gene mutants indicating *PHE1* as a down-stream target of FIS protein complex (Kohler et al., 2003). FIS class proteins are a group of proteins that interact with each other to form PcG protein complex and repress expression of target genes and there have been many reports about the importance of epigenetic modification caused by FIS protein complex in seed development (Kohler and Makarevich, 2006). In WT, *PHE1* is not expressed in central cell before pollination but the expression is induced 1 – 2 days after pollination (DAP) in seeds containing pre-globular stage embryo (Kohler et al., 2005). In *fis* class mutants, *PHE1* expression is observed much earlier, already expressed directly after pollination (0DAP) because repression from FIS protein complex is absent in the *fis*-class mutants (Kohler et al.,

2005). Subsequent ChIP assay showed FIS class proteins MEA and FIE can bind *PHE1* promoter and confirmed *PHE1* as direct target of FIS protein Pc-G complex (Kohler et al., 2005). When *PHE1* expression was down-regulated in *mea* mutant back-ground, seed abortion phenotype of *mea* mutant was rescued confirming role PHE1 plays in seed development (Kohler et al., 2005).

Examples of *AGL80* and *PHE1* further supports the hypothesis that plant type I MADS-box genes do play a role in development. Interestingly both *AGL80* and *PHE1* belong to same sub-clade of type I MADS-box clade (Parenicova et al., 2003) and since both of them have function in seed development, research is underway to study the role of other members within this subclade in seed development (Portereiko et al., 2006).

1.6 Seed development

In their life cycles, plants alternate a haploid (n) gametophytic phase and diploid (2n) sporophytic phase (Berger, 2003). In *Arabidopsis thaliana*, sporophytic phase consists most of the plant's life-time and haploid gametophytic phase is reduced to only short time at the reproductive organ developmental stage before fertilization (Berger, 2003). Once Arabidopsis plants reach reproductive state, the diploid plants produce special cell lineages that undergo meiosis to produce mega-gametophyte (ovule) and micro-gametophyte (pollen) (McCormick, 1993; Drews and Yadegari, 2002). In ovule development, meiosis is followed by 3 syncytial divisions and cellularization to produce the mature 7-celled ovule consisting of egg cell, central cell, 2 synergid cells and 3 antipodal cells (Faure et al., 2002). In pollen development, meiosis is followed by two rounds of mitosis – pollen mitosis I (PMI) and pollen mitosis II (PMII) – to give rise to 2 sperm cells (McCormick, 1993).

Following the formation of these female and male gametophytes, a process termed double fertilization occurs which is a fusion of 2 haploid gametes and represents the end of gametophytic phase (Faure et al., 2002). After pollen grain is dropped onto style, pollen grain grows pollen tube to deliver 2 sperm cells to the ovule (Faure et al., 2002). As sperm cells are delivered to ovule, 3 antipodal cells that acted as a guide marker become degenerated and ovule now contains 4 cells (Faure et al., 2002). Once sperm cells are delivered to ovule, one of this male gamete fertilizes egg cell to produce diploid embryo (2n) (Faure et al., 2002). The other male gamete fertilizes central cell (2n) and

give rise to triploid endosperm (3n) that acts to deliver maternally produced nutrients to the embryo for embryo growth (Faure et al., 2002).

Endosperm development is mainly divided into 2 phases - synctium division and cellularization (Berger, 2003). Synctium division is a process in which mitotic division occurs without cell wall formation to produce multi-nucleate endosperm (Berger, 2003). Experiment has shown that 5 hours after pollination (H.A.P), one of the two synergid cells degenerate as pollen sperm cells are released into ovule (Berger, 2003). At this stage, change in egg cell polarity occurs and egg cell nucleus moves away from central cell (Faure et al., 2002). Then central cell nucleus becomes elongated along micropyle-chalazal axis (Faure et al., 2002). Six to seven HAP, central cell nucleus contains 2 nucleoli with the one in a micropylar position always being smaller than the one in chalazal position (Faure et al., 2002). After this stage, 2 nucleoli can also be observed in egg cell. Again they observed significant difference in size between the bigger nucleoli and smaller nucleoli (Faure et al., 2002). In central cell, 2 nucleoli have not changed in size or position at this stage (Faure et al., 2002). Fertilized central cell undergo first nuclear division as early as 7 HAP along the micropyle-chalazal axis (Faure et al., 2002). The outcome of this division is a 2-nucleate endosperm with its nuclei being adjacent to each other which can be observed at 8 to 9 HAP (Faure et al., 2002). Second and third nuclear division follows 12 and 24 HAP producing 8-nucleate endosperm at the end of third nuclear division (Faure et al., 2002). Mitotic divisions without cell wall formation continues to take place until 3 – 4 DAP at which stage multinucleate endosperm containing over 100 nuclei have developed representing the end of synctium division (Faure et al., 2002).

1.6.1 PcG complex mediated epigenetic control on seed development

Until recently only little was known about the molecular mechanisms underlying fertilization. However the recent sequencing of entire genome of *Arabidopsis* coupled with improvement in isolation of mutants has given further insight into molecular processes that take place during fertilization (Arabidopsis Genome, 2000; Dresselhaus, 2006). When pollen sperm cell (n) fertilizes egg cell (n), 2 nuclei of these gametes need to fuse in a same cell-cycle phase in order to avoid aneuploidy and to allow proper subsequent embryo growth (Dresselhaus, 2006). Therefore, there must be some sort of

cell cycle regulation in both the male and female gametes (Dresselhaus, 2006). Unlike in animal gametes, regulation of cell cycle in plant gametes remains largely unknown (Faure et al., 2002). Recently it was suggested that in *Arabidopsis*, sperm cells enter new S-phase after PMII and the sperm cells are likely to be in G2 phase when they are delivered to ovule (Dresselhaus, 2006). Since two nuclei from male and female gametes must fuse in same cell-cycle phase, simplest assumption was that female gametes are also in G2 phase at the time of fertilization (Dresselhaus, 2006). In female gametes, cell-cycle arrest is mediated by polycomb group (PcG) complex (Drews and Yadegari, 2002). These PcG complex genes include *FERTILIZATION INDEPENDENT SEED* (*FIS*), *MEDEA* (*MEA*), and *FERTILIZATION INDEPENDENT ENDOSPERM* (*FIE*) and loss-of-function mutation of these genes all result in autonomous initiation of cell division in central cell without the fertilization (Luo et al., 2000). The PcG complex genes are shown to be involved in epigenetic control of gene expression (Luo et al., 2000). Many evidences point to the importance of epigenetic control of gene expression on seed development (Luo et al., 2000; Makarevich et al., 2006). Particularly the regulation of imprinted gene expression in developing endosperm by methylation was shown to play significant role in seed development in *Arabidopsis* (Luo et al., 2000). Polycomb group (Pc-G) genes initially identified in *Drosophila melanogastor* are group of genes that maintain repression of target gene expression by methylating the target histone lysine residue and hence remodeling the chromatin structure (Muller et al., 2002). In *Drosophila*, pc-g mutants failed to maintain transcriptional repression of homeo-box genes (Muller et al., 2002). Pc-G genes assemble in 2 complexes to carry out their function of repressing expression of target genes (Muller et al., 2002). First, Polycomb Repressive Complex 2 (PRC2) or E(Z)/ESC complex that consists of 4 Pc-G proteins – Enhancer of Zeste (E(z)), Extra Sex Comb (ESC), Supressor of Zeste 12 (Su(z)12) and NURF-55 – bind to histone H3 and methylate lysine27 residue (H3K27) to create epigenetic mark (Muller et al., 2002). The second complex PRC1 is then recruited to H3K27 via interaction of its component with the methylated H3K27 (Muller et al., 2002). PRC1 complex also consists of 4 proteins – Polycomb (PC), Polyhomeotic (PH), Posterior Sex Combs (PSC), and dRing (Muller et al., 2002). After binding to target gene via methylated H3K27, PRC1 interacts with SWI/SNF chromatin remodeling complex to modify chromatin structure in a way that transcription initiation is blocked (Muller et al., 2002).

In plants, Pc-G genes were first identified in a genetic screen for mutants with defective seed development (Luo et al., 2000). Several Pc-G genes have been isolated since and presence of PRC2 like complex was confirmed (Makarevich et al., 2006). However, there is still no evidence for existence of PRC1-like complex in plants yet (Makarevich et al., 2006). To date, 3 PRC2 like complexes have been characterized FIS – complex, CLF complex and VRN complex (Reyes and Grossniklaus, 2003). Interestingly all 3 plant PcG complexes seem to have MADS-box genes as their target (Michaels et al., 2003). For example, VRN complex regulates expression of *FLC* and have a role in flowering time and CLF complex repress the expression of floral organ identity gene *AG (*Michaels *et al., 2003)*. Of these 3 complexes, FIS complex has been shown to play important role in gametophyte and early seed development (Makarevich et al., 2006). FIS complex consists of 3 proteins – MEA, FIS2, and FIE – and these 3 proteins are homologues of E(z), Su(z)12, and ESC respectively (Luo et al., 2000). All 3 *mea*, *fis2*, and *fie* mutants show common phenotypes – autonomous endosperm development in absence of fertilization, arrest of embryo development at heart stage and failure of mutant seeds to develop beyond endosperm cellulalization (Luo et al., 2000). These common phenotypes indicated that MEA, FIE, and FIS2 work in a same pathway to suppress various aspects of seed development in absence of fertilization and that fertilization inactivates FIS complex to initiate seed development (Luo et al., 2000). Since MEA and FIE are homologues of E(z) and ESC respectively and E(z) and ESC physically interact in *Drosophila* to repress target genes' expression, one can postulate that MEA and FIE do the same in *Arabidopsis* as well (Pien and Grossniklaus, 2007). Yeast-2-hybrid assay was carried out and it was shown MEA and FIE do indeed interact with each other (Kohler et al., 2003). The assay also showed FIS2 does not physically interact with neither MEA nor FIE (Kohler et al., 2003). This implies yet unidentified component in FIS complex that attaches FIS2 to the rest of complex (Kohler et al., 2003).

1.7 Plant type I MADS-box mutants

Plant type I and type II MADS-box genes have a few more differences apart from their MADS-domain structures (Parenicova et al., 2003). One difference is their distribution on the chromosomes (Johansen et al., 2002). *Arabidopsis thaliana* genome is composed of 5 chromosomes (Arabidopsis Genome, 2000). Type II plant MADS-box genes are evenly

distributed across all the 5 chromosomes whereas most plant type I MADS-box genes are located on chromosomes I and V (Parenicova et al., 2003). It has been estimated that up to 83% of the plant type I MADS-box genes are located on these 2 chromosomes (Parenicova et al., 2003).

For many eukaryotic transcription factor families, gene duplication that gave rise to them occurred predominantly between different chromosomes (Parenicova et al., 2003). Plant type I and II MADS-box genes differ in this aspect as well (Parenicova et al., 2003). Analysis of closely related members of type II and type I plant MADS-box genes showed that 53% of the type II probably originated from gene duplications between 2 different chromosomes (Parenicova et al., 2003). On the other hand, 82% of the type I genes are thought to have risen by internal gene duplication within the chromosome (Parenicova et al., 2003). In other eukaryotes, recent gene duplication occurred more frequently within chromosomes and hence this difference may be suggesting diversity originated more recently in subclades of plant type I MADS-box gene (Becker and Theissen, 2003; De Bodt et al., 2003b; Martinez-Castilla and Alvarez-Buylla, 2004). Consistent with this, type I plant MADS-box genes are further classified into several subclades by several researchers (Kofuji et al., 2003; Parenicova et al., 2003; Martinez-Castilla and Alvarez-Buylla, 2004). For example, Parenicova et al divided plant type I MADS-box genes into Mα, Mβ, Mδ, and Mγ subgroups (Parenicova et al., 2003).

Also in terms of exon/intron numbers, plant type II MADS-box genes typically contain 7 to 9 exons while type I groups tend to have only 1 or 2 exons (Nam et al., 2003).

On top of all these differences, type II MADS-box genes are much more well characterized than type I clades in plants (Alvarez-Buylla, 2001; De Bodt et al., 2003b; de Folter et al., 2005). This is quite Surprising since there are much more type I plant MADS-box genes than type II (Johansen et al., 2002). It may be possible that because plant type II MADS-box genes have been known for a longer time than the type I, more experiments including reverse genetic analysis have been carried out with the type II (Ostergaard and Yanofsky, 2004). However, statistically it is unlikely that this is the only reason for the unequal identification of mutants between type I and type II genes (Parenicova et al., 2003). One possible reason for not being able to identify type I MADS-box mutant in plant is that type I plant MADS-box genes are either non-functional or pseudogenes (Parenicova et al., 2003). There is evidence that at least one

plant type I MADS-box gene (*At5g49490*) is a processed pseudogene (De Bodt et al., 2003a). However, Expression profile and protein – protein interaction studies have shown that most type I MADS-box genes are expressed in various tissues of plants and are interacting with other proteins which strongly rejects this idea (Johansen et al., 2002; de Folter et al., 2005). Also phylogenetic studies and bioinformatics showed that homologous genes have been identified in rice (*Oryza sativa*) which indicates that MADS-box genes were present before the divergence of dicots and monocots (Kofuji et al., 2003; Martinez-Castilla and Alvarez-Buylla, 2004). Because type I MADS-box genes are expressed, conserved between species, and are capable of encoding group specific protein domains, it is unlikely that plant type I MADS-box genes are non-functional or pseudogenes. Another possibility of not being able to identify the type I MADS-box phenotypic mutant is that loss-of-function mutation in these genes give embryonic lethal phenotype (Parenicova et al., 2003). However, several T-DNA or transposon insertion mutants have been identified in the type I genes which rejected this idea (Parenicova et al., 2003). Redundancy could be another reason for not identifying type I MADS-box mutant in plants (Pelaz et al., 2001; Pinyopich et al., 2003). In type II plant MADS-box genes, several very clear examples of redundancy was observed. In Arabidopsis flower formation, B and C class genes controlling petal, stamen and carpel identity are functionally dependent on 3 very similar MADS-box genes, *SEP1, SEP2*, and *SEP3* and only when all of the 3 *SEP* genes were knocked out in *sep1 sep2 sep3* triple mutant was there a loss of petal, stamen, and carpel observed (Pelaz et al., 2000; De Bodt et al., 2003a). Another example of redundancy was observed in *SHP1* and *SHP2* (Rounsley et al., 1995; Liljegren et al., 2000). Single mutants of either *shp1* or *shp2* showed phenotype indistinguishable from that of WT but when these mutants were crossed to produce the *shp1 shp2* double mutant, disturbance in dehiscence zone development was observed in the double mutant fruit which resulted in a failure to release seeds (De Bodt et al., 2003a). These examples indicated that redundant activity is a common phenomenon in plant type II MADS-box genes, which could also occur in type I MADS-box genes.

1.8 AGL40 & AGL62

Of approximately just over 100 plant MADS-box genes, about 60 of them belong to type I lineage (Parenicova et al., 2003). As previously mentioned, type I lineage can be further sub-divided into several subclades based on their sequence similarity (Parenicova et al., 2003). Both *AGL40* and *AGL62* belong to Må subclade and share expression domain in inflorescence and sillique (fruit) (Parenicova et al., 2003). Because MADS-box genes tend to function in a site of expression (De Bodt et al., 2003a), their sequence similarity and similar expression pattern indicates *AGL40* and *AGL62* are most closely related to each other and that they may have redundant activities. Yeast 2 hybrid assay showed that AGL40 and AGL62 can physically interact with each other in embryo further pointing out to the possibility of AGL40 and AGL62 playing role in embryo development (de Folter et al., 2005).

1.9 Aim & Hypothesis

So far, only little functional characterization has been achieved with plant type I MADS-box clade (De Bodt et al., 2003a). However, functional characterization has been achieved extensively in other clades of MADS-box genes (De Bodt et al., 2003a) and knowing that plant type I MADS-box genes share the conserved MADS-box domain with other MADS-box gene clades, one can hypothesize that plant type I MADS-box genes also play a role in plant development. Recent findings that AGL80 and PHE1 play a role in seed development in plants also support this hypothesis (Kohler et al., 2005; Portereiko et al., 2006).

Therefore our aims in this project are

1) To give further evidence of plant type I MADS-box genes playing role in plant development and further the understanding of a role plant type I MADS-box genes play in plant development by studying functions of AGL40 and AGL62 – 2 most closely related members within plant type I MADS-box clade – using T-DNA knock out lines.

2) Identify spatial and temporal expression pattern of *AGL40* and *AGL62 in planta* and use the information obtained for functional characterization of *AGL40* and *AGL62*

Chapter 2 Materials & Methods

2.1 Media

All media were prepared using milliQ (MQ) water supplied from MILLI-Q® Reagent grade water system (MILLIPORE®, Billerica, MA USA) and autoclaved at 120°C and 15 psi for 20 minutes unless otherwise stated. Solid media were cooled to 45°C before addition of antibiotics. Liquid media were cooled to room temperature before the addition of antibiotics.

2.1.1 Luria-Bertani (LB)-medium

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. 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

MS agar plates were prepared following recipe provided by The Arabidopsis information resource (TAIR) web page (http://www.arabidopsis.org/). The MS agar plates contained 0.44g MS salt, 0.05g MES salt and 1g sucrose per 100ml of media. 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, autoclaved MS-agar solution and poured into plates.

2.1.3 Antibiotics addition to the media

All stock antibiotics were prepared and added to media according to instructions given in Molecular Cloning Laboratory manual (Sambrook J. 2001). Ampicillin was used to select for *E. coli* transformed with cloning vectors such as pGEM T-EASY (PROMEGA, Madison, WI USA) and pBSKS (PROMEGA, Madison, WI USA), Kanamycin and Rifampicin was used to select for *Agrobacterium tumefaciens* (*Agrobacterium*) transformed with binary vector pCAMBIA1381Xc (CAMBIA, Canberra Australia), and Hygromycin was used to select for plants that were successfully transformed with pCAMBIA1381Xc (CAMBIA, Canberra Australia). Stock concentration of all antibiotics

used were 10X and the final concentration of antibiotics in both solid and liquid media were 1X - 50ug/ml Ampicillin, 50ug/ml Kanamycin, 50ug/ml Rifampicin and 20ug/ml Hygromycin.

2.2 Buffers

Unless otherwise stated, MQ water supplied from MILLI-Q® Reagent grade water system (MILLIPORE®, Billerica, MA USA) was used to make buffers. Buffers were then autoclaved at 120°C and 15 psi for 20minutes. Where solution was not autoclaved before use, autoclaved MQ water was used. Unless otherwise stated, all the buffers were prepared according to Molecular Cloning Laboratory manual recipe (Sambrook J. 2001).

2.2.1 STET solution

The STET solution contained 8% sucrose, 5% triton X-100, 50mM EDTA and 50mM Tris at pH 8.0. The buffer was then filter-sterilized using sterile 30ml syringe (Becton Dickinson & Co., Franklin lakes NJ USA) instead of autoclaving.

2.2.2 TNE buffer

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 quantitation with fluorometer.

2.2.3 MOPS buffer

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

2.2.4 Ligation Buffer

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

2.2.5 TE buffer

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

2.2.6 DNA Extraction Buffer

DNA extraction buffer was prepared following recipe of Edwards K (Edwards 1991). The extraction buffer contained 200mM tris-HCl (pH 7.5), 250mM NaCl, 25mM EDTA (pH 8.0), 0.5% SDS per 100ml buffer.

2.2.7 Gel Loading Buffer

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.

2.2.8 TBE buffer

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

2.2.9 GUS staining buffer

GUS staining buffer was prepared following recipe of Miguel B. The GUS staining solution contained 0.2% triton X-100, 50mM sodium phosphate buffer (pH 7.0), 1mM potassium ferrocyanide, 1mM potassium ferricyanide and 2mM X-Gluc (PROBIOGEN biochemicals).

2.3 Solution

Unless otherwise stated, MQ water supplied from MILLI-Q® Reagent grade water system (MILLIPORE®, Billerica, MA USA) was used to make solutions. Solutions were then autoclaved. Where solutions were not autoclaved before use, autoclaved MQ water was used. All solution were prepared according to recipe of Molecular Cloning Laboratory manual (Sambrook J. 2001) unless stated otherwise.

2.3.1 1 kb Plus DNA Ladder

1kb Plus DNA ladder (Invitrogen Carlsbad, CA USA) contained 50μl of the stock ladder (1mg/ml), 100μl of 10X loading buffer (section 2.2.7) and 850ul of TE buffer (Section2.2.5).

2.3.2 Agarose gel

1g of agarose were added to 100ml 0.5X TBE buffer (Section 2.2.8) to make a 1% (w/v) gel. The solution was microwaved until the agarose completely melted. 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 of 0.5ug/ml. The gel solution was then poured into the gel running apparatus 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.

2.3.3 FAA

FAA contained 50% ethanol, 10% formaldehyde and 5% glacial acetic acid.

2.3.4 Glycerol 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°C.

2.3.5 *Agrobacterium* infiltration solution

Agrobacterium infiltration media was prepared following recipe of Andrew B (Bent 1998). The infiltration solution contained 6.6g of MS salt, 0.75g of MES salt and 75g of

sucrose in 1.5l of water. pH was adjusted to 5.6-5.8 with 1M KOH. 15μl of BAP (1mg/ml) and 300ul Silwet 77 (Osi Specialties, Inc., Danbury, CT, USA) were then added.

2.3.6 Competent *E.coli* cells

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.

2.3.7 Electro-competent *Agrobacterium* cells

Electro-competent *Agrobacterium* cells were prepared following manual of Katinger H (Diethard Mattanovich 1989). Single colony of *Agrobacterium* was inoculated in LB media containing Rifampicin (10ug/ml) at 30°C for 2 days. 2 ml of the preinnoculum 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.

2.4 Bacteria strains

2.4.1 *E.coli*

2.4.1.1 Strains and growth conditions

The *E.coli* strain used for transformation was XL1-Blue. *E.coli* cells were grown on LB-agar plates (solid) or LB-broth (liquid) supplemented with an appropriate antibiotic for selection and incubated at 37°C for 16-18 hours. Liquid culture was incubated with constant movement on a shaker at 200 rpm.

2.4.1.2 Heat shock transformation of E.coli

All transformation of the *E.coli* was done using the heat shock method following the Molecular Cloning Laboratory manual (Sambrook J. 2001). 100μl of the competent *E.coli* cells were thawed on ice. Then the ligation product was added to the cells and left on ice for 20 minutes to allow introduction of vectors to *E. coli*. The cells were then placed in a 42°C water bath for 90 seconds to heat shock the cell membrane to allow penetration of vectors into cell. The cells were transferred to ice for 1 minute before 1ml of autoclaved LB-broth was added to it. This mixture was then incubated at 37°C for at least 1 hour to enable the *E.coli* cells to recover and express the antibiotic resistance marker gene. After the recovery period, the transformed *E.coli* cells were spun down in the centrifuge at high speed for 10 minutes. The LB supernatant was pipetted out but leaving about 300μl in the tube. Then the pelleted cells were re-suspended by gentle vortexing. These resuspended cells were then plated on LB-agar plates containing an appropriate antibiotic to allow selection of successfully transformed cell.

2.4.1.3 Blue-White selection

The blue-white selection was used to separate antibiotic resistant *E.coli* colonies, which were transformed with just the empty vector from the colonies transformed with the recombinant vector (vector + insert). Blue – white selection was only possible where *E. coli* was transformed with vectors that had the lacZ operon such as pGEM-T EASY, and pBSKS II. This operon was designed such that when the plasmid was transcribed, the product of the lacZ gene, an enzyme could utilize X-gal and IPTG to give a blue colored product. Since the multiple cloning site (MCS) was engineered within the lacZ operon, if

the ligation was successful and insert was incorporated within the lacZ gene, the lacZ operon should be disrupted and the *E.coli* could not make the lacZ enzyme. Then *E. coli* cannot use X-gal to give the blue product. Hence giving recombinant *E. coli* a white appearance. However, if the ligation was not successful and the digested ends of the vector re-ligate without incorporating the insert DNA, then the lacZ operon should still be functional and thus *E.coli* transformed with just the empty vector can result in a blue colony when the *E.coli* was grown on the plates.

The Blue-White selection was carried out following protocol from Molecular Cloning Laboratory manual (Sambrook J. 2001). Before the transformed *E.coli* was plated on LB antibiotic agar plate, 40μl of 2% X-gal and 10μl of 20%IPTG was spread over the plate and placed at 37°C for 5 minutes to allow agar to absorb IPTG and X-gal. The transformed cells were then placed on the plates and incubated as stated above (section 2.4.1.2).

2.4.2 *Agrobacterium*

2.4.2.1 Strains and growth conditions

Agrobacterium strain used was LBA4404. Unless stated otherwise, *Agrobacterium* cells were grown on LB-agar plates or LB-broth supplemented with appropriate antibiotic selection and incubated at 30°C for 48hours. Liquid culture was incubated under constant motion on a shaker at 300 rpm.

2.4.2.2 Transformation of *Agrobacterium* cells through electroporation

All transformation of *Agrobactreium* cells were carried out using electroporation method following protocols of Katinger H (Diethard Mattanovich 1989). 100μl of the electro-competent *Agrobacterium* cells were thawed on ice. 500ng-1μg of plasmid DNA was added to the *Agrobacterium* cells and mixed thoroughly by pipetting. The mixture was placed inside the electroporation cuvette between the two steel plates and placed on ice for 5-10minutes. The cuvette was wiped dry and placed in the holder of the PULSE CONTROLLER electroporater 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 2-3hours before plating on an appropriate antibiotic plate.

2.4.2.3 Preparation of *Agrobacterium* cells for floral dipping plant transformation

5ml culture with appropriate antibiotics was made from glycerol stock (Section 2.3.4) and incubated at 30°C for 48hours on a shaker. The culture was poured into a 2.5L flask containing 500ml of LB-broth and incubated for 12-16hours (overnight) for transformation into plants. On the day of plant transformation, the overnight culture was centrifuged at 6000 rpm for 10 minutes and the pellet was suspended in the infiltration media (section 2.3.5).

2.5 *Arabidopsis* growth and tissue preparations

2.5.1 Plant strains and growth conditions

For wild type (WT) *Arabidopsis* plants, the Colombia ecotype was used. T-DNA insertion lines for both *AGL40* (SALK107011) and *AGL62* (SALK022148) were obtained from SALK institute (www.signal.salk.edu/). *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/)

2.5.2 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 for 2 days to allow stratification of seeds so that they germinate at the same time. Seeds were then placed on the surface of the soil. 5 seeds were 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 were placed under the growth light and grown under condition as mentioned (section 2.5.1). The plastic bags were removed when the germinated seedlings had produced at least 2-3 true leaves, usually after 7days. The soil was kept moist by adding water to the base of the tray.

2.5.3 Leaf tissue preparation

For nucleic acid extraction, 2 to 3 rosette leaves were harvested and snap frozen in liquid nitrogen following Kasajima's protocol (ICHIRO KASAJIMA 2004). An autoclaved pestle and eppendorf tube was used to homogenize leaf tissues for DNA extraction. A tube lack was placed inside Styrofoam box to prevent from warming. Liquid nitrogen was poured on the tube lack and allowed to evaporate 3 times in order to cool the lack. Liquid nitrogen was added to the tube containing frozen leaf tissues just enough to cover the tissues completely. Then leaf tissues were ground with the pestle to a fine powder. The homogenized tissue was stored at -80°C or used immediately for DNA or RNA extraction.

2.6 Molecular biology

Unless stated otherwise, all molecular biological methods were carried out following protocols from Molecular Cloning Laboratory manual (Sambrook J. 2001).

2.6.1 Genomic DNA extraction

Genomic DNA extraction from Arabidopsis leaf tissue was carried out as described by Kasajima (ICHIRO KASAJIMA 2004). To extract genomic DNA, 700μl of DNA extraction buffer (section 2.2.6) was added to every 100μg of homogenized plant tissue and the mixture was vortexed briefly at room temperature to mix for 10 minutes. Then 700μl of phenol-chloroform-isoamyl alcohol (25:24:1) was added and vortexed again. The mixture was centrifuged at 13000 rpm for 10 minutes and the supernatant was transferred to a new Eppendorf tube. 0.1-volume 3M sodium acetate (pH5.2) and 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 20μl TE buffer (Section 2.2.5). Genomic DNA extractions were stored at –20°C until use

2.6.2 Polymerase Chain Reaction (PCR)

PCR reactions were carried out using an Eppendorf Mastercycler® (Eppendorf South Pacific, North Ryde NSW Australia). To optimise 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 analysed by gel electrophoresis and visualised by UV trans-illuminator (section 2.6.3).

2.6.2.1 Oligonucleotide primers

The primers were ordered from Gibco BRL, Sigma Corporation or Life Technologies. Each primer was re-suspended in TE buffer (Section 2.2.5) 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. Table 2.1 lists the primers used in this study.

Table 2.1 Oligonucleotide primers used in this study

Primer name	Sequence (5 →3)	Application
001AGL40	GAGAAAATGGTGAGAAGTACC	Cloning AGL40
002AGL40	GTTCTTCTAGCTTCAGACTCTAGC	Cloning AGL40
001AGL62	CATTCTACAAACACCAATCC	Cloning AGL62
002AGL62	GTTGGATCACACAACAAAGGG	Cloning AGL62
003AGL62	CACTTCTGCACCACAAAGTGTGC	Cloning AGL62
004AGL62	CAATCCATCAAATGAGATAGGGGAGC	Cloning AGL62
005AGL62	GGTGGATCTTTCTGGCAGATTTGG	Cloning AGL62
006AGL62	GCGTAAACGTAACCATGGCAGAC	Cloning AGL62
LBa1	TGGTTCACGTAGTGGGCCATCG	Cloning T-DNA fragment in AGL40 & AGL62
LBb1	GCGTGGACCGCTTGCTGCAACT	Cloning T-DNA fragment in AGL40 & AGL62
001GUSAr	GTTGGGGTTTCTACAGGACG	Checking insertion of GUS construct in plants
M13F	TTGTAAAACGACGGCCAG	Sequencing
M13R	CAGGAAACAGCTATGACCATG	Sequencing

2.6.2.2 Standard PCR conditions

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^{TM} High Fidelity polymerase (Roche, Indianapolis, IN USA). The thermo cycle 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.

2.6.3 Agarose gel electrophoresis (DNA and RNA)

Agarose gel electrophoresis was used to separate DNA or RNA following PCR, restriction endonuclease digestion and extraction. 10X loading dye (Section 2.2.2.7) was mixed with DNA sample to reach a final concentration of 1X and loaded onto the agarose gel (Section2.3.2). The agarose gels were run at 80-110V for 1hour until the desired separation had been achieved. The DNA was then visualized using the ethidium bromide fluorescence under short-wavelength UV trans-illuminator and digital photographs were obtained using Quantity One software (Bio-Rad, Albany Auckland New Zealand). Size of the DNA fragments were determined by comparing the mobility of the DNA samples with that of the 1kb Plus DNA ladder.

For loading total RNA samples 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 to the sample. The mixture was heated at 65°C for 2minutes. Then 1.2μl of RNA dye and 1μl of 0.1mg/ml ethidium bromide was added before loading the RNA on the gel. 1% gel was made but without addition of ethidium bromide and run at 80V.

2.6.4 Gel purification of DNA samples

DNA samples (PCR or restriction endonuclease digest products) were purified after running on a 1% agarose gel (Section 2.6.3) for an appropriate time to achieve the desired separation of DNA sizes. The desired DNA band was cut under long UV light (366nm) using a scalpel. Gel pieces were placed in a 0.5ml Eppendorf tube which had a small pinhead-sized hole made at the bottom and covered with glass wool that itself was placed within 1.5ml Eppendorf tube and centrifuged at 6000 rpm for 10minutes to collect the solution containing DNA in 1.5ml eppendorf tubes. A phenol chloroform wash and DNA precipitation was then carried out for DNA purification. Equal volume of 1:1 phenol

chloroform was added to the DNA solution after collecting DNA at bottom of 1.5ml eppendorf tubes. Upon addition of phenol chloroform, the mixture was centrifuged at high speed for 5min and the supernatant was moved to a new tube. The DNA was precipitated by adding 0.1 volumes 3M sodium acetate (pH 5.2) and 2.5volume 100% ethanol and placing on ice for 10minutes. After centrifugation at high speed for 10minutes, the pellet was washed with 70% ethanol, air dried and re-suspended in 10 – 20ul TE buffer (Section 2.2.5) and stored at -20°C until use.

2.6.5 Gel DNA quantification

The amount of purified DNA present in each sample was quantified before ligation reactions were set. 1μl of both the vector and insert DNA was loaded onto separate wells in the gel and electrophoresed for 1 hour. DNA concentrations in the two samples were estimated by comparing the brightness of the bands observed on gel. The number of molecules in each of the samples were estimated using the above information and the known size of the vector and insert fragment. This information was then used to determine the amount/ratio of vector and insert DNA used in the ligation reactions.

2.6.6 Ligation reaction (DNA)

DNA ligation reactions were performed to incorporate the insert DNA (e.g. the gene) into the vector. Vectors used include pGEM T-EASY (PROMEGA, Madison, WI USA), pBSKS (PROMEGA, Madison, WI USA), and pCAMBIA1381Xc (CAMBIA, Canberra Australia). The quantity (μg) of the vector and insert DNA used depended on the amount present in the extracted samples. Generally, 1:3 to 1:8 ratio of the vector DNA to the insert DNA was used to drive the ligation reaction. On some occasions, a ratio of vector to insert DNA went up to as high as 1:10 and ligation reaction was heated at 65°C for 5minutes to allow incorporation of insert DNA into vector. The vector and insert DNA was mixed with an appropriate amount of a 2X ligation buffer (Section 2.2.4) 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 overnight for blunt end ligation or at room temperature for 2-16 hours for sticky end ligation.

2.6.7 Boiling Lysis (Mini Prep) for plasmid DNA extraction

After ligation, recombinant vector was transformed into *E. coli* as stated above (section 2.4.1.2) and recombinant *E. coli* was selected using antibiotic resistance and blue – white selection (section 2.4.1.3). Recombinant vector was isolated from these *E. coli* by boiling lysis method following protocol developed by Holmes, S (Quigley 1981) and restriction digested with appropriate restriction endonucleases to confirm presence of insert DNA in the vector. 1-1.5ml of a recombinant *E. coli* o/n culture was placed into an Eppendorf tube and centrifuged at high speed for 15-20 seconds to pellet the bacteria. The LB supernatant was discarded and the pellet was re-suspended in 200μl of STET buffer (Section 2.2.2.1). 3μl of RNAse (10mg/ml) (for every 200μl of STET) was added to the resuspended cells followed by 10μl of lysozyme (10mg/ml). The tubes were inverted gently to mix. The mixture was immediately placed in boiling water (100°C) for 1minute and centrifuged at high speed (13000 rpm) for 10minutes. The slimy pellet was removed with a sterile toothpick and discarded. 0.7 volume of isopropanol was added to the supernatant to precipitate DNA. The mixture was centrifuged again at high speed for 10minutes and the supernatant was discarded. White DNA pellet at the bottom of the tube was washed with 70% ethanol, air-dried and re-suspended in 20μl TE buffer (Section 2.2.5). DNA was either used immediately for restriction endonuclease digest or stored at -20°C.

2.6.8 Restriction endonuclease digestion

Restriction endonuclease digests of plasmids were performed to determine the success/failure of ligations of inserts into vector 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 1hour to overnight at 37°C. Digested DNA was then run and separated on agarose (1%) gel electrophoresis (Section 2.6.3).

2.6.9 Plasmid DNA isolation using a Quantum Miniprep kit

Plasmid DNA extraction using a Quantum Prep Plasmid Minipre 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 neutralisation 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 sterilised water to the matrix and centrifuged for 1 minute.

2.6.10 Fluorometer DNA quantification

DNA was quantified using a fluorometer for sequencing. A Hoefer DyNA Quant 200 (Amersham, Greenlane Auckland New Zealand) was used to this end. 2ml of 1X TNE containing Hoechst dye (Sectioin 2.2.2) 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 1X TNE was placed in the cuvette and 2μl of the sample (test) DNA was added. DNA concentration in ng/μl was recorded.

2.6.11 Automated DNA sequencing

DNA was sequenced at Alan Wilson Center (AWC) (www. awcmee.massey.ac.nz/) by the dideoxynucleotide chain termination method (Sanger *et al.*, 1977), using Big-Dye chemistry (PE Applied Biosystems, Foster City, CA, USA) and oligonucleotide primers synthesised 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).

2.6.12 Total RNA extraction

Total RNA was extracted as instructed in protocol of TRIzol® reagent (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 xg for 10 minutes at 4°C. The supernatant was mixed vigorously with 0.2 ml chloroform, kept at room temperature for 3 minutes, and then centrifuged at 12,000 xg 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 xg 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 is stored at -80 °C until use.

2.6.13 Nanodrop RNA quantification

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.

2.6.14 RT-PCR

RNA was reverse transcribed into cDNA following SuperScript™ III First-Strand Synthesis System manufacture's instructions (Invitrogen, Carlsbred, CA, USA). RNA/primer mix solution was prepared by mixing 1ug RNA, 1ul of 50uM oligo(dT)20 primers, 1ul of 10mM dNTP mix and topped up to 10ul with DEPC water. RNA/primer mix solution was then heated for 5minutes at 65°C and cooled on ice for 1 minute. 10ul cDNA synthesis mix solution containing 2ul 10X RT buffer, 4ul 25mM Magnesium chloride, 2ul 0.1M DTT, 1ul RNaseOUT™ (40U/ul) and 1ul SuperScript™ RT (200U/ul) was then added to RNA/primer mix solution. The mixture was mixed gently and incubated at 50°C for 50minutes for reverse transcription of mRNA into cDNA. The reaction was then terminated by heating at 85°C for 5minutes and chilled on ice for 1

minute. 1ul of RNase H (2U/ul) was added to the reaction and incubated at 37°C for 20 minutes to remove the RNA template.

1 µl of each cDNA sample was PCR amplified in a reaction mixture of 50 µl using forward and reverse gene specific primers. The thermocycle conditions were 1 cycle of 2 min at 94°C, 35 cycles of 30 sec at 94°C, 30 sec at 55°C and 1 min at 72°C, and 1 cycle of 5 min at 72°C. 10ul of PCR product was run on gel for analysis (Section 2.6.3).

2.7 Plant transformation

2.7.1 Plant transformation

Plants were transformed with binary vector pCAMBIA1381Xc (CAMBIA, Canberra Australia) using the floral dip method following protocol of Andrew, B (Bent 1998). Bolting plants about 3 weeks old were decapitated 5-7 days 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 excessively. On the day of the transformation, the plants were clipped to remove all siliques and flowers older than stage 16 of development (David R. Smyth 1990) to minimize non-transformed seeds in the batch. The plants were placed upside down on a wire net to allow flowers to be dipped in the infiltration solution (Section 2.3.5) containing the re-suspended *Agrobacterium* cells with the desired construct. The plants were dipped for 45min. The plants were then placed on their sides on a tray, covered with plastic to maintain humidity and left out of direct light for 24hrs. The transformed plants were then placed upright in the growth light and allowed to grow/develop normally. Seeds were collected as usual from dried siliques.

2.7.2 Sterilization and plating of seeds

The seeds were sterilized in a laminar flow following protocol of TAIR (http://www.arabidopsis.org/). 1ml of seeds were vernalized at 4°C overnight. The seeds were then washed in 50ml of 70% ethanol for 5 minutes and then in 50ml of 40% bleach for 10 minutes to surface sterilize. The seeds were washed several times with autoclaved water. The seeds were divided into 3 portions and 30ml of cool but molten

0.3% autoclaved agar was added to each portion. 10ml of the agar containing the seeds were plated on MS-phytoagar plates supplemented with the appropriate antibiotics for selection of transformed plants (kanamycin for T-DNA insertion and hygromycin for pCAMBIA1381Xc (CAMBIA, Canberra Australia)). The MS-phytoagar plates were sealed with parafilm and placed under the growth light and allowed to germinate.

2.7.3 Selection of resistant seedlings and transplanting

Resistant seedlings can be detected after 7-10 days after plating following protocol of TAIR (http://www.arabidopsis.org/). The roots of the resistant seedlings were able to grow and elongate toward the base of the plates by penetrating the agar whereas the antibiotics included in MS agar killed non-transformed seedlings. The surviving seedlings were carefully removed from the agar; the roots were lightly washed with water to remove traces of the antibiotic and transferred to soil. The seedlings were allowed to grow as usual (Section 2.5.1).

2.8 Phenotypic analysis

2.8.1 Manual pollination

2.8.1.1 Emasculation

Mature plants were emasculated to inhibit self-pollination and then cross pollinated with plants with different genetic back-ground to generate multiple knock-out mutants following instruction given by TAIR (http://www.arabidopsis.org/). Flowers at stage 10 or 11 (David R. Smyth 1990) of development were lightly pressed/constricted at the base of the sepals using fine forceps to expose the floral organs. With a pair of fine forceps, the 6 anthers were carefully removed without damaging the other organs of the flower, especially the stigma and ovary.

2.8.1.2 Pollination

Emasculated flowers were pollinated after 1-2days once the flower reached stage 12 of flower development (David R. Smyth 1990) which corresponded to when the stigmatic papillae had fully developed and was receptive to pollen. Freshly dehisced anthers

containing mature pollen were picked from flowers at stage 13 of development (David R. Smyth 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, the pollinated stigmas were allowed to develop to maturity and the seeds were collected.

2.8.2 Microscopic observation of tissues

2.8.2.1 Embedding plant tissues

In order to embed plant tissue in paraplast for sectioning, the tissues were prepared following TAIR's protocol (http://www.arabidopsis.org/). The tissues were first harvested and chemically fixed in FAA solution (Section 2.3.3) overnight at 4°C. Then tissues were passed through an ethanol series of 50%, 70%, 85%, and 95% for 1-1.5 hour each. Then the tissues were placed in 100% ethanol for a total of 4.5 hours, changing the 100% ethanol every 1.5 hour. The tissues were then passed through a graded xylene:ethanol series containing 25%, 50%, 75% xylene for 1 hour each and finally in 100% xylene for a total of 3hr, changing 100% xylene every 1 hour. Tissues were then immersed in xylene:paraplast (50:50) solution. The tissues were placed in a 60°C oven to melt the paraplast. Tissues were moved to 100% molten paraplast and the paraplast was changed several times with 3 hours interval until all traces of xylene were removed. The tissues were then embedded in paraplast on the embedding machine.

2.8.2.2 Sectioning

10μm thick sections of the embedded tissues were made using the microtome. 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 xylene before staining the tissues or before observing under the microscope.

2.8.2.3 Alexander staining

Alexander stain was prepared following protocol of Alexander, M (Alexander 1969). 10ml 95% Ethanol, 10mg (1ml of 1% solution in 95% ethanol) Malachite Green, 50ml MQ water, 25ml Glycerol, 5g Phenol, 5g Chloral hydrate, 50mg (5ml of 1% solution in

water) Acid fuchsin, 5mg (0.5ml of 1% solution in water) Orange G and 1ml Glacial acetic acid were added in a given order. The solution was then mixed thoroughly and stored at room temperature in a bottle wrapped in aluminium foil to protect from light until use.

Malachite green is a weakly basic dye and included in the Alexander dye to stain the pollen wall in green while acid fuchsin was included in dye to stain protoplasm in red to deep red. Addition of orange G to the stain was shown to improve the differentiation. Since the aborted pollens only consist of a pollen wall, malachite green stains them in green whereas the viable pollens are stained red by acid fuchsin on inside and green on the outside wall by malachite green.

Anthers were picked from the individual flowers at various developmental stages using forceps and tapped onto slide glass to release pollens on a slide. Drop of Alexander stain was added to the slide and covered with a cover slip, then warmed over a small flame. The stained pollen was then observed under microscope and number of green aborted pollens and red viable pollens were counted.

2.8.2.4 Hoyer's medium tissue clearing

Hoyer's tissue clearing method was carried out following protocol of Stangeland, B (Zhian 2002). Hoyer's medium was prepared by mixing the chemicals in a following order - 30g Chloral hydrate, 10ml Water, 1ml glacial acetic acid and 2ml 100% glycerol. Hoyer's medium was wrapped in aluminium foil and stored at room temperature to protect from light.

This protocol pre-treats the tissue in ethanol:acetic acid (1:1) before clearing the tissues in Hoyer's medium. Pre-treatment with ethanol: acetic acid solution significantly improved the clearing and allowed clear detection of GUS stains in ovules.

Sillique either stained with GUS or not were harvested and slit longitudinary to release the developing seeds into epitubes. Ethanol:acetic acid (1:1) were added and incubated at room temperature. Time of incubation depended on the ages of seeds and younger stage seeds containing up to heart stage embryo were incubated for 1 – 4 hours while the older stage seeds containing embryos after heart stage were incubated for 4 – 8 hours. Ovules were then washed twice with water and then cleared in Hoyer's medium overnight at

room temperature covered with aluminium foil. Cleared ovules were then mounted in Hoyer's medium and observed under Zeiss microscope equipped with differential interference contrast (DIC) optics either for studying embryo development or detecting GUS stains in ovules or seeds.

2.9 GUS reporter gene assay

2.9.1Construction of GUS fusion genes

1.5kb upstream region and genomic region of *AGL62* was first PCR amplified using gene specific primers 004AGL62 (5' CAATCCATCAAATGAGATAGGGGAGC 3') and 002AGL62 (5' GTTGGATCACACAACAAAGGG 3') and cloned in PGEM T-EASY (PROMEGA, Madison, WI USA) for sequencing. After sequence was examined, 2241bp fragment containing upstream and genomic region of *AGL62* was sub-cloned into binary pCAMBIA 1381Xc vector (CAMBIA, Canberra Australia) for in frame fusion of AGL62 and GUS.

2.9.2 GUS histochemical assay

Anthers, carpels, and fruits were collected at a various developmental stages for histochemical staining (David R. Smyth 1990). For observing GUS stains in ovules or seeds, histochemical assay was carried out as described by (Richard A. Jefferson 1987). Carpel was dissected at bottom using razor blade to allow entry of GUS staining solution into the ovules. These tissues were incubated in 90% acetone at –20°C for 20 minutes for prefixing. Tissues were then washed in 100mM sodium phosphate buffer 3 times to remove acetone and then vacuum filtrated twice - first with GUS staining buffer without X-gluc for 15minutes and then 5minutes in fresh GUS staining buffer with X-gluc. Tissues were incubated for 3 days at 37°C after vacuum filtration to allow blue precipitate formation.

Chapter 3. AGL62::GUS temporal & spatial expression pattern

3.1 Introduction

In vitro microarray data (www.genevestigator.ethz.ch/) indicate that *AGL62* is expressed in various plant organs with high levels of expression in roots and male reproductive organs suggesting a possible role of AGL62 in regulation of their development (table 3.1). To investigate the temporal and spatial expression pattern of *AGL62* in detail, a reporter fusion construct was made between β-glucuronidase (GUS) reporter gene and promoter and genomic regions of *AGL62* (AGL62::GUS) and introduced *in planta*. GUS originates from *E. coli* and is a stable tetramer protein with a molecular weight of 68200 Da that catalyzes the break down of X-gluc (5-bromo-4-chloro-3indolyl glucuronide) to produce an indoxyl derivative that undergoes a subsequent oxidative dimerization to form a blue precipitate (Jefferson et al., 1987). Therefore GUS gene fusion assays allow the identification of site of *AGL62* expression by observing blue precipitate after the treatment of AGL62::GUS transgenic plant with a solution containing X-Gluc.

Plant tissue	Expression level
Root	High
Leaf	Low - High
Stem	Low
Inflorescence	Moderate
Pollen	High
Ovule	Low
Fruit	Moderate
Seed	Moderate

Table 3.1 *In vitro* microarray *AGL62* expression data. Spatial expression pattern of *AGL62* was calculated in web-based *in vitro* microarray assay (www.genevestigator.ethz.ch/). The assay showed high levels of *AGL62* expression in root and pollen and moderate level of expression in seed.

3.2 Construction of AGL62::GUS reporter gene system

The AGL62::GUS construct used to transform the plant for this assay is illustrated in figure 3.1. A 1.5kb upstream region and 1kb genomic region of *AGL62* was fused to GUS N-terminal coding region through sub-cloning. *AGL62* has 1kb genomic region that spans 2 exons and 1 intron. In the construct, all of 1.5kb upstream region and almost all of genomic region was included except for the stop codon in exon 2 to cover all the possible *AGL62* regulatory regions.

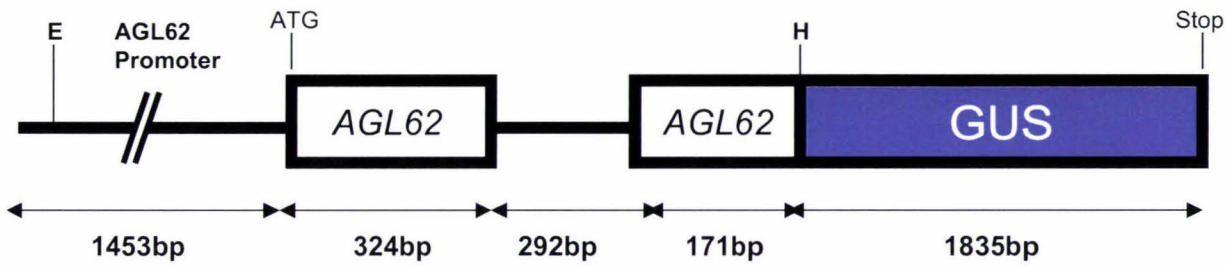


Figure 3.1. AGL62::GUS construct used to transform WT *Arabidopsis thaliana*. White rectangles represent exons of *AGL62* while blue rectangle indicates GUS coding region. Thick black lines represent intron region of *AGL62* between 2 exons and approximately 1.5kb upstream region of *AGL62*. E and H represent the Eco RI and Hind III restriction sites respectively. Size of each segment of construct is shown under each region that is delineated by a double-headed arrow line. Start codon and stop codon are indicated by ATG and Stop respectively.

Approximately 1.5kb upstream region and genomic region of *AGL62* was first PCR amplified using gene specific primers 004AGL62 and 002AGL62 in Expand High fidelity system and cloned in PGEM T-EASY for sequencing (chapter 2.6.2.2). Insertion of the *AGL62* upstream and genomic regions in the pGEM T-EASY was confirmed by single restriction digest with *Eco*R1 and also with double restriction digest with *Eco*R1 and *Xho*1 (Figure 3.2A) (Chapter 2.6.8). The single *Eco*R1 restriction digest released approximately 2.6kb *AGL62* upstream and genomic band from the 3kb pGEM T-EASY plasmid and double restriction fragmented the 2.6kb band into approximately 2.1kb and 500bp bands in accordance with *AGL62* restriction map (Appendix1). After the sequence fidelity was confirmed, recombinant PGEM T-EASY was then double digested with

EcoRI and Hind III to release a 2241bp fragment containing upstream and genomic region of *AGL62* except for the stop codon (Appendix 1). This fragment was then purified and sub-cloned into a binary vector for *Agrobacterium tumefaciens* (*Agrobacterium*) mediated plant transformation (chapter 2.7.1). Binary vectors are plasmids that are specifically designed to facilitate the production of transgenic plants and so called because of its ability to replicate in both *E. coli* and *Agrobacterium* (Hellens et al., 2000). pCAMBIA 1381Xc is a binary vector approximately 10kb in size that already contains the GUS coding region for in-frame fusion and was used in this assay for in frame fusion of *AGL62* and GUS. Insertion of the 2.241kb *AGL62* upstream and genomic band in 10kb pCAMBIA 1381Xc was confirmed by restriction digest (Chapter 2.6.8) of the recombinant pCAMBIA 1381Xc with *Eco*R1 and *Hind*III that released the 2.241kb band (Figure 3.2B).

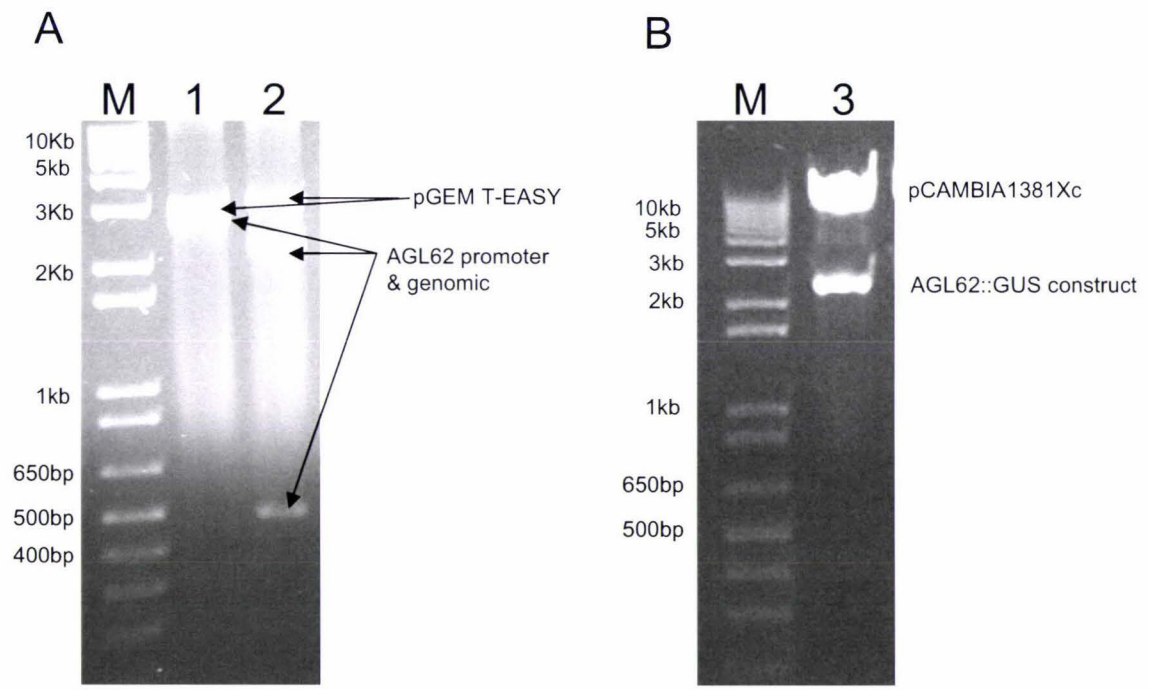


Figure 3.2. Restriction digest of recombinant vector to confirm the presence of AGL62::GUS construct in vector. A) M - 1Kb plus DNA marker, lane 1 - 3Kb pGEM T-EASY cloning vector digested with EcoR1 to release approximately 2.6Kb *AGL62* Genomic fragment, lane 2 - pGEM T-EASY cloning vector digested with EcoR1 and Xho1 showing 2.1kb and 500bp *AGL62* genomic fragment. B) M – 1kb plus DNA marker, lane 3 – Confirmation of 2.6kb of *AGL62* genomic region in pCAMBIA1381Xc when digested with *Eco*R1 and *Hind*III

3.3 Plant transformation with AGL62::GUS construct

In order to make transgenic plants, the AGL62::GUS construct then needed to be moved to *Agrobacterium*. *Agrobacterium* is a genus of bacteria that is capable of transferring DNA between itself and plant and is widely used in plant biology to make transgenic plants (Hellens et al., 2000). The recombinant pCAMBIA 1381Xc containing chimerical *AGL62::GUS* fusion construct was isolated from *E. coli* and transformed to *Agrobacterium* strain LBA4404 by electroporation (chapter 2.4.2.2). WT (T0) plants were transformed with recombinant *Agrobacterium* harboring AGL62::GUS construct by floral dipping method (Chapter 2.7.1). Since pCAMBIA1381Xc confers hygromycin resistance (Hellens et al., 2000), T1 seeds were sterilized and plated on MS Hygromycin (50ug/ml) plates. The green surviving seedlings were transferred onto soils to mature. To verify the presence of the AGL62::GUS construct in the *Arabidopsis* genome of the hygromycin resistant plants, PCR genotyping was performed using *AGL62* gene specific primer 004AGL62 and 001GUSAr (Chapter 2.6.2.2). Since 001GUSAr binds at the start of GUS coding regions, the size of the AGL62::GUS construct to be amplified should be roughly equivalent to the promoter/genomic region of *AGL62*. The PCR genotyping amplified approximately 2kb AGL62::GUS construct region from hygromycin resistant plants indicating the presence of the construct in these plants (Figure 3.3). Following the identification of transgenic plants, 13 T1 individuals were chosen for subsequent histochemical assay With X-Gluc (Chapter 2.9.2).

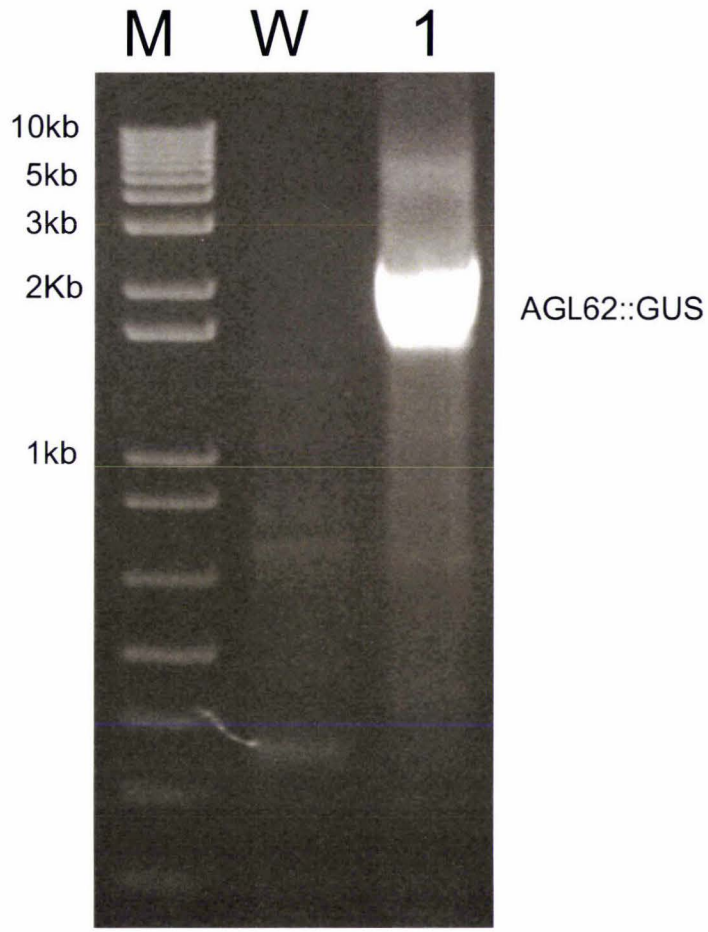


Figure 3.3. PCR genotyping confirmation of presence of AGL62::GUS construct in transgenic plant that showed Hygromycin resistance. Primers 004AGL62 (5' CAATCCATCAAATGAGATAGGGGAGC 3') and 001GUSAr (5' GTTGGGGTTTCTACAGGACG 3') were used to detect approximately 2kb AGL62::GUS construct. Marker used was 1Kb plus. M – 1Kb plus DNA marker, W – WT, 1 – AGL62::GUS transgenic line.

3.4 *AGL62* expression in Seedling root

According to the microarray data (www.genevestigator.ethz.ch/), *AGL62* expression level was high in seedling root and pollen and modest level of expression was also observed in seeds (table 3.1). To assess *in vivo* expression pattern of AGL62::GUS in the transgenic seedlings, seeds were collected from 2 T1 transgenic lines that showed the strongest AGL62::GUS expression and T2 seeds were plated on MS Hygromycin (50ug/ml) plates and collected 3 to 9 days after germination. Unless otherwise stated, all the subsequent GUS assays were carried out using these 2 lines. Collected tissues were assayed for GUS expression (Chapter 2.9.2) as shown in figure 3.4 and 3.5.

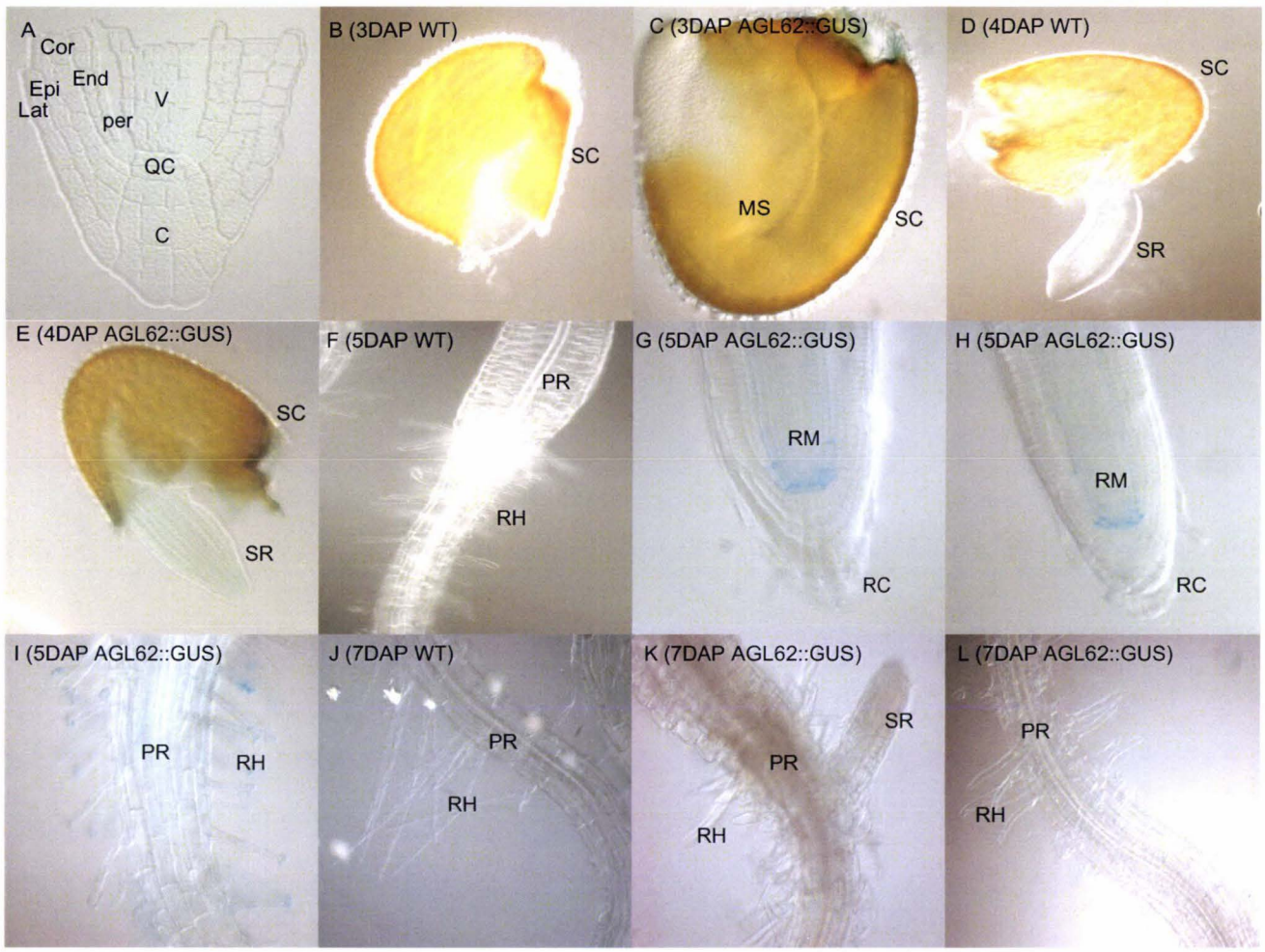


Figure 3.4. GUS assay in AGL62::GUS seedling root. A) Structure of *Arabidopsis* root tip modified from www.arabidopsis.org/ .B) WT seed 3days after plating (DAP) on MS media. C) Ab597-6 seed 3DAP stained for GUS. D) WT seed 4DAP. E) Ab597-6 seed 4DAPstained for GUS. F) WT root 5DAP. G) Ab597-6 root 5DAP stained for GUS. H) Ab552-3 root 5DAP stained for GUS. I) Ab597-6 root 5DAP stained for GUS. J) WT root 7DAP. K) Ab597-6 root 7DAP stained for GUS. L) Ab52-3 root 7DAP stained for GUS. SC - Seed coat, MS - Mature seedling, SR - Seedling root, PR - Primary root, SR - Secondary root, RH - Root hair, RC - Root cap, RM – Root meristem, QC - Quiescent center, V - Vascular bundle, C - Columella, Lat - Lateral root cap, Epi - Epidermis, Cor - Cortex, End - Endodermis, Per - Pericycle

Before germination, no AGL62::GUS expression was observed in seed (data not shown) from the transgenic plants. After 3 days of growth on MS media, the seed coat just began to open and tips of the seedling root became visible. At this stage, still no GUS expression was observed in transgenic seedling (Figure 3.4C). At 4 days after plating, the seedling root elongated and penetrated into the MS media but no GUS expression was observed yet (Figure 3.4E). At 5 days after plating, AGL62::GUS expression was

observed at the tips of root in the root meristem region (Figure 3.4G and 3.4H). Within root meristm region, GUS expression was evident in the quiescent center, pericycle and endodermal initials. (Figure 3.4A, 3.4G, and 3.4H) indicating *AGL62* expression in these tissues. At 5 days of growth on MS media, *AGL62* expression was also observed in root hairs (Figure 3.4I). After 7 days of growth however, AGL62::GUS expression disappeared in both root meristematic region and root hair tissues (Figure 3.4K and 3.4L) indicating *AGL62* expression in root tissues is transient and is found mainly 5 days after germination. After 9 days of growth on MS media, AGL62::GUS transgenic seedlings were stained for GUS expression but no expression was observed (data not shown).

3.5 *AGL62* expression in Seedling shoot

The same 2 T2 transgenic lines that had been used for observing AGL62::GUS expression in seedling roots were used to look at the AGL62::GUS expression in seedling shoots. After 5days of growth on MS media, seed had germinated and formed seedling shoots. No GUS expression, however, was detected in 5-day-old seedling shoots (Figure 3.5B). At 7 days of growth on MS media, the seedling had formed its first set of true leaves. GUS assay of the seedling at this stage showed *AGL62* was strongly expressed in newly formed true leaves (Figure 3.5D, 3.5E and 3.5F). Also vasculature within cotyledon was shown to express *AGL62* at this stage (Figure 3.5D). As the leaves matured, *AGL62* expression in true leaves was no longer detected at 9 days of growth on MS media (Figure 3.5H and 3.5I).

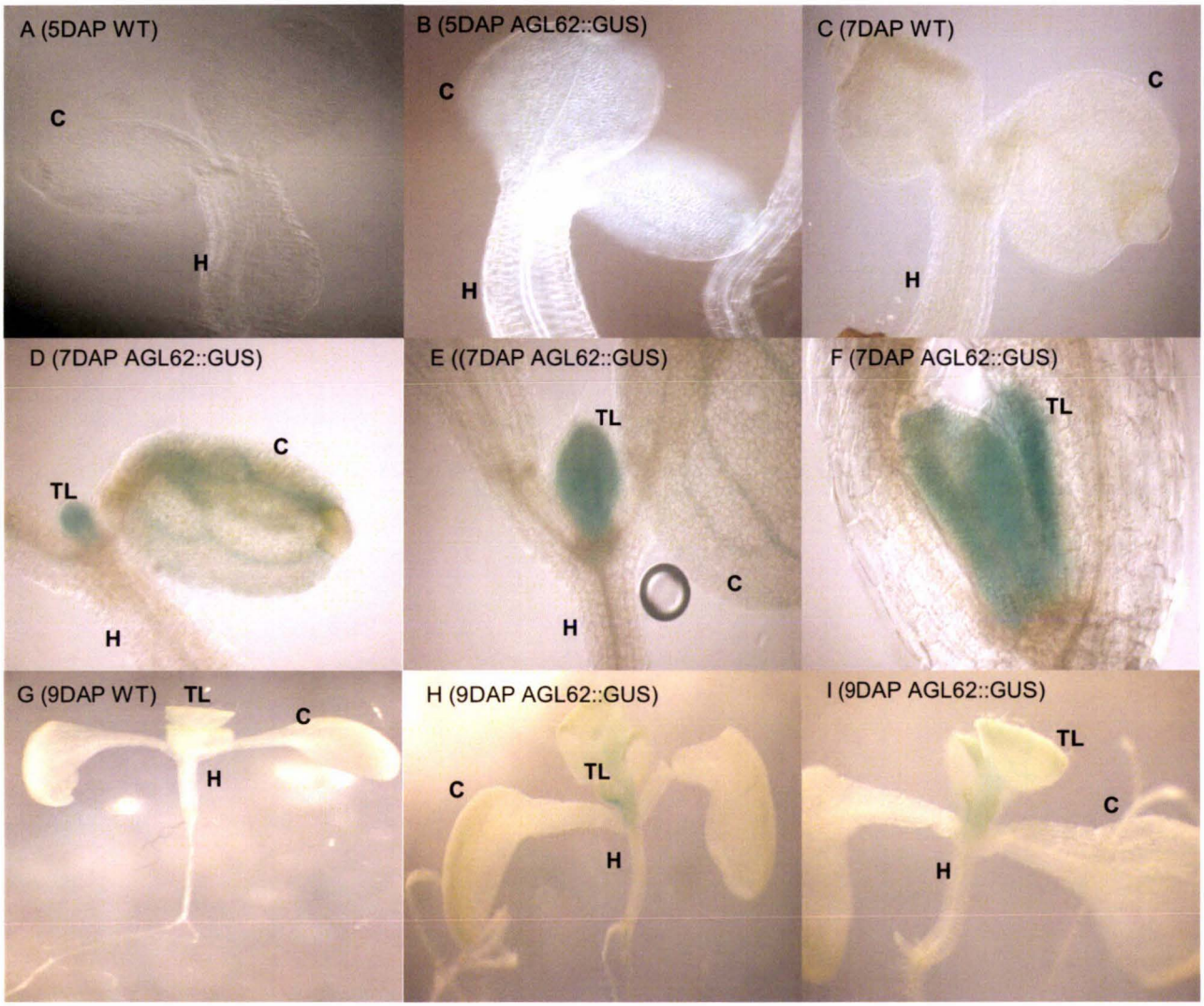


Figure 3.5. GUS assay in AGL62::GUS seedling shoot. A) WT germinating seedling 5days after plating (DAP) on MS medium. B) Ab597-6 seedling 5DAP seedling stained for GUS. C) WT germinating seedling 7DAP. D) Ab552-3 seedling 7DAP seedling stained for GUS. E) Ab597-6 seedling 7DAP seedling stained for GUS. F) Close up photo of E). G) WT germinating seedling 9DAP. H) Ab597-6 seedling 9DAP seedling stained for GUS. I Ab552-3 seedling 9DAP seedling stained for GUS.C - Cotyledon, H - Hypocotyl, TL - True leaves

3.6 *AGL62* expression in pollen

Microarray data (www.genevestigator.ethz.ch/) (table 3.1) showed the highest level of *AGL62* expression in pollen. Mature tricellular pollen contains 2 pollen nuclei and one nucleus will fuse with egg cell to form the embryo and the other will fuse with the central

cell to form the triploid endosperm during double fertilization. Pollen plays a significant role in fertilization and subsequent embryo development and the microarray data (www.genevestigator.ethz.ch/) (table 3.1) indicated possible role of AGL62 in pollen development. Therefore the 2 T2 generation transgenic lines were used to look at the AGL62::GUS expression in developing pollen to investigate whether AGL62 has any role in regulation of pollen development.

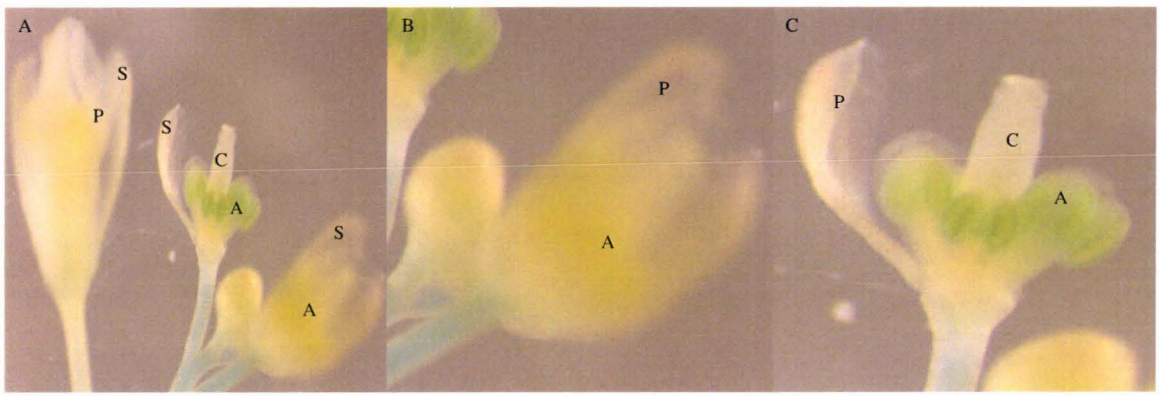


Figure 3.6. GUS assay in developing pollen. A) Pollen developmental stages. B) GUS staining was observed in pollen when the flower was at stage 12. C) When flowers were at stages earlier or later than stage 12, no GUS staining was observed in pollen. D) Close up of figure B. A - Anther, C - Carpel, S - Sepal, P – Petal. The flower showing GUS expression had its sepals and petals removed manually to allow better observation of AGL62::GUS expression in pollen.

After GUS staining (Chapter 2.9.2), AGL62::GUS expression was evident in anther that contained pollen when the flowers were at stage 10 (Smyth et al., 1990) (Figure 3.6B and 3.6D). Sepals and petals were manually removed from the flowers expressing AGL62::GUS in anther to allow better observation of GUS stains in anther. The GUS staining appeared strongest around margins of the anthers, which could be indicating expression of AGL62::GUS in tapetum. However, AGL62::GUS expression in tapetum was transient as pollen from flowers earlier or later stage than stage 10 did not express AGL62::GUS (Figure 3.6C).

3.7 *AGL62* expression in developing Seed

Microarray data (www.genevestigator.ethz.ch/) (table 3.1) also showed modest level of *AGL62* expression in seeds at early stage of development from after fertilization to 4 days after pollination. In double fertilization, egg cell nuclei (n) are fused with pollen nuclei (n) and develop into embryo (2n). The developing embryo is nurtured by endosperm (3n) – a second product of double fertilization that is formed as a result of fusion between pollen nuclei and central cell and continues to grow in seed until maturity. To investigate detailed temporal and spatial expression pattern of *AGL62* in developing seeds, seeds were dissected out from fruits and assayed for GUS expression (Chapter 2.9.2). Seeds were collected from fruits at early stages of development at 0 to 4 days after flowering according to the microarray data (www.genevestigator.ethz.ch/) (table 3.1). The 2 T2 independent lines that showed strongest AGL62::GUS signal were used to look at the AGL62::GUS expression in developing seeds. To enable observation of GUS staining inside the developing seed, GUS stained seeds were cleared in Hoyer's medium (Chapter 2.8.2.4). Seeds collected from flowers that had just opened (0 day after flowering) showed no GUS expression in any part of the seed (Figure 3.7A). However, after 1 day of flowering, seed started to show AGL62::GUS expression in the chalazal region and aleurone layer (Figure 3.7B and 3.7C). At 2 days after flowering *AGL62* expression appeared reduced in the aleurone layer compared to 1day after flowering (Figure 3.7C and D) but GUS staining became stronger in the chalazal region (Figure 3.7C and D). Later, AGL62::GUS expression was no longer detected in the aleurone layer (Figure 3.7E). AGL62::GUS expression in chalazal region persisted in the seed until 3 days after flowering when the embryo was at the late globular stage of development (Figure 3.7F). After 4 days of flowering, the embryo enclosed within the seed reached heart stage of development. At this stage, expression of AGL62::GUS was no longer detected in chalazal region (Figure 3.7G).

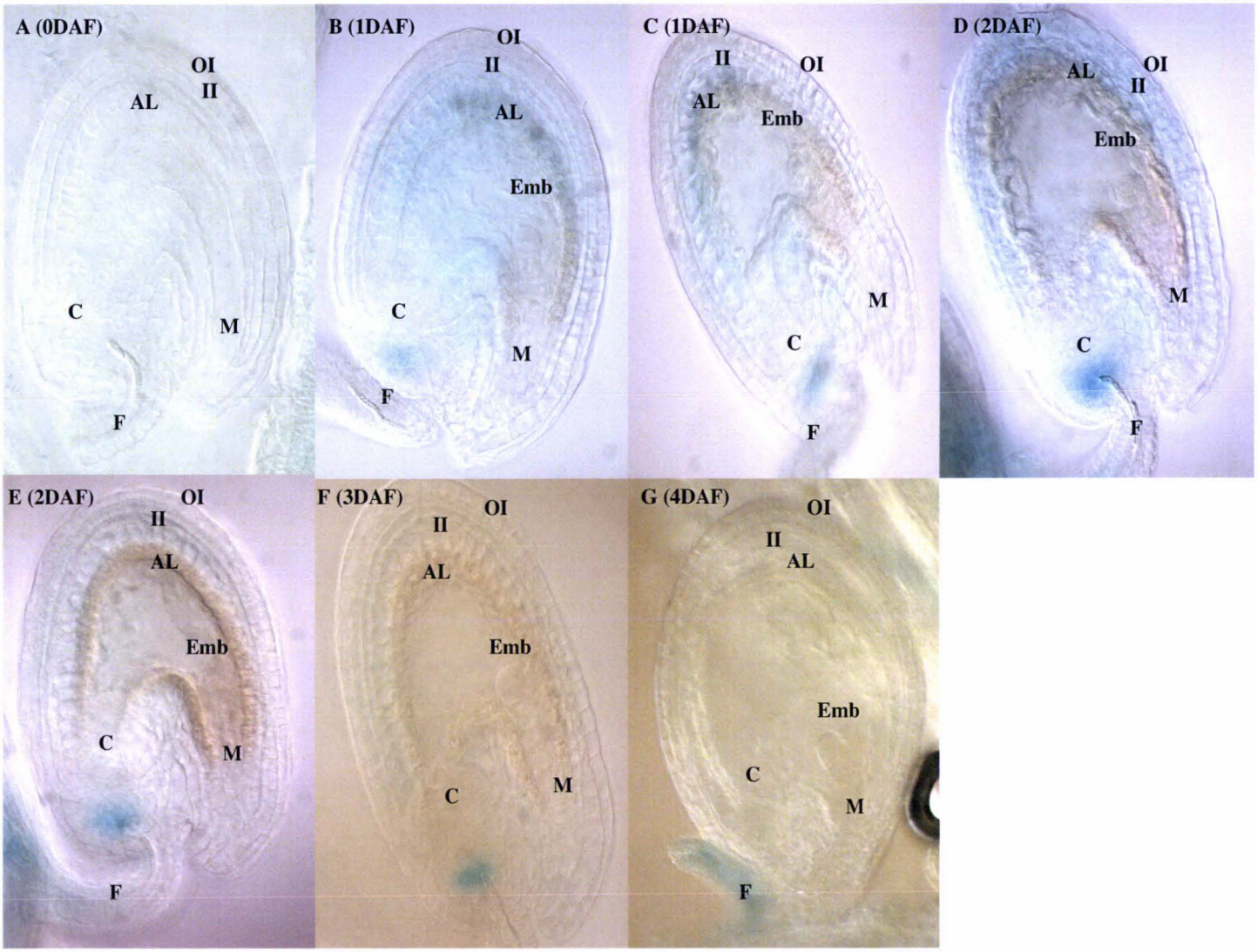


Figure 3.7. AGL62::GUS expression pattern in developing seeds. A) Seeds from flower that has just been opened (0day after flowering (DAF)). B) Seed from 1DAF fruit, embryo is 1-celled. C) Seed from 1DAF fruit, embryo is 2-celled. D) Seed from 2DAF fruit. E) Seed from 2DAF fruit. F) Seed from 3DAF fruit, embryo is at globular stage. G) Seed from 4DAF fruit, embryo reached heart stage. Emb - Embryo, M - Micropyle, C - Chalazal endosperm, F - Funicules, OI - Outer Integument, II - Inner Integument, AL - Aleurone Layer

3.8 Discussion

3.8.1 AGL62::GUS expression in general

AGL62::GUS reporter gene assay was carried out to investigate *in planta* temporal and spatial expression pattern of *AGL62*. Using *in vitro* microarray data (www.genevestigator.ethz.ch/) (table 3.1) as a preliminary data, possible *AGL62* expression domains were identified as seedling root, pollen, inflorescence and seeds at early stages of development. Inflorescence is an organ composed of 4 whorls of organs – sepal, petal, stamen, and carpel. In microarray, only a moderate level of *AGL62* expression was detected in inflorescence but GUS assay showed high level of *AGL62* expression in pollen but not in sepal or petal. Since microarray experiment used inflorescence tissue that consists of 4 whorl organs, moderate level of *AGL62* expression in inflorescence by microarray can be explained if sepals and petals that do not express AGL62 diluted high level of signals from pollen. This indicated that specifically it is the pollen that expresses *AGL62*. Both AGL62::GUS assay and microarray showed that there was no *AGL62* expression in ovule before fertilization so AGL62 is unlikely to have a role in female gametophyte development. However, after double fertilization, *AGL62* expression was observed in seeds 2 to 3 days after the fertilization in chalazal region indicating AGL62 may have a role in seed development.

3.8.2 AGL62::GUS reporter gene construct

AGL62::GUS fusion gene was constructed so that it included 1.5kb upstream and almost all the genomic region of *AGL62* except for the stop codon at the end of exon 2. In this assay intron was included in the AGL62::GUS construct in relation to *AG* – a plant MADS-box gene that was shown to have regulatory sequence within its intron (Sieburth and Meyerowitz, 1997). Although exact promoter region for *AGL62* have not been identified, including 1.5kb upstream region in the construct should be sufficient to cover the AGL62 promoter region. Both microarray and AGL62::GUS assays showed similar *AGL62* expression pattern in root, shoot, pollen and seeds indicating that AGL62::GUS expression pattern likely reflect the endogenous *AGL62* expression pattern. Therefore detailed phenotypic analyses on the specific tissues that express *AGL62* would be required to investigate whether AGL62 plays a role in regulation of development of these tissues.

Initially, 13 AGL62::GUS transgenic plants have been created and expression of AGL62::GUS was assayed in all the 13 plants. However 8 lines showed very weak signals after the GUS staining possibly due to low copy number inserted in the genome. It is also possible that including intron in the AGL62::GUS construct may have reduced the AGL62::GUS signal as some plant MADS-box genes are known to contain the target sequences for negative regulatory proteins in their introns (Sieburth and Meyerowitz, 1997). To investigate whether *AGL62* contains target site for negative regulators in its intron, construct can be made that contains only *AGL62* promoter region fused to the GUS coding region in a future and then compare the expression pattern of this construct to our AGL62::GUS construct.

Only 2 lines showed strong AGL62::GUS signals to be observed therefore these 2 lines that showed the strongest GUS signals were used in all subsequent GUS assays. Although these 2 lines showed the similar expression pattern and their expression patterns were also similar to that of mircoarray's, it may still be important to look at expression of AGL62::GUS in several more AGL62::GUS lines so that possibility of positional effect in which site of insertion of the construct have effect on the expression pattern of the construct can be denied.

3.8.3 AGL62::GUS expression in seedling root and shoot

In seedling roots, AGL62::GUS expression was observed in root meristematic regions specifically in quiescent center (QC), pericycle and endodermal initials (Figure 3.4). QC is located within root meristem and is a group of undifferentiated mitotically inactive cells that divide slowly and give rise to sets of initial cells (Mayer and Jurgens, 1998). The initial cells on the other hand, are mitotically active and differentiate to various cells (Mayer and Jurgens, 1998). During root development, QC regulates balance between the numbers of undifferentiated stem cells and differentiated ones to establish and maintain the root meristem for subsequent root growth (Mayer and Jurgens, 1998). In number of mutants that show defectiveness in QC, root meristem development is disrupted (Mayer et al., 1993). In addition laser ablation surgery showed that disrupting QC cells alters development of the adjacent initial cells and in turn affect root growth (Mayer et al., 1993). Therefore, the expression of AGL62::GUS in the root meristem region particularly in QC indicates possible role of AGL62 in regulation of root development. However,

expression of AGL62::GUS in root meristamatic regions were transient and only observed after 5 days of germination indicating role of AGL62 if any in regulation of root development occur in very short time frame after 5days of germination. Phenotypic analysis on roots from *agl62* mutant would be required to show whether AGL62 truly has a function in root development particularly after 5days of germination. AGL62::GUS expression was also observed in root hairs as well giving rise to possibility that AGL62 has a role in root hair development. In the phenotypic analyses on root, seedling root hairs will be looked at to address this possibility. AGL62::GUS expression was also observed in newly formed leaves in seedlings but the expression was quickly turned off in older leaves (Figure 3.5) indicating role of AGL62 in development of leaves at very early stage but its expression was no longer required for later developmental stages in leaf.

3.8.4 AGL62::GUS expression in pollen

Microarray assays indicated that *AGL62* is expressed highly in developing pollen. Pollen development takes place within an anther that is a part of stamen or the male reproductive organ. During pollen development, Pollen mother cell (PMC) (2n) first undergoes meiosis and give rise to tetrad haploid cells (n) (Bedinger, 1992). Later the haploid cells (n) undergo 2 mitosis to produce mature pollen grain that contains 2 sperm cells (Bedinger, 1992). GUS staining assays also showed AGL62::GUS expression in developing anther from flowers that were at stage10 (Smyth et al., 1990). Anther consists of 4 layers – epidermis, endothelium, middle layer and tapetum (McCormick, 1993). Within anther, strongest AGL62::GUS expression was observed around margins of the anther that corresponds to tapetum. Tapetum is known to have several important roles in microsporogenesis. Tapetum is located adjacent to locules containing pollen mother cells that undergo meiosis (McCormick, 1993). After the meiosis, the newly formed haploid cells (n) are enclosed in callose walls but are released from the wall by the action of callase enzyme secreted from the tapetum layer. The callase enzymes produce and secreted from tapetum breaks down the callose wall of pollen mother cell and releases tetrad haploids as free microspores that undergo subsequent mitoses to grow into mature pollen (McCormick, 2004). Before the release of the microspores, tapetum is also known to play a nutritive role for developing microspores (McCormick, 2004). The third role tapetum plays in pollen development is that it aids the production of precursors for

synthesis of the outer pollen wall – exine (Bedinger, 1992). After completion of the meiosis, pollen cell wall synthesis begins in the haploid cells before they are released as free microspores (McCormick, 1993). The pollen wall consists of 2 layers – inner intine and outer exine (McCormick, 1993). At stage 10 of flower development, flower bud is still closed but tips of petals have just become visible (Smyth et al., 1990). During this stage of flower development, pollens have already undergone meiosis and developing into mature pollens. Since AGL62::GUS expression was observed in tapetum that is known to have several important roles in regulation of pollen development, it is possible that AGL62 may have a function in pollen development. However, anthers from flowers at stages earlier or later than stage 10 showed no AGL62::GUS expression indicating AGL62::GUS expression in tapetum is transient and that possible role of AGL62 in regulation of pollen development may occur specifically when flower is at stage10. Phenotypic analysis will therefore be carried out with pollens from flowers at stage 10 and onwards to study if there is any defect in the developing pollen in *agl62* mutants.

3.8.5 AGL62::GUS expression in seed

Developing seeds at early stages of development was also shown to express AGL62::GUS. The strong expression was observed in chalazal regions located adjacent to funicules 1 to 3 days after pollination. AGL62::GUS expression in this region was observed between 1 to 3 days old seeds but the expression was turned off in later stage seeds indicating possible role of AGL62 in regulation of seed development between 1 to 3 days after pollination. Between 1 to 3 days old seeds, many developmental events occur in seeds mainly in embryo and endosperm. After double fertilization fusion of egg cell (n) and sperm cell nucleus (n) results in formation of 1-celled zygote (2n). The zygote cell elongates followed by asymmetric cell division to generate smaller apical cell and larger basal cell (Faure et al., 2002). Basal cell then undergoes horizontal cell division and give rise to suspensor cells which allows the transfer of nutrients to developing embryo (Faure et al., 2002). On the other hand, smaller apical cell undergoes vertical and horizontal cell division and give rise to 2-celled, 4-celled and eventually to 8-celled octant stage embryo (Faure et al., 2002). At octant stage, apical and basal axis of embryo has already established and upper parts are called upper tier and lower parts are called lower tier respectively (Drews and Yadegari, 2002). Embryo continues to divide and 3days after pollination, embryo reaches late globular stage where number of cell in

embryo exceed 100 (Yadegari and Drews, 2004). On the other hand, endosperm (3n) is also formed by double fertilization as a result fusion between pollen nuclei (n) and central cell nuclei (2n). Endosperm functions as a nutrient reserve for the developing embryo and nurtures the embryo until maturity (Faure et al., 2002). Chalazal endosperm region is located opposite to embryo and is adjacent to funicules that is connecting the developing seeds to the mother fruit tissue. As well as the nutrients obtained from endosperm, chalazal endosperm region allows the transfer of nutrients from maternal tissues to the developing embryo for its growth (Yadegari and Drews, 2004). Chalazal endosperm region therefore plays significant role in embryo development and strong expression of AGL62::GUS in this region suggests possible role of AGL62 in embryo development. Phenotypic analysis on the developing seeds from *agl62* mutant will therefore be required to investigate whether AGL62 has a role in embryo development.

3.9 Conclusion

AGL62::GUS construct was made and GUS staining assays were carried out along with the *in vitro* microarray data (www.genevestigator.ethz.ch/) to compare the expression domains of *AGL62*. The comparison allowed confirmation of the fidelity of the microarray data and AGL62::GUS. GUS staining assays were able to identify more detailed spatial and temporal expression patterns of *AGL62* in root, leaf, anther and seeds. To investigate whether AGL62 has a role in development of any of these organs, phenotypic analyses on these organs in AGL62 loss-of-function mutants will be required.

Chapter 4. Identification of *AGL40* and *AGL62* T-DNA insertion lines

4.1 Introduction

Both AGL40 and AGL62 are known to belong to type I MADS-box clade according to their structural motifs of MADS-domain (Parenicova, de Folter et al. 2003). Approximately 60 out of 108 plant MADS-box genes in *Arabidopsis* belong to type I lineage and type I lineage can be further sub-divided into several subclades based on their sequence similarity (Parenicova, de Folter et al. 2003). Both *AGL40* and *AGL62* belong to Må subclade and share expression domain in inflorescence and fruit (Parenicova, de Folter et al. 2003). Phylogenetic analyses indicate that *AGL40* and *AGL62* are most closely related to each other (Figure 4.1). In addition MADS-box genes tend to function in a site of expression (De Bodt, Raes et al. 2003). Furthermore, compared to other model organisms, *Arabidopsis* is known to be gene-rich due to relatively high rates of gene duplication in its genome (Arabidopsis Genome 2000). Gene duplication is a process where DNA sequence that contains gene is duplicated possibly due to mistakes in homologous recombination, retro transposition event or duplication of entire chromosome (Arabidopsis Genome 2000). The second copy of the gene produced as a result of gene duplication is often free from selective pressure (Arabidopsis Genome 2000). In other words, mutation of the copy has less or no deleterious effects to its host organism. Thus the copy mutates faster than a functional original copy gene and over time, it results in the copy obtaining new or at least partially different function – a process called functionalization (Arabidopsis Genome 2000). Similarly, it is possible that a gene can become a pseudo gene in which a gene loses its function because it lost the sequence required for the gene to be expressed (Arabidopsis Genome 2000). In *Arabidopsis*, half of its genome is covered by gene duplication and since *AGL40* and *AGL62* share similar sequence and expression pattern, AGL40 and AGL62 may have risen from a common ancestor and therefore they may share redundant activities in regulation of plant development.

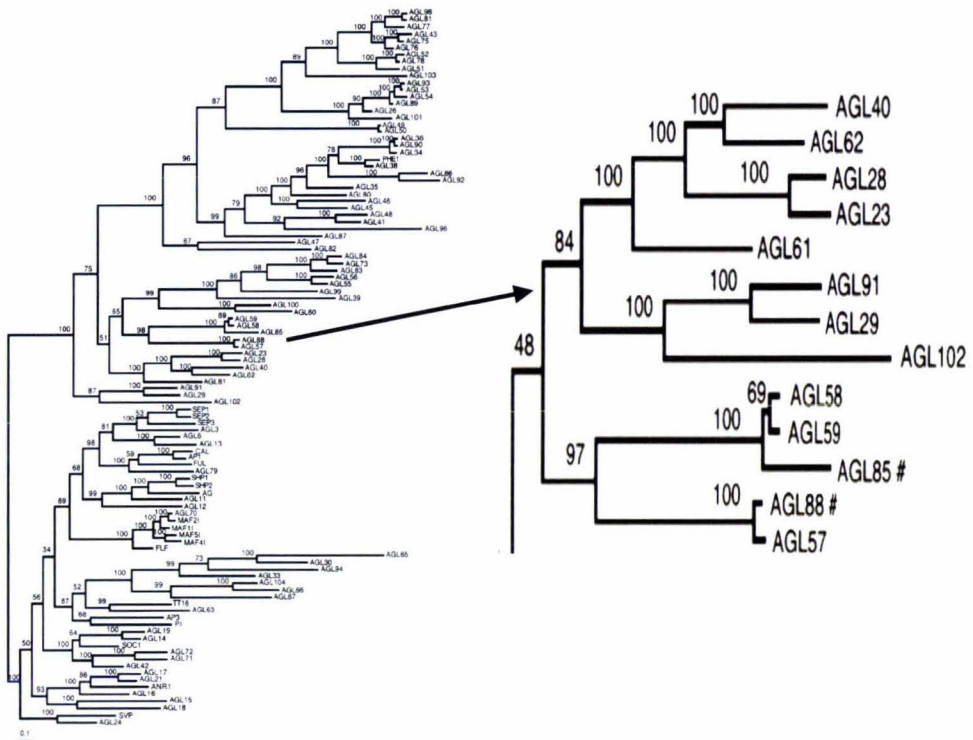


Figure 4.1. *Arabidopisis thaliana* MADS-box gene phylogenetic tree. Taken and modified from (Martinez-Castilla and Alvarez-Buylla 2004). Left tree - There are 108 MADS-box genes in *Arabidopsis* and 64 of them are known to belong to type I lineage according to their structural motifs in MADS-domain. Right tree - Mα subclade of type I MADS-box genes showing *AGL40* and *AGL62* are most closely related to each other. Arrow indicates the Mα subclade of type I MADS-box genes in all MADS-box gene phylogenetic tree. Numbers at the top of branches indicates the number of nucleotide changes that had occurred.

To understand the role of AGL40 and AGL62 may play in plant development, functional analyses of putative loss of function alleles of *AGL40* and *AGL62* were required. T-DNA insertion mutant lines are publicly available and commonly used for functional characterization of genes in *Arabidopsis* (www.signal.salk.edu/). T3 generation T-DNA insertion lines for both *AGL40* and *AGL62* were obtained from SALK institute T-DNA collections (http://signal.salk.edu/index.html) and used in subsequent molecular and phenotypic studies. The SALK database was screened for T-DNA insertions in *AGL40* (At4g36590) and *AGL62* (At5g60440). SALK line 107011 showed a T-DNA insertion in *AGL40* while SALK line 022148 showed a T-DNA insertion in *AGL62* (Figure 4.2). These 2 lines were obtained from the Arabidopsis biological resource center (ABRC) (www.biosci.ohio-state.edu/pcmb/Facilities/abrc/abrchome.htm). Gene specific primers

and a primer for the left border of T-DNA (LBb1) were designed (Table 2.1) and used to identify the T-DNA in each locus as a first step in functional characterization of AGL40 and AGL62.

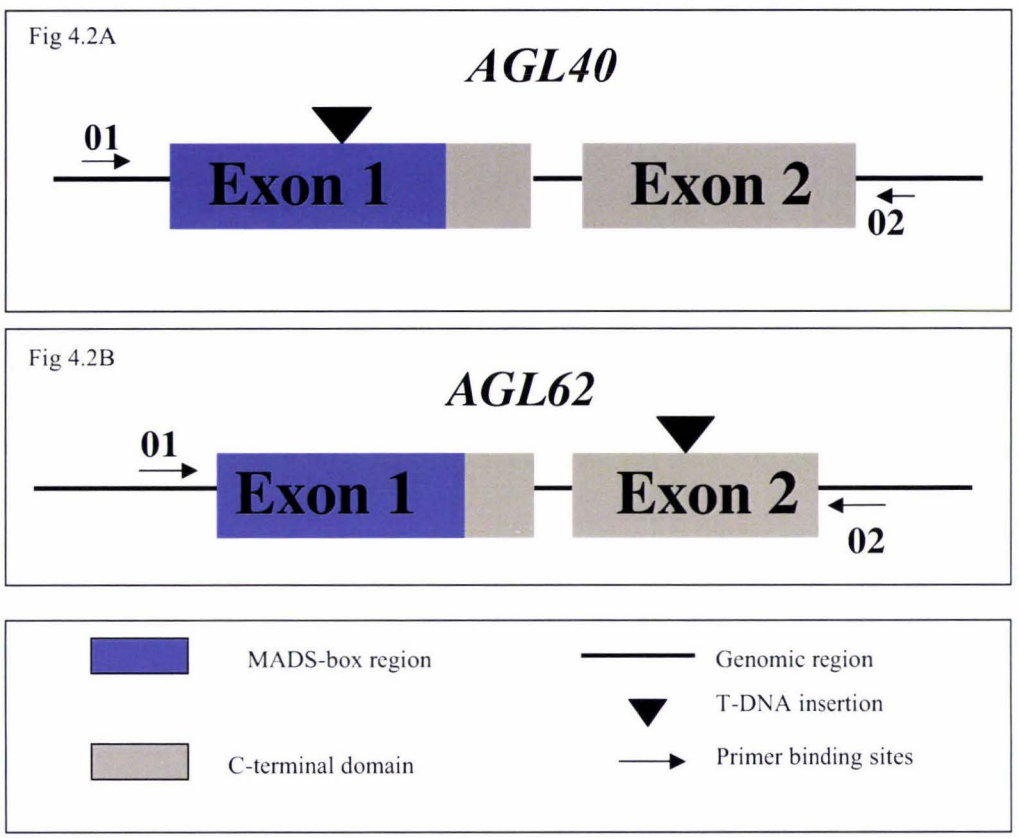


Figure 4.2. Schematic diagram of PCR design to amplify *AGL40* and *AGL62* sequences using combination of gene specific primers and T-DNA primer. Boxes represent exons. Blue box represents the MADS-box domain and gray boxes represent C-terminal domain. Thick black lines indicate introns and genomic regions. Triangle represents the T-DNA insertion and arrows indicate the primer-binding site. Approximate site of T-DNA insertion was estimated by SALK institute (http://signal.salk.edu/index.html) where the T-DNA insertion lines had been obtained. In figure 4.2A, 01 – 001AGL40, 02 – 002AGL40 while in figure 3.2.B, 01 – 001AGL62, 02 – 002AGL62. In both cases, Lb – LBb1

4.2 Confirmation of T-DNA insertion in *AGL40* by PCR genotyping

AGL40 is a member of type I MADS-box gene that belongs to Ma subclade (Parenicova, de Folter et al. 2003). *AGL40* share similar expression pattern with *AGL62* in inflorescence and fruit (Table 4.1) (Parenicova, de Folter et al. 2003) and is most closely related to *AGL62* (Figure 4.1).

Tissue	*AGL40* expression	*AGL62* expression
Root	Absent	Present
Leaf	Absent	Present
Inflorescence	Present	Present
Fruit	Present	Present

Table 4.1 Expression domains of *AGL40* and *AGL62*. Data obtained from (Parenicova, de Folter et al. 2003). Expressions of *AGL40* and *AGL62* in root, leaf, inflorescence and fruit were checked by RT-PCR and tabulated as above. If expression was observed, it was described as present while if expression was not detected, it was described absent.

To determine the role AGL40 may play in plant development, lines homozygous for a T-DNA insertion in *AGL40* needs to be identified. As a first step in identifying *AGL40* homozygous mutant PCR genotyping was carried out to confirm the insertion of T-DNA fragment in *AGL40* as described in figure 4.2A. SALK institute T-DNA collections (http://signal.salk.edu/index.html) indicated the insertion of T-DNA in exon 1 in the MADS-box domain of *AGL40*. Genomic DNA was extracted from 10 *AGL40* T-DNA insertion lines and using a combination of gene specific primer 001AGL40 and T-DNA border primer LBb1, approximately 550bp T-DNA band was amplified from *AGL40* T-DNA insertion lines (Figure 4.3).

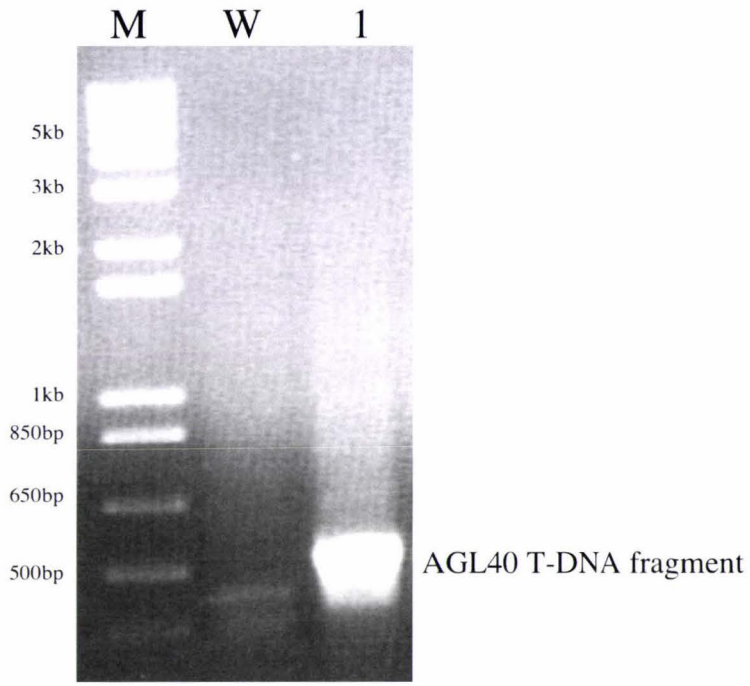


Figure 4.3. PCR genotyping confirmation of T-DNA insertion in *AGL40*. Using 001AGL40 and LBb1 primers, approximately 550bp band was amplified from *AGL40* T-DNA insertion line genomic DNA, which was absent in WT plants confirming the presence of T-DNA insertion in *AGL40* T-DNA insertion lines. M - 1kb plus DNA marker, W – WT control, 1 – *AGL40* T-DNA insertion plant

4.3 *AGL40* T-DNA fragment sequencing

To verify that that the 550bp fragment amplified was in fact a T-DNA insertion in *AGL40*, the 550bp PCR product was gel purified, cloned into pBSKS cloning vector EcoRV site and transformed into *E. coli* using heat shock method. Recombinant *E. coli* were selected, the plasmid isolated and restriction digested with EcoRV. EcoRV released full-length 550bp insert from the recombinant vector and confirmed the presence of 550bp insert in the pBSKS (Figure 4.4).

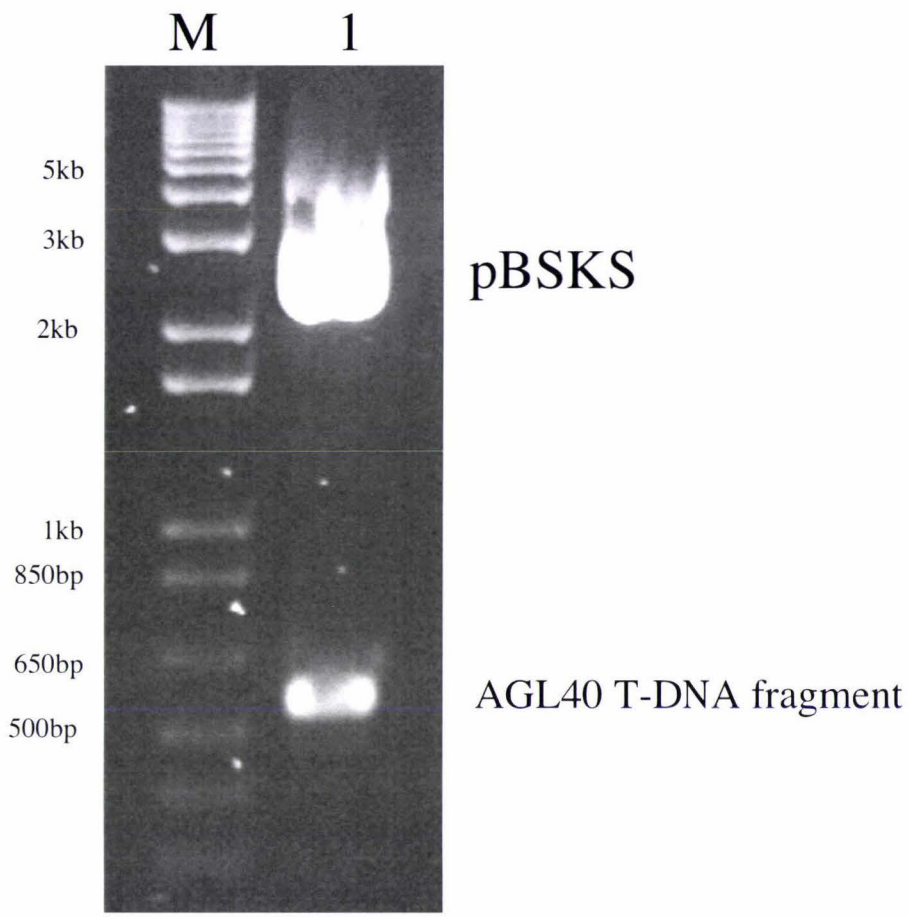
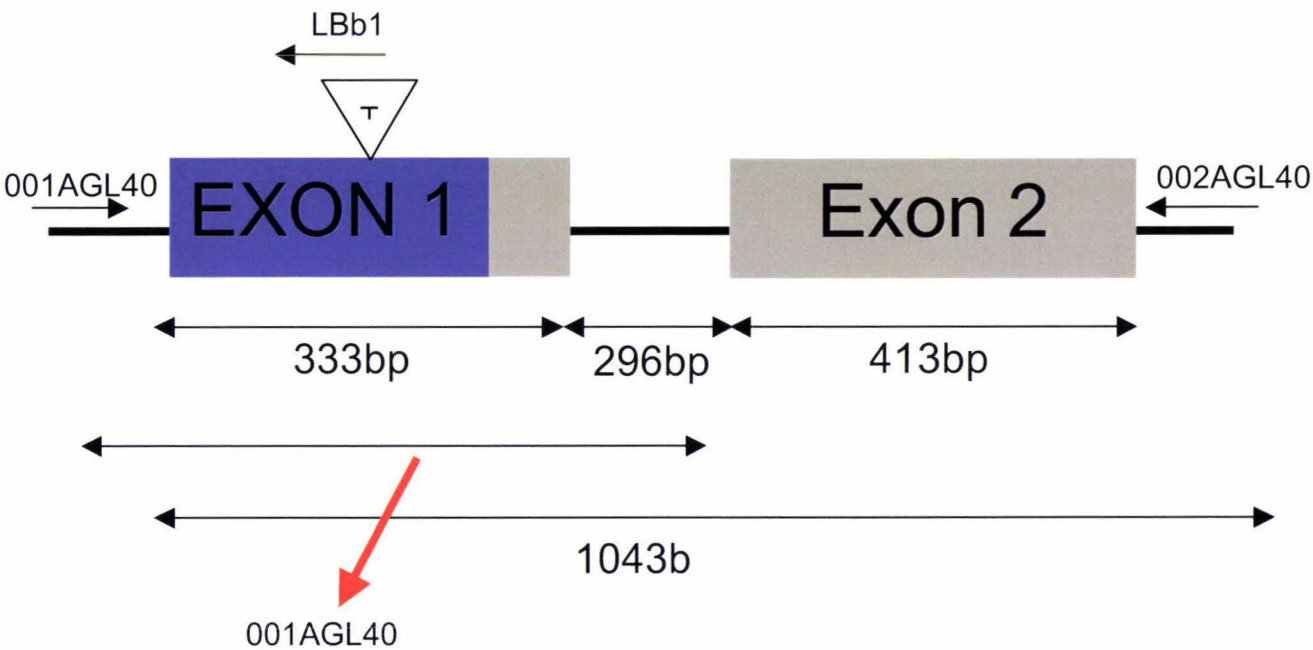


Figure 4.4. Restriction digestion of recombinant pBSKS for confirmation of *AGL40* T-DNA insert in the vector. Selected recombinant vectors were isolated and restriction digested with *Eco*RV to release the 550bp *AGL40* T-DNA fragment insert from 3kb pBSKS cloning vector confirming the presence of insert in vector. M – 1kb plus DNA marker, 1 – recombinant pBSKS digested with *Eco*RV.

The recombinant vector was purified by using Quantum Miniprep kit (Bio-Rad) and then sequenced to confirm the insertion of T-DNA in *AGL40* and to identify the precise location of T-DNA insertion in *AGL40* to state that any phenotypic defects observed in *agl40* mutant is likely due to insertion an in *AGL40*. The purified plasmid was sent to Allan Wilson center (AWC) (http://awcmee.massey.ac.nz/) for sequencing. After sequencing, insertion of T-DNA in *AGL40* was confirmed as well as precise site of T-DNA insertion in MADS-domain (Figure 4.5). The line in which T-DNA insertion in *AGL40* was confirmed by sequencing was designated as *agl40-1* and its segregating

population was used in subsequent studies involving molecular and phenotypic studies on *agl40* mutant.



GAGAAAATGG TGAGAAGTAC CAAAGGTCGT CAGAAAATAG AGATGAAAAA
AATGGAAAAC GAAAGCAACC TTCAGGTTAC TTTCTCAAAA AGAAGATTCG
GTCTTTTCAA AAAAGCTAGT GAACTTTGCA CATTAAGTGG TGCAGAGATT
CTGTTGATTG TGTTCTCTCC TGGTGGGAAA GTGTTTTCTT TTGGCCATCC
AAGTGTTCAA GAACTCATTC ATCGCTTTTC GAATCCTAAC CATAATTCTG
CCATTGTCCA TCATCAGAAC AACAATCTCC AACTTGTTGA AACCCGTCCG
GATAGAAATA TCCAATATCT CAACAATATA CTCACTGAGG TAAACAATTT
TGTTTTAAAG AAAAATGCAA GAGAATTATA TATATA*TATT GTGGTGTAAA
CAAATTGACG CTTAGACAAC TTAATAACAC ATTGCGGACG TTTTTAATGT
ACTGGGGTGG TTTTTCTTTT CACCAGTGAG ACGGGCAACA GCTGATTGCC
CTTCACCGCC TGGCC**CTGAG AGAGTTGCAG CAAGCGGTCC ACGC**

LBb1

Figure 4.5. Diagram of T-DNA insertion site in *AGL40* confirmed by sequencing. Boxes represent exons. Blue box represents the MADS-box domain and gray boxes represent C-terminal domain. Introns and genomic regions are drawn as lines. Triangle shows the site of T-DNA insertion and black arrows indicates the primer-binding site. Sizes of each region are indicated below the regions flanked by line with double arrowheads. Red arrow points to the *AGL40* exon 1 where T-DNA was inserted and its sequence was given below. * Indicates the insertion of T-DNA at second exon of *AGL40*. Sequences after the * mark represents the sequences from *AGL40* exon 2 while sequences before * corresponds to that of T-DNA. Sequences with underline are the DNA pieces where the corresponding primers were binding.

4.4 Identification of homozygous *AGL40* T-DNA insertion lines

To obtain a plant homozygous for the T-DNA insertion in *AGL40* in a segregating population of *agl40-1*, 10 seeds were plated and PCR genotyping using 01AGL40, 02AGL40, and LBb1 primers was carried out to distinguish between homozygous T-DNA insertion and heterozygous T-DNA insertion lines (Figure 4.2A). Using primers 001AGL40 and LBb1 in PCR, the primers should only amplify the 550bp T-DNA fragment from T-DNA insertion lines and therefore separate all the AGL40 T-DNA insertion lines from WT lines. In a second set of PCR using forward gene specific primer (001AGL40) and reverse gene specific primer (002AGL40), heterozygous T-DNA insertion lines and homozygous T-DNA insertion lines are distinguished because in homozygous T-DNA insertion line, *AGL40* PCR amplification should be abolished due to insertion of large fragments of T-DNA in both DNA strands whereas in heterozygous T-DNA insertion lines, approximately 1kb WT *AGL40* band should still be amplified. After PCR genotyping, 2 individuals out of 9 surviving plants that were genotyped showed amplification of 550bp T-DNA band and no WT 1kb *AGL40* band indicating that these 2 lines were homozygous for the T-DNA insertion in *AGL40* (Figure 4.6).

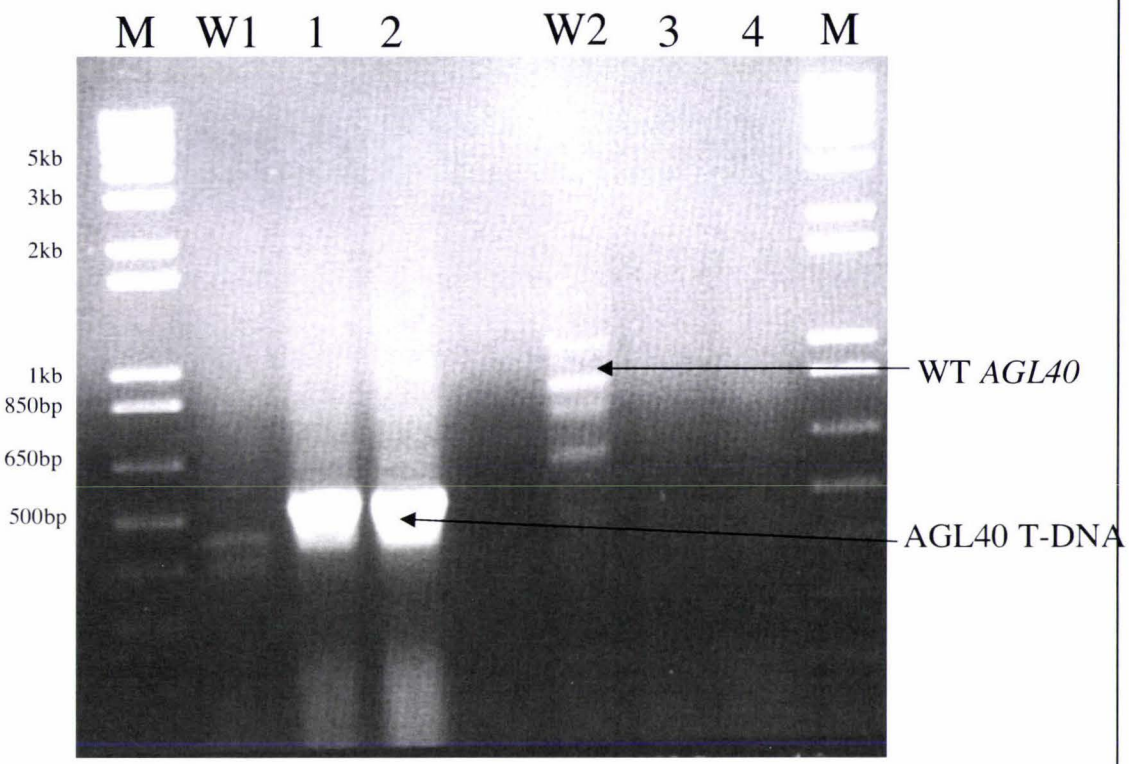


Figure 4.6. PCR genotyping identification of homozygous *AGL40* T-DNA insertion mutant. 2 AGL40 T-DNA insertion mutants were shown to have T-DNA insertion in *AGL40* as evident by PCR amplification of 550bp band using 001AGL40 and LBb1 (lanes 1adn 2). The same 2 lines were genotyped using 001AGL40 and 002AGL40 and shown to have homozygous T-DNA insertion in *AGL40 (lanes 3 and 4).*

As *AGL40* is predominantly expressed in the inflorescence (Table 4.1), total RNA was extracted from WT and *agl40-1*inflorescence tissues and RT-PCR was carried out using gene specific primers 001AGL40 and 002AGL40 to see if plants homozygous for the T-DNA insertion in *AGL40* (*agl40-1/agl40-1*) result in a loss-of-function for AGL40. As a control to see that RT-PCR is working, expression of 400bp ubiquitin was tested using ubiquitin primers 001UBQ and 002UBQ in the same RT-PCR. RT-PCR was unable to detect the WT 750bp *AGL40* cDNA band in homozygous T-DNA insertion lines indicating that *agl40-1* lines with homozygous T-DNA insertion in *AGL40* were likely AGL40 loss-of-function mutants (figure 4.7).

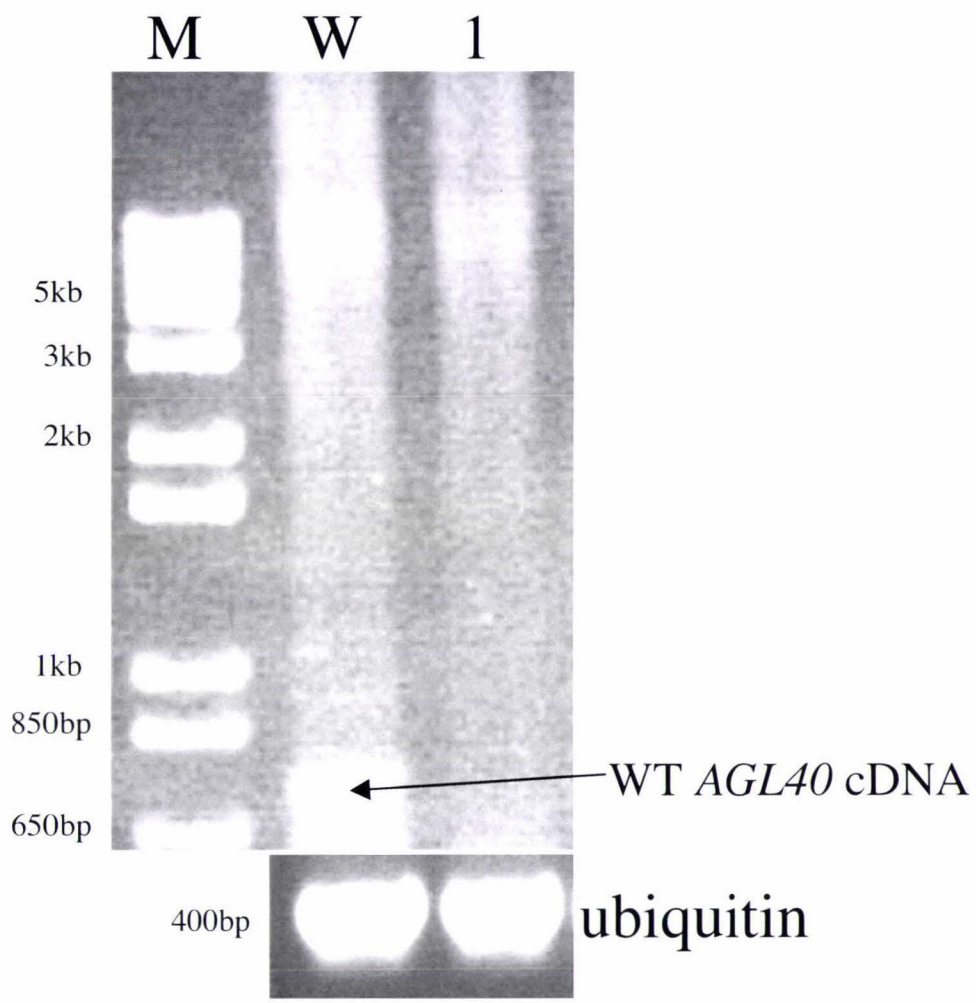


Figure 4.7. RT-PCR confirmation of loss of *AGL40* expression in homozygous *AGL40* T-DNA insertion mutant. Top gel - RT-PCR was carried out with total RNA extract from WT and homozygous *AGL40* T-DNA insertion mutant using 001AGL40 and 002AGL40. In homozygous *AGL40* T-DNA insertion mutant, 750bp *AGL40* cDNA was not amplified indicating loss of AGL40 expression in homozygous *AGL40* T-DNA insertion mutant. M - 1kb plus DNA marker, W - WT, 1 - homozygous *AGL40* T-DNA insertion mutant. Bottom gel - Loading control (400bp ubiquitin)

Gross morphological analyses including observation of inflorescence structure, fruit morphology, and seed numbers were carried out with the *agl40-1* AGL40 loss-of-function mutant lines to observe if there is any phenotype caused by loss of AGL40. However, the phenotypic analyses showed no phenotypic defects in agl40 mutants (data not shown). It is possible that AGL40 has redundant or overlapping activities with AGL62 in *Arabidopsis* plant development.

4.5 Confirmation of T-DNA insertion in *AGL62* by PCR genotyping

Confirmation of T-DNA insertion in *AGL62* was necessary to show that the T-DNA insertion was in fact in *AGL62* and that the gene specific primers worked for PCR genotyping. To this end, PCR based genotyping using a combination of gene specific primer (02AGL62) and left border primer (LBb1) designed for T-DNA insertion was carried out to show the presence of T-DNA fragments in *AGL62* as shown in figure 4.2B. From SALK institute T-DNA collections (http://signal.salk.edu/index.html), we expected the T-DNA sequence to be inserted in exon 2 of *AGL62*. Genomic DNA was extracted from 5 *AGL62* T-DNA insertion lines and using combination of 02AGL62 gene specific primer and LBb1 T-DNA left border primers, approximately 700bp band was amplified from *AGL62* T-DNA insertion lines that was absent from the WT plants (Figure 4.8). There were also 1kb band and approximately 500bp bands amplified from *AGL62* T-DNA insertion lines but these bands were also amplified in WT plants indicating that these bands were amplified as a result of the primers' non-specific binding elsewhere in the genome.

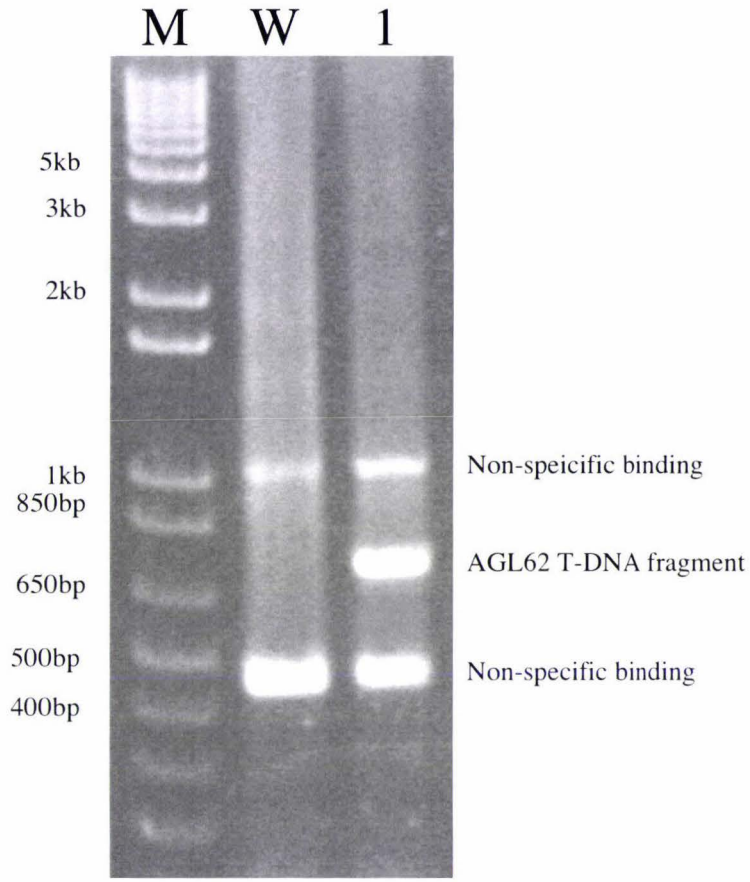


Figure 4.8. PCR genotyping confirmation of T-DNA insertion in *AGL62*. Gene specific primer 002AGL62 and T-DNA left border primer LBb1 were used in PCR to amplify *AGL62* T-DNA insertion fragment from the *AGL62* T-DNA insertion lines. 500bp band and 1kb band observed in both WT and AGL62 T-DNA insertion lines were due to mispriming of primers elsewhere on genomic DNA. 700bp band amplified only in *AGL62* T-DNA insertion line represents the *AGL62* T-DNA fragment.

The 700bp band was then gel purified, cloned into pBSKS cloning vector at *Eco*RV site and transformed into *E. coli* using heat shock method. Recombinant *E. coli* was selected and the recombinant vector was isolated from *E. coli*. To confirm that the vector contained fragment of interest, restriction digest with EcoRV and *Hind*III was carried out. EcoRV was used to release the inserts from vector and HindIII cuts approximately in the middle of exon 2 and thereby fragmenting 700bp band into 2 pieces. Double restriction digest fragmented approximately 700bp AGL62 T-DNA insert into approximately 400 and 300bp bands and released the fragments from the 3kb pBSKS vector (Figure 4.9).

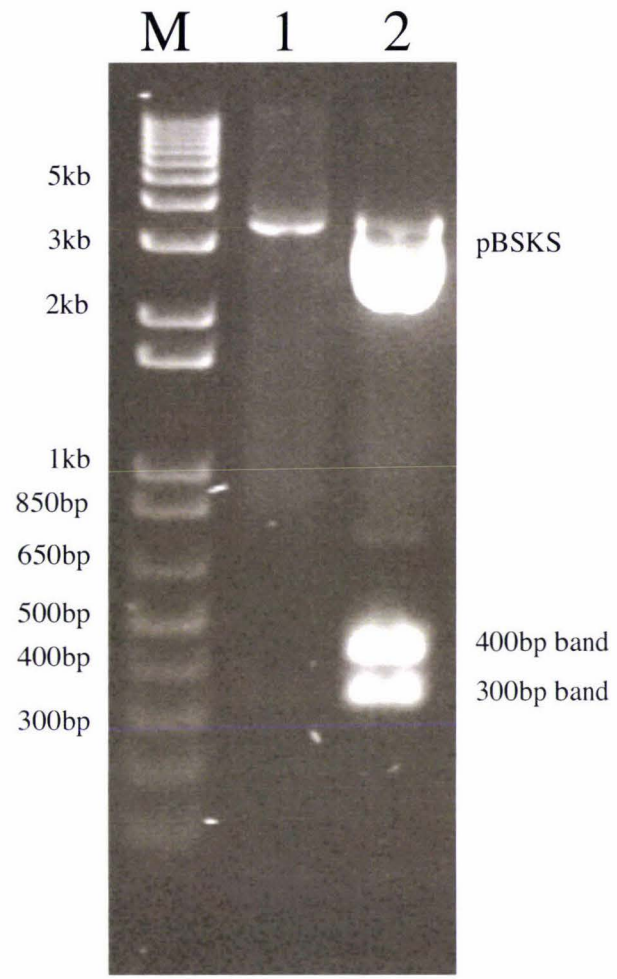
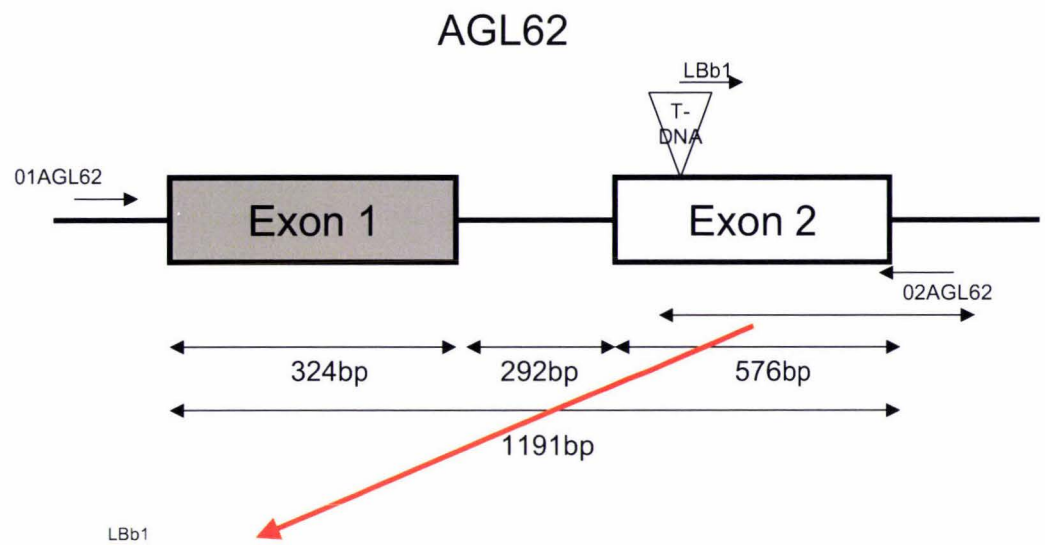


Figure 4.9. Double restriction digestion of recombinant vector to confirm insertion of *AGL62* T-DNA. 3kb pBSKS was double restriction digested with *Eco*RV and *Hind*III to release *AGL62* T-DNA insert. Double digestion with *Eco*RV and *Hind*III fragmented approximately 700bp fragment into 400bp and 300bp DNA pieces. M - 1Kb plus DNA ladder, 1 - undigested pBSKS, 2 - pBSKS double digested with *Eco* RI and *Hind* III

Confirmation of the T-DNA insertion in *AGL62* by sequencing was necessary to identify the exact site of T-DNA insertion in *AGL62* and to state that any phenotypic defects observed in *agl62* mutant is likely due to insertion in *AGL62*. Therefore the recombinant vector was purified by using Quantum Miniprep kit (Bio-Rad) and then sent to Allan Wilson center (AWC) (http://awcmee.massey.ac.nz/) for sequencing.



LBb1

CGAT**GCGTGG ACCGCTTGCT GCAACT**CTCT CAGGGCCAGG CGGTGAAGGG CAATCAGCTG TTGCCCGTCT CACTGGTGAA
AAGAAAAACC ACCCCAGTAC ATTAAAAACG TCCGCAATGT GTTATTAAGT TGTCTAAGCG TCAATGTGTT TATTTGTATT
TATTTG * TAGA AACAGAGAAA AAGAAGTACG ACGAGTTAAA GAAAATAAGA GAAAAGACAA AAGCCCTTGG GAATTGGTGG
GAAGATCCCG TTGAGGAACT TGCGTTATCT CAACTCGAGG GGTTCAAAGG TAATCTTGAA AATTTGAAGA AAGTAGTTAC
AGTCGAAGCT TCCAGATTTT TTCAGGCAAA TGTTCCAAAC TTCTATGTGG GAAGTTCTAG TAATAATGCT GCTTTTGGGA
TTGATGATGG TAGTCATATC AACCCTGATA TGGATCTCTT TAGCCAAAGA AGAATGATGG ACATAAATGC CTTCAACTAC
AACCAGAACC AAATTCACCC TAATCATGCA TTACCACCCT TTGGAAACAA TGCTTATGGT ATTAATGAAG GGTTTGTTCC
AGAATACAAT GTGAACTTCA GACCAGAGTA TAACCCAAAC CAAAACCAAA TCCAAAACCA AAATCAAGTT CAAATCCAAA
TCCAAAACCA GAGTTTTAAG AGAGAAAACA TCTCTGAATA TGAACATCAT CATGGKTATC CTCCCCAGTC TAGATCTGAT
TACTATTAA **CCCTTTGTTG TGTGATCCAA C**ATCA

02AGL62

Figure 4.10. Diagram of T-DNA insertion site in *AGL62* confirmed by sequencing. Boxes represent exons. Blue box represents the MADS-box domain and gray boxes represent C-terminal domain. Introns and genomic regions are drawn as lines. Triangle shows the site of T-DNA insertion and black arrows indicates the primer-binding site. Sizes of each region are indicated below the regions flanked by line with double arrowheads. Red arrow points to the *AGL62* exon 2 where T-DNA was inserted and its sequence was given below. * Indicates the insertion of T-DNA at second exon of *AGL62*. Sequences after the * mark represents the sequences from *AGL62* exon 2 while sequences before * corresponds to that of T-DNA. Sequences with underline are the DNA pieces where the corresponding primers were binding.

After sequence analysis, the T-DNA insertion in *AGL62* and the precise location of insertion was confirmed (Figure 4.10). The T-DNA was shown to be at the start of exon 2 of AGL62 as expected from the SALk institute T-DNA data (http://signal.salk.edu/index.html). After the confirmation of T-DNA insertion in *AGL62* by sequencing, the line was designated as agl62-1 and its segregating populations were used in all subsequent molecular and phenotypic studies.

4.6 Identification of *AGL62* loss-of-function mutant

Preliminary PCR analyses of *agl62-1* indicated that the SALK line was segregating for the T-DNA insertion in *AGL62*. To identify a putative AGL62 loss-of-function mutant line, segregating population of *agl62-1* was screened for plants with homozygous T-DNA insertion in *AGL62*. In a segregating population of *agl62-1*, 20 seeds were planted on soil

and PCR genotyping was carried out with the surviving plants to identify *agl62-1* homozygous plants. PCR genotyping using 01AGL62, 02AGL62, and LBb1 primers was carried out to distinguish between homozygous T-DNA insertion and heterozygous T-DNA insertion (Figure 4.2B). Since primer combinations 002AGL62 and LBb1 amplifies 700bp T-DNA fragment inserted within *AGL62*, PCR genotyping reactions using these primers should only amplify the T-DNA fragments from plants carrying T-DNA and thereby separate all the *AGL62* T-DNA insertion lines from WT lines. In second sets of PCR using forward gene specific primer (001AGL62) and reverse gene specific primer (006AGL62), heterozygous T-DNA insertion lines and homozygous T-DNA insertion lines were distinguished because in homozygous T-DNA insertion line, *AGL62* PCR amplification should be abolished due to insertion of large fragments of T-DNA in both DNA strands whereas in heterozygous T-DNA insertion lines, approximately 1.5kb WT *AGL62* band should still be amplified. After PCR genotyping the 15 individuals from a segregating population of *agl62-1* with both sets of primers, 12 out of 15 individuals showed amplification of 700bp T-DNA bands using primers 002AGL62 and LBb1 indicating these individuals carry T-DNA in *AGL62* (Figure 4.11A). However, second set of PCR genotyping using 001AGL62 and 006AGL62 amplified 1.5kb WT *AGL62* band from all the 15 individuals, which indicated all the individuals had heterozygous T-DNA insertion in *AGL62*. (Figure 4.11B).

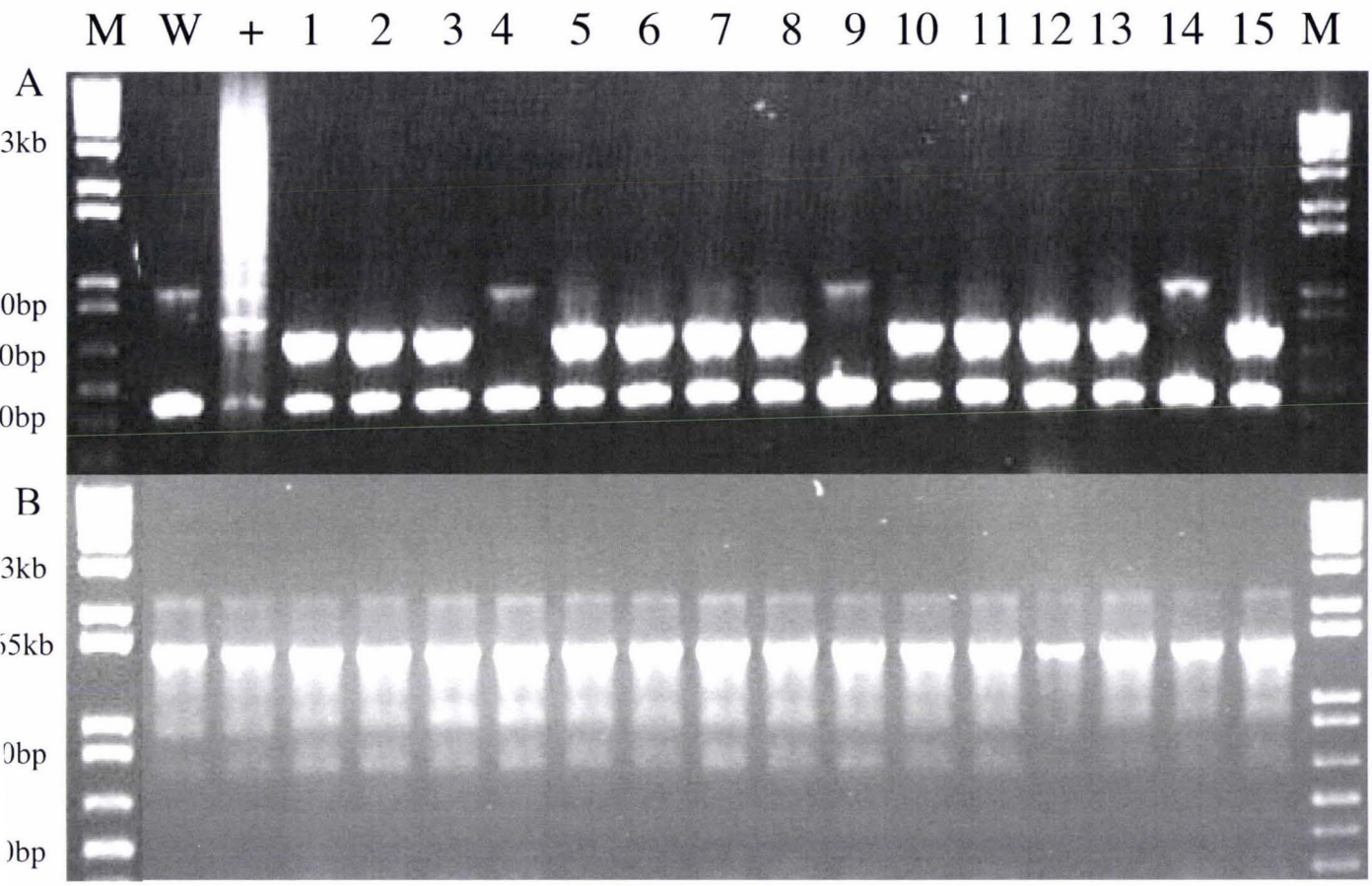


Figure 4.11. PCR genotyping for distinguishing between heterozygous and homozygous *AGL62* T-DNA insertion lines. Top gel - 15 individual plants were selected from *AGL62* T-DNA insertion line segregating population and PCR genotyped for T-DNA insertion using gene specific primer 002AGL62 and T-DNA border primer LBb1. If T-DNA fragment is inserted in *AGL62,* approximately 700bp band should be amplified. Bottom gel - Same individuals that had been PCR genotyped were PCR genotyped again with gene specific primers 001AGL62 and 006AGL62 to distinguish between heterozygous and homozygous T-DNA insertion lines. All individuals show same banding pattern as WT indicating all the plants tested were heterozygous T-DNA insertion lines. M - 1kb plus DNA marker, W - WT control, + - positive (known agl62-1) control, 1 - 15 - 15 individual plants selected from segregating population of agl62-1.

In an attempt to recover homozygous *AGL62* T-DNA insertion lines, we suspected that combination of primers 001AGL62 and 006AGL62 may not be an ideal combination as amplification of *AGL62* bands achieved with these 2 primers were not consistent (data not shown). Therefore, another gene specific primer, 005AGL62, was designed and in combination with 002AGL62 and LBb1, used to distinguish heterozygous and homozygous T-DNA insertion lines. However, homozygous *AGL62* T-DNA insertion

line was still not recovered. Larger scale PCR genotyping was then carried out as before with the 151 plants from the segregating population of *agl62-1* to observe if any homozygous line can be recovered (Table 4.2) (See Appendix 2). The table showed that for approximately 150 PCR genotyping reactions, no homozygous T-DNA insertion mutant was recovered (Appendix 2). Instead, the PCR genotyping showed out of 151 plants from a segregating population of *agl62-1*, 49 of them was WT and the other 102 plants carried heterozygous T-DNA insertion in *AGL62*. This ratio of heterozygous T-DNA insertion:WT in segregating population was 2:1 instead of the normal 1:2:1 indicating loss-of-function-mutation of *AGL62* may result in embryonic lethal phenotype.

No. WT [a]	No. T-DNA insertion line [a]	No. Heterozygous T-DNA insertion line [B]	No. Homozygous T-DNA insertion line [b]
49	102	151	0

Table 4.2. Large scale PCR genotyping result. 151 individuals from *AGL62* T-DNA insertion line were PCR genotyped to identify homozygous T-DNA insertion line. Using gene specific primer 002AGL62 and T-DNA border primer LBb1, WT and T-DNA insertion line were separated. Then using forward and reverse gene specific primers 001AGL62 and 006AGL62, heterozygous and homozygous T-DNA insertion was distinguished. No homozygous T-DNA insertion in *AGL62* was recovered and ratio of WT to heterozygous T-DNA insertion line was 1:2. a – primers used were 002AGL62 and LBb1, b – primers used were 001AGL62 and 006AGL62.

4.7 Production of *agl40;agl62* double knock out mutant

Lack of any phenotype in AGL40 loss-of-function mutant *agl40-1* indicated redundancy might exist between *AGL40* and other closely related members of plant type I MADS-box genes. Since the most closely related gene to *AGL40* is *AGL62*, *agl40;agl62* double knock out mutant was produced to study the effect knocking out the functions of AGL40 and AGL62 have on plant development and thereby functionally characterize AGL40 and AGL62.

AGL40 loss-of-function line *agl40-1* was crossed with *AGL62* heterozygous T-DNA insertion line *agl62-1* (Chapter 2.8.1) and F1 seeds were collected. F1 seeds were PCR genotyped for T-DNA insertion in *AGL62* using primers 002AGL62 and LBb1 (data not

shown) as before (see Figure 4.8 for example). No PCR genotyping was carried out for confirmation of T-DNA insertion in *AGL40* since one of the parent (*agl40-1*) had homozygous T-DNA insertion in *AGL40* and hence all F1 plants should have heterozygous T-DNA insertion in *AGL40*. F1 plants with heterozygous T-DNA insertion in both *AGL40* and *AGL62* were selfed and F2 seeds were collected. 100 F2 seeds were planted on soil and allowed to mature. PCR genotyping was carried out to screen the 68 surviving F2 plants to identify F2 plant with *AGL40* homozygous T-DNA insertion and *AGL62* T-DNA insertion. First the F2 plants were PCR genotyped using 001AGL40 and LBb1 to distinguish between WT plants and plants carrying T-DNA insertion in *AGL40* (Figure 4.12A) as before (see Figure 4.6 lanes 1 and 2 for example). Then the F2 plants with *AGL40* T-DNA insertion were genotyped using 001AGL40 and 002AGL40 to distinguish heterozygous and homozygous T-DNA insertion in *AGL40* (Figure 4.12B) as before (see Figure 4.6 lanes 3 and 4 for example). F2 plants that was shown to have homozygous T-DNA insertion in *AGL40* was finally genotyped for T-DNA insertion in *AGL62* using primers 002AGL62 and LBb1 (Figure 4.12C) as before (Figure 4.8).

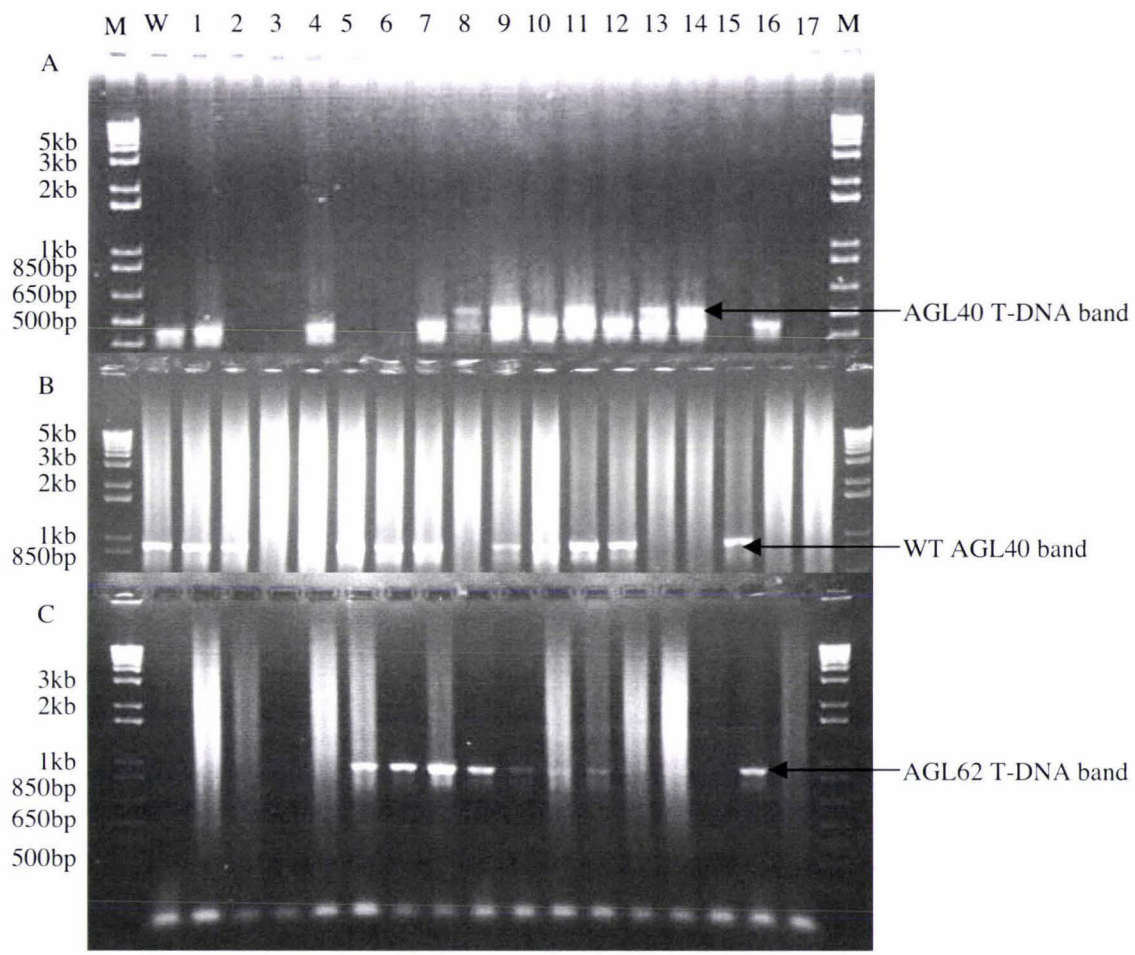


Figure 4.12. Representative gel of PCR genotyping of F2 plants. Top gel - F2 plants have been PCR genotyped for T-DNA insertion in *AGL40* using primers 001AGL40 and LBb1. The primer combination allows amplification of 550bp T-DNA fragment inserted within *AGL40* from the lines carrying T-DNA in *AGL40*. Middle gel - F2 plants have been PCR genotyped for homozygous T-DNA insertion in AGL40 using primers 001AGL40 and 002AGL40. If T-DNA insertion in AGL40 was homozygous, then the amplification of 1kb WT *AGL40* band should be disrupted. Bottom gel - F2 plants have been PCR genotyped for T-DNA insertion in *AGL62* using primers 002AGL62

PCR genotyping of F2 plants for homozygous T-DNA insertion in *AGL40* and T-DNA insertion in *AGL62* using the respective primers were able to recover 1 line that showed both homozygous T-DNA insertion in *AGL40* and T-DNA insertion in *AGL62* (Figure 4.12 lane 8). Phenotypic analyses were carried out with the line to observe the effect of

knocking out the function of 2 closely related members of type I MADS-box genes - *AGL40* and *AGL62* - on plant development. Since *AGL40* and *AGL62* share their expression domain in inflorescence and fruits, phenotypic analyses were carried out with inflorescence and fruit tissues, which included observation of inflorescence structure, fruit morphology, and seed numbers. However, these assays showed that phenotypes observed in the *agl40-1/agl40-1;agl62-1/+* mutant was not significantly different from that of single *agl62-1* mutants in terms of fruit size, seed number and seed sizes (data not shown).

4.8 Discussion

4.8.1 *AGL40* T-DNA insertion lines

To study functions AGL40 and AGL62 may have in plant development, loss-of-function mutants of AGL40 and AGL62 were necessary. T-DNA is a short piece of *Agrobacterium* origin DNA that is inserted in host plant genome and is used commonly in plant biology to knock out the function of target genes (Clough and Bent 1998). Arbidopsis plants are known to be amenable to *Agrobacterium* transformation making T-DNA insertion a powerful tool to study function of target genes in *Arabidopsis* (Clough and Bent 1998). Many T-DNA insertion lines are available to public and SALK institute (http://signal.salk.edu/index.html) T-DNA database lists publically available T-DNA insertion lines. In this study, SALK institute T-DNA insertion lines for both *AGL40* and *AGL62* were obtained from ABRC (www.biosci.ohio-state.edu/pcmb/Facilities/abrc/abrchome.htm) for molecular and phenotypic analyses. However, the problem of T-DNA insertion is that insertion is random. Furthermore, SALK institute T-DNA database (http://signal.salk.edu/index.html) indicated T-DNA insertion in *AGL62* to be in C-terminal domain of the *AGL62*, which are the least conserved domains in MADS-box gene family. Therefore to confirm loss-of-function mutation of AGL40 and AGL62, PCR genotyping were used to check insertion of T-DNA in *AGL40* and *AGL62*.

The genotyping was able to identify *AGL40* homozygous T-DNA insertion lines using gene specific primers 001AGL40 and 002AGL40 (Figure4.6) and subsequent expression analysis by semi-quantitative RT-PCR (Figure 4.7) identified loss-of-function mutant of AGL40 but the subsequent phenotypic analyses of tissues that express *AGL40* failed to

detect any phenotype indicating AGL40 may have redundant role in development of these genes with the other closely related members of type I MADS-box genes. It is also possible that AGL40 expression in planta is very specific or transient and function of AGL40 in plant development is subtle that our phenotypic analyses could not detect any phenotypes. In a future, production of AGL40 over-expression lines may be useful in understanding AGL40 function. Also construction of AGL40::GUS line may be necessary to show detailed *in planta AGL40* temporal and spatial expression pattern to characterize AGL40 function.

4.8.2 *AGL62* T-DNA insertion lines

On the other hand, after more than 100 PCR genotyping reactions, we were only able to recover heterozygous T-DNA insertion in *AGL62* (Table 4.2 and appendix2). Instead of the usual Mendelian genetics ratio of 1:2:1 (Homozygous:Heterozygous:WT), 2:1 ratio of Heterozygous T-DNA insertion:WT was recovered (Table 4.2). This gave rise to a possibility that AGL62 may function in regulation of gametophyte development or embryo development. If AGL62 had a role in either male or female gametophyte development then, disruption of AGL62 function results in aberrant fertilization, which lead to aborted embryo. If this was the case, then ratio of WT:*agl62-1/+* in segregating population of *agl62-1* is expected to be 1:1 as both *agl62-1/agl62-1* and one of *agl62-1/+* gives embryonic lethal phenotype. However, the ratio of WT:*agl62-1/+* in a segregating population of *agl62-1/+* was roughly 1:2 (Table 4.2), which indicated AGL62 might have a role in embryo development. If AGL62 had a function in embryo development, then agl62-1/agl62-1 results in embryonic lethality but agl62-1/+ can still give rise to WT embryo resulting in WT:agl62-1/+ ratio of 1:2. Since the completion of *Arabidopsis* genome sequencing (Arabidopsis Genome 2000) many embryonic lethal mutants or gametophyte lethal mutants have been identified in *Arabidopsis* (Drews and Yadegari 2002; Faure, Rotman et al. 2002; McCormick 2004). Many of such mutants were identified in a mutagenesis assay screening for a defective seed phenotype (Mayer and Jurgens 1998; Ostergaard and Yanofsky 2004). Therefore phenotypic analyses needs to be carried out in next chapter to investigate whether AGL62 plays a role in gametophyte or embryo development. Such analyses include setting up cross between *agl62-1/+* and WT and screening the progeny for ratio of *agl62-1/+* to WT to determine if there will be

any deviation from the expected ratio. Also to be able to state that embryonic lethal phenotype observed in this study was due to loss of function of AGL62, we must in a future investigate the possibility of other T-DNAs being inserted elsewhere in genome of *agl62-1* plants that could affect embryo viability in agl62-1/+ as T-DNA insertion is a random event. To this end, back-crossing the *agl62-1/+* to WT several times and selecting for T-DNA insertion in *AGL62* in segregating population each time should be sufficient to remove any unwanted T-DNA insertions elsewhere in genome.

4.8.3 Double *agl40;agl62* knock out mutant

Double *agl40;agl62* knock out mutant was also produced in our study to further investigate the roles AGL40 and AGL62 may play on plant development and whether they act redundantly to regulate plant development especially since yeast-2-hybrid assay showed physical interaction capability between AGL40 and AGL62 (de Folter, Immink et al. 2005). Crosses were made between *agl62-1/+* and *agl40-1/agl40-1* and F1 plants with agl40-1/+;agl62-1/+ was selfed. F2 seeds were planted to screen for *agl40-1/agl40-1;agl62-1/+* by PCR genotyping. F2 plants were PCR genotyped for homozygous T-DNA insertion in *AGL40* and *AGL62* as well because mutation in *AGL40* may rescue *AGL62* expression and thereby rescue embryo lethal phenotype of *agl62-1* mutant. Out of 68 surviving F2 plants that were PCR genotyped, only 1 line showed *agl40-1/agl40-1;agl62-1/+* genotype (Figure4.12) and no *agl40-1/agl40-1;agl62-1/agl62-1* genotype was detected. We screened for homozygous T-DNA insertion in *AGL40* (one quarter should have *agl40-1/agl40-1*) and heterozygous T-DNA insertion in *AGL62* (one half should be *agl62-1/+*), therefore one eights should have *agl40-1/agl40-1;agl62-1/+* genotype. It was unexpected to only get 1 plant out of 68 plants that showed *agl40-1/agl40-1;agl62-1/+* genotype. However, the phenotypes observed in the double *agl40-1/agl40-1;agl62-1/+* mutant was comparable to that of single *agl62-1/+* mutant. This suggested that AGL62 has its own role in regulation of embryo development that it does not share with AGL40 and phenotypic analyses will be carried on developing embryo in *agl62-1* plants to study the role AGL62 may play in regulation of embryo development.

Chapter 5. Genetic & Phenotypic analyses on segregating population of *agl62-1* plant

5.1 Introduction

For over one hundred plants from segregating population of *agl62-1*/+ plant, PCR based genotyping using gene specific primers 001AGL62 and 002AGL62 (table 2.1) were carried out (Section 2.6.2) but homozygous T-DNA insertion in *AGL62* (*agl62-1/agl62-1*) was never recovered (Appendix 2). The PCR genotyping were repeated using newly designed gene specific primers 005AGL62 and 006AGL62 to confirm that the first result was not due to the contamination of DNA extracts from agl62/+ segregating population by the WT DNA extracts since DNA was extracted from WT plants along with agl62-1/+ segregating population to be used as a control. The second PCR genotyping still did not recover any homozygous T-DNA insertion in AGL62 (data not shown) indicating knocking out the function of AGL62 may result in lethal phenotype. To functionally characterize AGL62, it is necessary to understand at which stage the *AGL62* homozygous T-DNA insertion lines are aborted.

In their life cycle, plants alternate between haploid (n) phase and diploid (2n) phase. In *Arabidopsis*, diploid (2n) or sporophytic phase consists of most of its life cycle and haploid (n) or gametophytic phase is reduced to limited period during which male and female reproductive organs develops. Gametophytic phase initiates with development of microspore (male reproductive organ) and megaspore (female reproductive organ) and ends at the double fertilization where pollen developed from microspore fertilizes egg cell and central cell developed from megaspore to initiate seed development (Figure 5.1). Based on the AGL62::GUS expression pattern, there are 3 possible stages where the agl62-1/agl62-1 could abort during development (Figure 5.1). These stages include pollen development (Figure 5.1 arrow 1), seed development (Figure 5.1 arrow 2) and seed germination (Figure5.1 arrow 3).

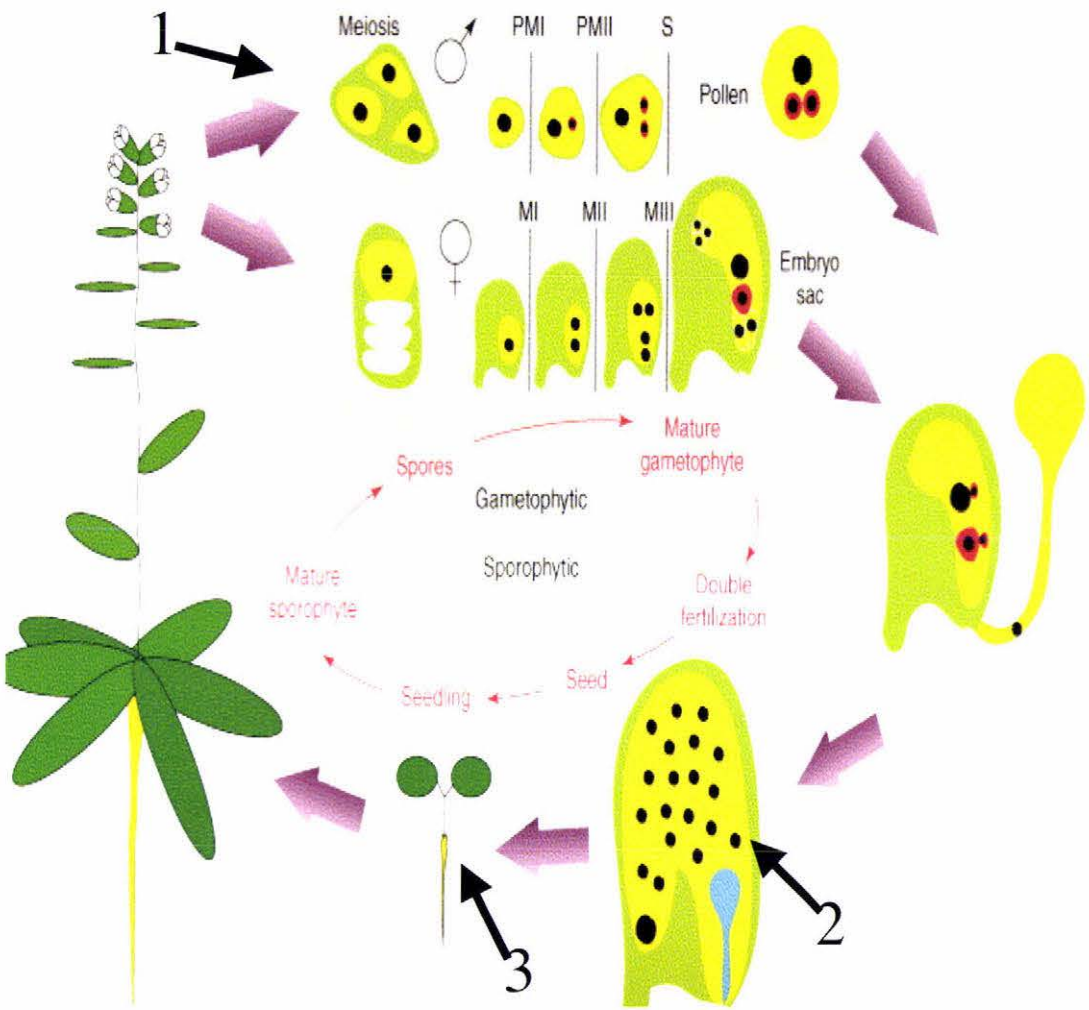


Figure 5.1 Life cycle of *Arabidopsis thaliana*. Figure modified from (Berger, 2003). *Arabidopsis* plants alternate between diploid (2n) sporophytic phase and haploid (n) gametophytic phase. In *Arabidopsis*, the haploid gametophytic stage is reduced to the time of reproductive organ development which include microspore (male) and megaspore (female) development to the double fertilization which represents the onset of seed development while diploid sporophytic phase starts at double fertilization and consists of the rest of its life including seed stage, seedling stage and mature vegetative state. Black arrows with numbers 1 to 3 points to the stage where *AGL62* homozygous T-DNA insertion lines are possibly aborting. Arrow number 1 points to gametophytic phase, arrow number 2 points embryo developing stage and arrow number 3 points at the germination stage.

First, the microarray expression data for *AGL62* (Table 3.1) indicated that AGL62 might have a role in development of male (pollen) or female (ovule) reproductive organ (Figure

5.1 arrow 1). GUS staining assay on AGL62::GUS plants has also shown expression of *AGL62* in pollen further supporting this hypothesis (Figure 3.6). If this is true, agl62-1/agl62-1 mutants will lack functional gametophytes and hence homozygous T-DNA insertion in *AGL62* will not be recovered in the F1 population from self of agl62-1/+ instead, only WT and agl62-1/+ with a 1:1 ratio should be recovered (Figure 5.2). If AGL62 had no role in gametophyte development then both male and female reproductive organs will be formed then according to the Mendelian genetics, ratios between WT:agl62-1/+:agl62-1/agl62-1 should be1:2:1 (Figure 5.2 panel A). On the other hand, if AGL62 had a role in male or female reproductive organ, then since AGL62 function is knocked out, agl62-1/agl62-1 mutant will not be able to form proper reproductive organ. Furthermore, if AGL62 had a function in male reproductive organ development, then agl62-1/+ in which T-DNA is derived from male copy will fail to form proper pollen in which case the ratio between WT:agl62-1/+ will be 1:1 (Figure 5.2 panel B). Similarly if AGL62 had a role in female reproductive organ, then agl62-1/agl62-1 and agl62-1/+ in which T-DNA is inserted in female *AGL62* copy will fail to form proper female reproductive organ and the ratio between WT:agl62-1/+ will be 1:1 (Figure 5.2 panel C).

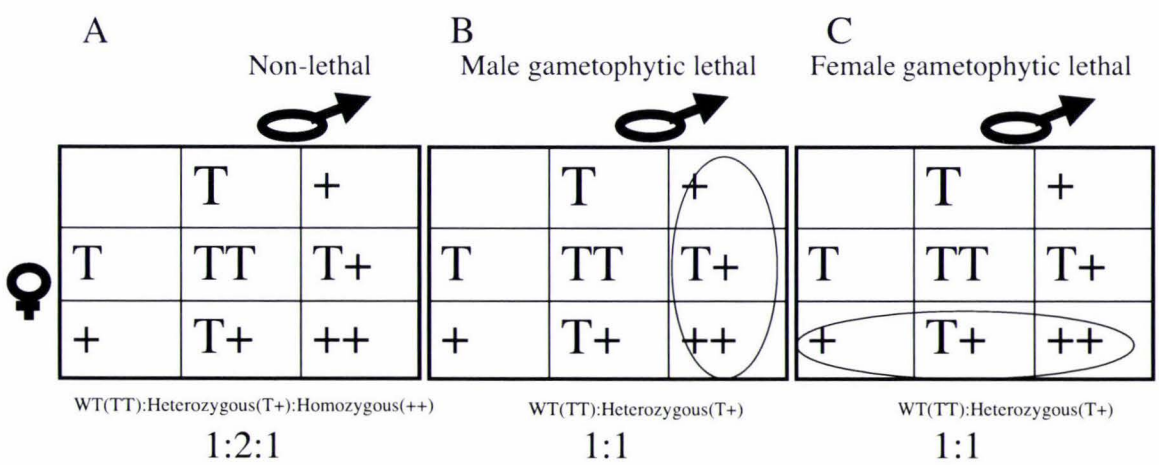


Figure 5.2 Punnet squares showing ratios of genotypes in progenies from Heterozygous *AGL62* T-DNA insertion lines selfed. A - If *AGL62* had no function in either male or female gametophyte development, then the self cross of heterozygous *AGL62* T-DNA insertion lines (T+) should produce WT (TT):Heterozygous T-DNA insertion (T+):Homozygous T-DNA insertion (++) in a 1:2:1 ratio. B - If *AGL62* had role in male gametophyte development, Homozygous T-DNA insertion and Heterozygous T-DNA insertion in which T-DNA is inserted in male copy should be embryonic lethal therefore, WT:Heterozygous *AGL62* T-DNA insertion is 1:1. Circles surrounding genotypes indicate the genotypes that result in lethal phenotype. C - Similarly, if *AGL62* had a role in female gametophyte development, Homozygous T-DNA insertion and Heterozygous T-DNA insertion in which T-DNA is inserted in female copy should be embryonic lethal therefore, WT:Heterozygous *AGL62* T-DNA insertion is 1:1. Circles surrounding genotypes indicate the genotypes that result in lethal phenotype.

The second possibility is that AGL62 has a role in seed development and knocking out AGL62 function results in embryonic lethality phenotype (Figure 5.1 arrow 2). If this was the case, then agl62-1/+ can still give rise to seed but the agl62-1/agl62-1 seed will have non-functional *AGL62* copies from both parents and results in lethal embryo and will not be recovered. Then instead of the normal Mendelian ratio of WT:agl62-1/+:agl62-1/agl62-1 of 1:2:1 if AGL62 had no function in seed development (Figure 5.3 panel A), the ratio between WT:agl62-1/+ will be 1:2 if AGL62 had a role in seed development because only the agl62-1/agl62-1 fails to develop the mature seed while agl62-1/+ can still produce mature seeds (Figure 5.3 panel B).

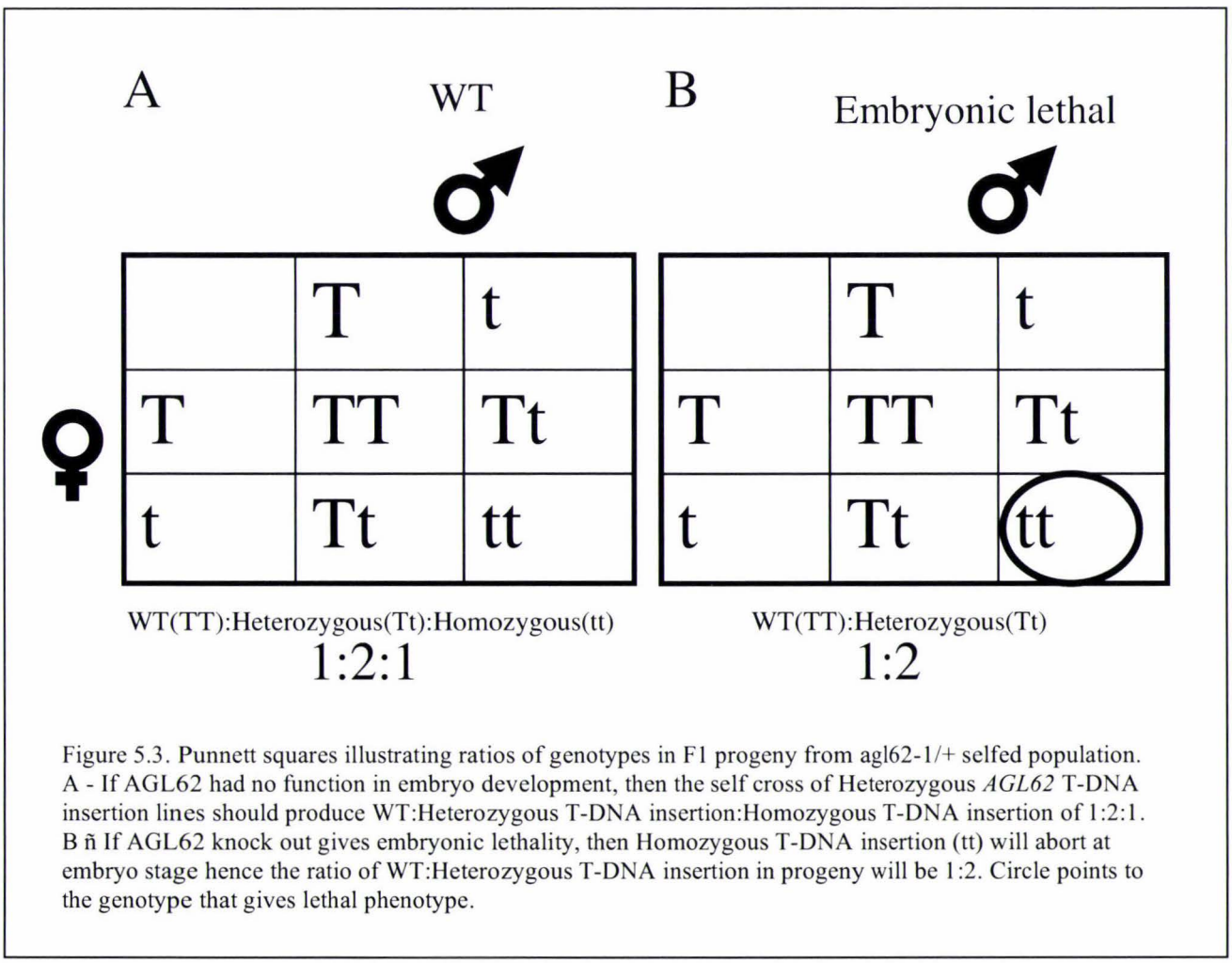


Figure 5.3. Punnett squares illustrating ratios of genotypes in F1 progeny from agl62-1/+ selfed population. A - If AGL62 had no function in embryo development, then the self cross of Heterozygous *AGL62* T-DNA insertion lines should produce WT:Heterozygous T-DNA insertion:Homozygous T-DNA insertion of 1:2:1. B ñ If AGL62 knock out gives embryonic lethality, then Homozygous T-DNA insertion (tt) will abort at embryo stage hence the ratio of WT:Heterozygous T-DNA insertion in progeny will be 1:2. Circle points to the genotype that gives lethal phenotype.

Result from the large-scale PCR genotyping (Table 4.2) gave roughly 1:2 ratio of WT:*agl62-1*/+ that supports this hypothesis.

The third possibility is that AGL62 functions at germination stage when seedlings develop from seed and in *agl62-1/agl62-1* mutants, seed does not germinate and hence mature Homozygous *AGL62* T-DNA insertion plant would never be recovered.

5.2 Gametophytic lethality assay on agl62-1/+ segregating population

5.2.1 Alexander staining

Microarray data (table 3.1) indicated expression of *AGL62* in pollen suggesting that there is a possibility that AGL62 may indeed play a role in male gametophyte (pollen) development. GUS staining assay on AGL62::GUS plants has also shown expression of AGL62 in developing pollen supporting this hypothesis (Figure 3.6). To further investigate this possibility, developmental abortion rate of pollen from WT and *agl62-1*/+ were compared using a simple Alexander staining technique to see if pollen from agl62-1/+ is aborting prematurely compared to the WT.

Alexander staining is capable of differentially staining aborted and non-aborted pollens. Alexander dye is a mixture of stains composed of malachite green, acid fuchsin and orange G. Malachite green is a weakly basic dye which stains the pollen wall in green while acid fuchsin stains protoplasm red to deep red. Orange G will improve the differentiation between the pollen wall and protoplasm. Since the aborted pollens only consist of a pollen wall, malachite green stains them in green uniformly whereas the viable pollens are stained red by acid fuchsin on inside and green on the outside wall by malachite green and thereby differentiating the viable from non-viable pollen. Four *agl62-1*/+ plants and 1 WT control plant were used to compare the abortion rates of pollen in each genotype. Pollen was collected randomly from various developmental stages of both WT and *agl62-1*/+ flowers ranging from just before opening (Stage10) to anthesis stage when pollination occurs (stage13) (Smyth et al., 1990) to observe at which developmental stage the pollen from agl62-1/+ may be aborting. The collected pollen were mounted on microscope slide, stained by Alexander dye and observed under light microscope. Throughout all stages, number of both green aborted pollen and deep red viable pollen were identified in pollen from segregating population of agl62-1/+ plants (Figure 5.4C). Compared to the viable pollen, which is shaped spherical (Figure 5.4B),

green aborted pollen showed oval shape due to the absence of protoplasm (Figure 5.4D). However, both aborted and non-aborted pollen were also identified in WT flowers as well (Figure 5.4A) indicating that abortion of pollen observed in agl62-1/+ segregating population may be a natural phenomenon rather than a result of malfunction in AGL62.

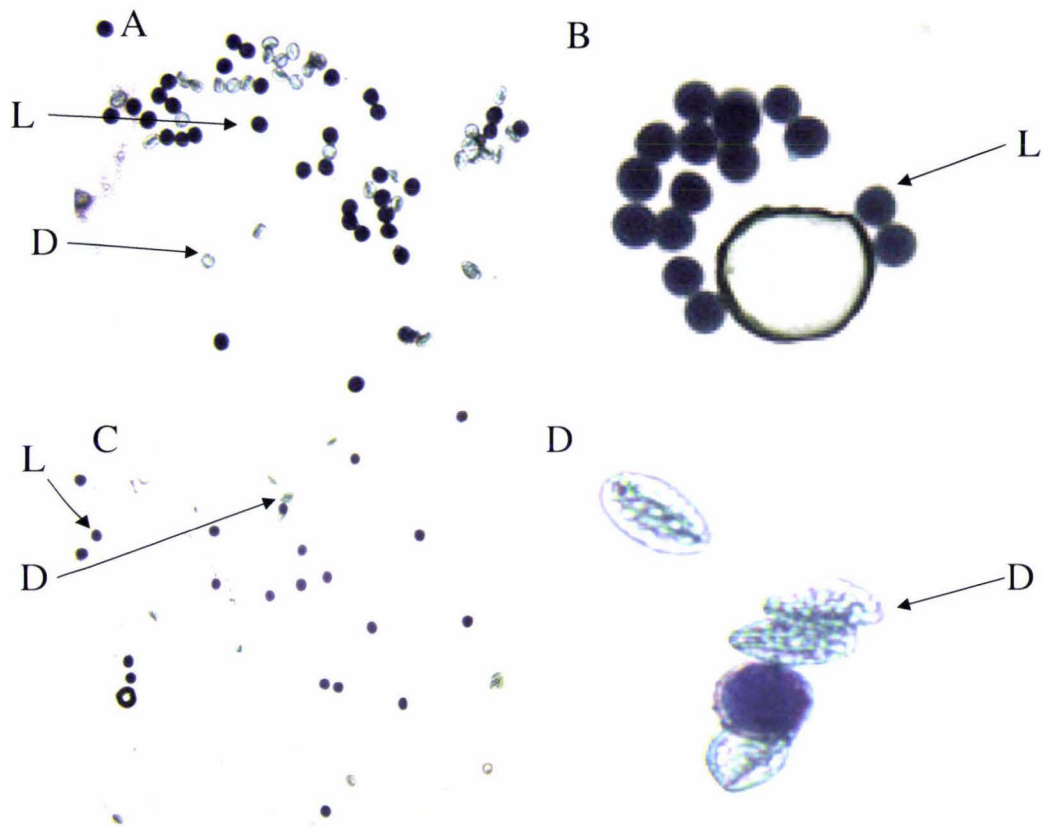


Figure 5.4. Alexander staining of pollen from WT and Heterozygous *AGL62* T-DNA insertion plants. Alexander staining can differentiate between live pollen and aborted pollen. Live pollen (L) has sphere shape with green membrane and purple inside whereas dead pollen (D) is oval and stained entirely in green. A - Pollen from WT stained with Alexander dye. B - Higher resolution picture of Alexander stained WT pollen. C - Pollens from heterozygous *AGL62* T-DNA insertion line stained with Alexander dye. D - Higher resolution picture of Alexander stained heterozygous *AGL62* T-DNA insertion line pollen.

More pollen were collected from both WT and agl62-1/+ plants and stained with Alexander dye to see whether the aborting pollen observed in agl62-1/+ plants are due to T-DNA insertion in *AGL62*. The numbers of aborted and non-aborted pollen from WT and agl62-1/+ plants were counted and tabulated (table 5.1). Out of approximately 800 WT pollen observed, roughly 8% of them were aborted which was a much higher abortion ratio than 2 of the agl62-1/+ plants observed. These 2 agl62-1/+ plants Ab435-1

and Ab435-3 respectively showed 1.47% and 1.09% pollen abortion rate respectively. However, the other 2 agl62-1/+ plants Ab435-4 and Ab435-5 showed pollen abortion rates of 9.34% and 8.22% respectively, which were similar to that of WT.

		No. Total pollen	No. Viable pollen	No. Dead pollen	% Aborted pollen
WT	Ab434	603	560	43	7.68
AGL62 hetero	Ab435-1	812	800	12	1.47
	Ab435-3	275	272	3	1.09
	Ab435-4	107	97	10	9.34
	Ab435-5	888	815	73	8.22

Table5.1 Number of aborted and non-aborted pollen in WT and Heterozygous *AGL62* T-DNA insertion plants. Pollen from 4 AGL62 Heterozygous T-DNA insertion plants (Ab435-1 - Ab435-5) and 1 WT (Ab434) were chosen at stages between 10 and 13 of flower development and stained with Alexander dye. Numbers of viable and dead pollens were counted and percentages of aborted pollen were calculated.

Microscopic observation of stained pollens showed that abortion rate of pollen in *agl62-1/+* was variable and not always significantly different from that of WT. To conclude, Alexander staining did not detect any phenotypic defects in developing pollen from *agl62-1/+* indicating putative loss-of-function of AGL62 alone does not affect the development of pollen.

5.2.2 Transmission assay on *agl62-1/+* segregating plants

Alexander staining of pollen from agl62-1/+ segregating population indicated that the putative loss-of-function of AGL62 does not affect pollen viability. To determine if there are any developmental defects in the developing male or female gametophyte in agl62-1/+ background, transmission assays were conducted. The transmission assays involved setting up a controlled cross between WT and agl62-1/+ mutants (Section 2.8.1) and screening segregating progenies for ratios of WT to agl62-1/+. If AGL62 does not affect the transmission of male or female gametophyte during development, then ratios of WT to agl62-1/+ will be 1:1 in the F1 progeny according to the Mendelian genetics (Figure 5.5 panels A and B). However, if AGL62 did affect the transmission of a gamete during development, then the gametes with the T-DNA in AGL62 (+) would not be transmitted

and the agl62-1/+ (T+) background would never be recovered in F1 offspring so F1 offspring would all be WT (TT). To determine if the T-DNA insertion in *AGL62* affects transmission of the male gamete (Figure 5.5 panel A) or the female gamete (Figure 5.5 panel B), crosses were set up to WT in both direction so that agl62-1/+ allele was either transmitted through the male or female gametophyte and F1 progeny were scored for WT:agl62-1/+ ratio.

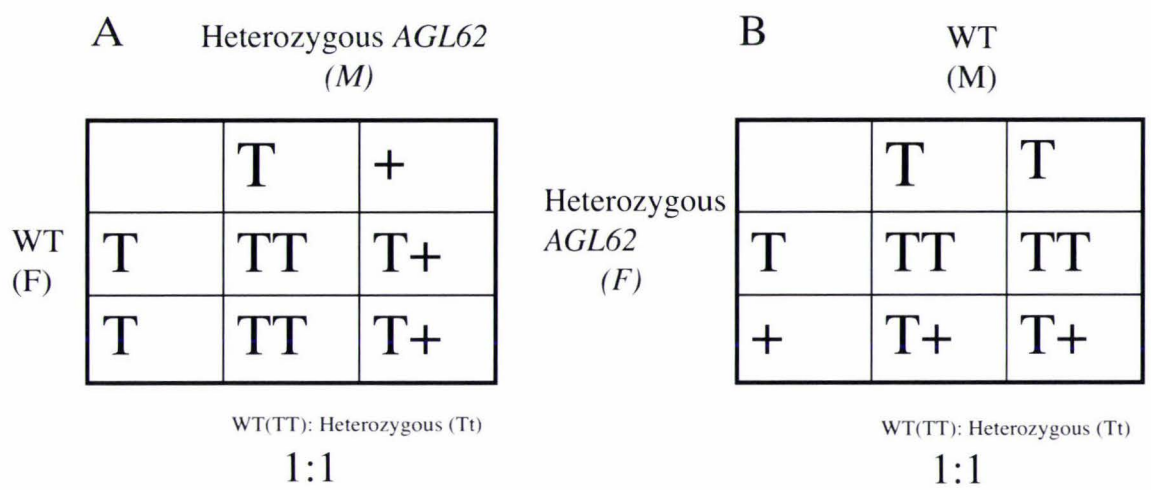


Figure 5.5. Punnet squares showing ratios of genotype in the F1 progeny from a cross between agl62-1/+ and WT. According to Mendelian genetics, if AGL62 does not affect transmission of either male or female gamete, ratio of genotype in a F1 progeny of a cross between WT and agl62-1/+ would be WT (TT):agl62-1/+ (T+) = 1:1. A - If AGL62 did have a role in transmission of male gamete, then F1 offspring with agl62-1/+ background from WT (female (F)) X agl62-1/+ (male (M)) cross should abort due to non functional male gametophyte development hence result in recovering all WT in the progeny. B - Similarly if AGL62 did have a role in transmission through female gamete, then Heterozygous F1 progeny from WT (male) X agl62-1/+ (female) cross should abort due to non-functional female gametophyte development hence result in recovering all WT in the progeny.

Ten crosses between WT and agl62-1/+ were made as described in method (Chapter 2.8.1). F1 seeds were plated on MS medium supplemented with kanamycin (50ug/ml) (Section 2.1.1 and 2.1.3) to select for kanamycin resistance in the progeny as a method to differentiate WT and agl62-1/+ plants. Since the T-DNA insertion confers plants with kanamycin resistance, kanamycin resistance was used to distinguish between WT and *AGL62* heterozygous T-DNA insertion plants. However, all the seeds that were plated on kanamycin (50ug/ml) supplemented medium died after germination indicating that they had no kanamycin resistance. This could mean that no T-DNAs were transferred to the

segregating population but it is also possible that kanamycin resistance was lost in this particular line. Fifty agl62-1/+ seeds were plated on MS kanamycin (50ug/ml) plate to assess whether the kanamycin gene had been silenced in agl62-1/+ plants. All seeds plated died after germination (Figure 5.6) indicating kanamycin resistance gene was silenced in agl62-1/+ plants.

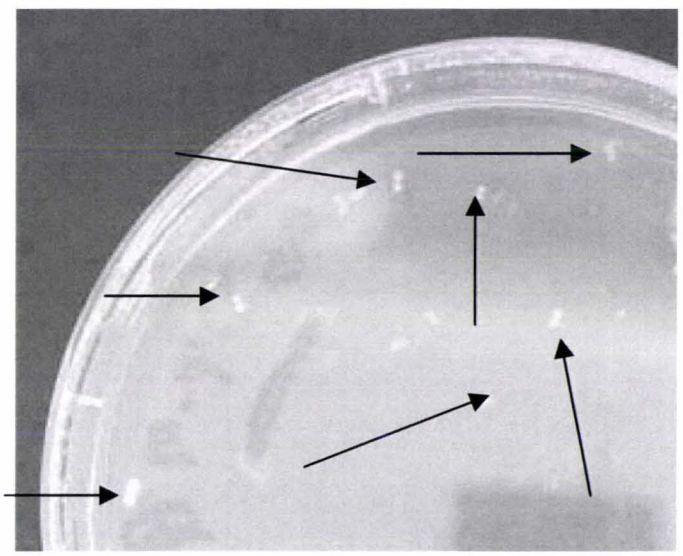


Figure 5.6. *agl62-1*/+ self-*progenies* plated on MS kanamycin (50ug/ml) plate. Fifty *agl62-1*/+ seeds have been plated on MS plate supplemented with kanamycin (50ug/ml). All 50 seeds plated have died indicating Npt II gene was silenced in AGL62 heterozygous T-DNA insertion line and kanamycin resistance was lost from the line. Arrows are showing positions of dead seedlings

Therefore, PCR genotyping was used as an alternative approach to assess the transmission of the agl62-1 allele (Section 2.6.2). Using a gene specific primer 002AGL62 and T-DNA border primer LBb1 (Table 2.1), PCR was able to amplify the AGL62 T-DNA band from heterozygous *AGL62* T-DNA insertion plants and therefore distinguished between WT and heterozygous *AGL62* T-DNA insertion plants in the F1progeny as tabulated (table 5.2).

Cross	No. WT	No. Heterozygous *AGL62* T-DNA insertion
Heterozygous *AGL62* T-DNA insertion line (F) × *WT* (M)	12	26
WT (F) × Heterozygous *AGL62* T-DNA insertion line (M)	24	20

Table 5.2. PCR genotyping identification of WT:Heterozygous *AGL62* T-DNA insertion ratios in progeny of the WT X *agl62-1* crosses. In WT (F) X Heterozygous *AGL62* T-DNA insertion line (M) cross, ratios of WT:Heterozygous *AGL62* T-DNA insertion in progeny were nearly 1:1. In Heterozygous *AGL62* T-DNA insertion line (F) × *WT* (M) cross, ratios of WT:Heterozygous *AGL62* T-DNA insertion in progeny was 1:2. F – female, M – male.

Fifty seeds from segregating population of WT female (F) X agl62-1/+ male (M) crosses and 50 seeds from segregating population of agl62-1/+ (F) X WT (M) crosses were planted on soil and PCR genotyping was carried out with surviving plants to screen for ratios of WT:heterozygous T-DNA insertion in *AGL62* in the segregating F1 progenies. In WT (F) X agl62-1/+ (M) cross, number of WT:heterozygous *AGL62* T-DNA insertion in progeny was 24:20, which roughly corresponds to 1:1 ratio. This indicated that AGL62 does not affect the transmission of male gametophyte. In agl62-1/+ (F) (WT (M) cross, number of WT:heterozygous *AGL62* T-DNA insertion in progeny was 12:26 which roughly corresponded to 1:2 ratio. Although the number of heterozygous *AGL62* T-DNA insertion in progeny was twice as high as WT in this cross, *agl62-1* transmission through female gametophyte seemed to be unaffected as we were able to recover heterozygous *AGL62* T-DNA insertion lines. This transmission assay indicated that the agl62-1 allele could be transmitted through the male and female gametophytes.

5.3 Whole mount assays on *agl62-1/+ segregating population*

5.3.1 AGL62 and germination

Since pollen from *agl62-1/+* plants did not show any significant phenotypic defect, it is possible that *agl62-1/agl62-1* mutants are aborting at either seed development or seed germination stage (Figure 5.1 arrows 2 and 3). To investigate if AGL62 plays any role in

the regulation of the seed germination, germination assays were carried out. Roughly 300 seeds were randomly selected from WT and *agl62-1*/+ F1 segregating population and plated on MS medium. At 3, 5, 7 and 12 days after plating WT and *agl62-1*/+ seeds on MS media, the development of both WT and *agl62-1*/+ plants were observed (Table 5.3).

	Germination after 3days	Seedling viability after 5days	Seedling viability after 7days	Seedling viability after 12days
WT	187/315 (59.4%)	183/315 (58.1%)	204/312 (65.4%)	101/131 (77.1%)
Agl62-1/+	196/348 (56.3%)	175/336 (52.1%)	218/334 (65.3%)	98/131 (74.8%)

Table 5.3. Germination ratios and seedling viability ratios of Heterozygous *AGL62* T-DNA insertion lines compared to that of WT. Seeds of WT and Heterozygous *AGL62* T-DNA insertion line segregating populations were plated on MS media and at 3, 5, 7, and 12 days after plating, the germination ratio and seedling viability ratio of WT and *agl62-1*/+ segregating population were compared. In the above table, number of germinated seeds or surviving seedlings were counted out of the total number of seeds plated and numbers in parentheses represent the percentages of seed germination or seedling viability. Both germination and seedling viability rate of Heterozygous *AGL62* T-DNA insertion segregating population were comparable to that of WT.

There was no significant difference between germination rates between WT and *agl62-1*/+ F1 segregating population after 3days of plating the seeds on MS media with 59.4% (WT) and 59.3% (*agl62-1*/+) germination rates respectively (Table 5.3) indicating that AGL62 alone does not function to regulate seed germination. Viability of the germinated WT and *agl62-1*/+ F1 segregating seedlings was observed in relation to the expression of AGL62::GUS in seedling leaves and roots (Figures 3.4 and 3.5). If AGL62 played a role in the regulation of seedling growth then since roughly one-quarter of the seeds plated on MS media should be *agl62-1*/*agl62-1*, F1 segregating population of *agl62-1*/+ should show 25% more seedling abortion rate than WT. However, not much difference in seedling viability was observed between WT and *agl62-1*/+ segregating population after 5, 7, and 12 days of plating the seeds on MS media (Table 5.3) implying at least AGL62 alone does not regulate germination of the seeds or the viability of the developing seedling.

5.3.2 *AGL62* and embryonic lethality

Since abortion of *agl62-1/agl62-1* gametophyte nor germinating seeds were not detected, the last possibility as to why no homozygous *AGL62* T-DNA insertion line can be recovered is that the loss of function of AGL62 results in embryo lethality (Figure 5.1 arrow 2). To study embryo and seed development of WT and *agl62-1/+* segregating population, Hoyer's clearing was used in order to visualize all layers of the seed and embryo and endosperm development (Section 2.8.2.4).

As a first step in investigating whether AGL62 plays a role in seed development, fruits 3 to 7 days after anthesis were collected and Hoyer's cleared (Figure 5.7). Fruit walls are sporophytic tissues but fruit walls enclose developing next generation seeds. In WT, fruits are produced as a result of self-pollination and all developing seeds within fruit are WT whereas in *agl62-1/+* plants, fruits bear F1 seeds that should segregate for WT, *agl62-1/+* and *agl62-1/agl62-1*. Therefore, if *agl62-1/agl62-1* background gives embryo lethality, number of seeds in *agl62-1/+* fruits should be 25% less than WT. This assay also gives rough estimation of at what stage the *agl62-1/agl62-1* embryo is aborting. Until 4days after anthesis, no difference in seed number was detected between fruits from WT and *agl62-1/+* (Figure 5.7A&B). However, after 5days of anthesis, reduction in number of seeds was observed in *agl62-1/+* fruits compared to the WT (Figure 5.7C) and many gaps in between seeds in fruits started to become evident 6 and 7days after anthesis (Figure 5.7D&E).

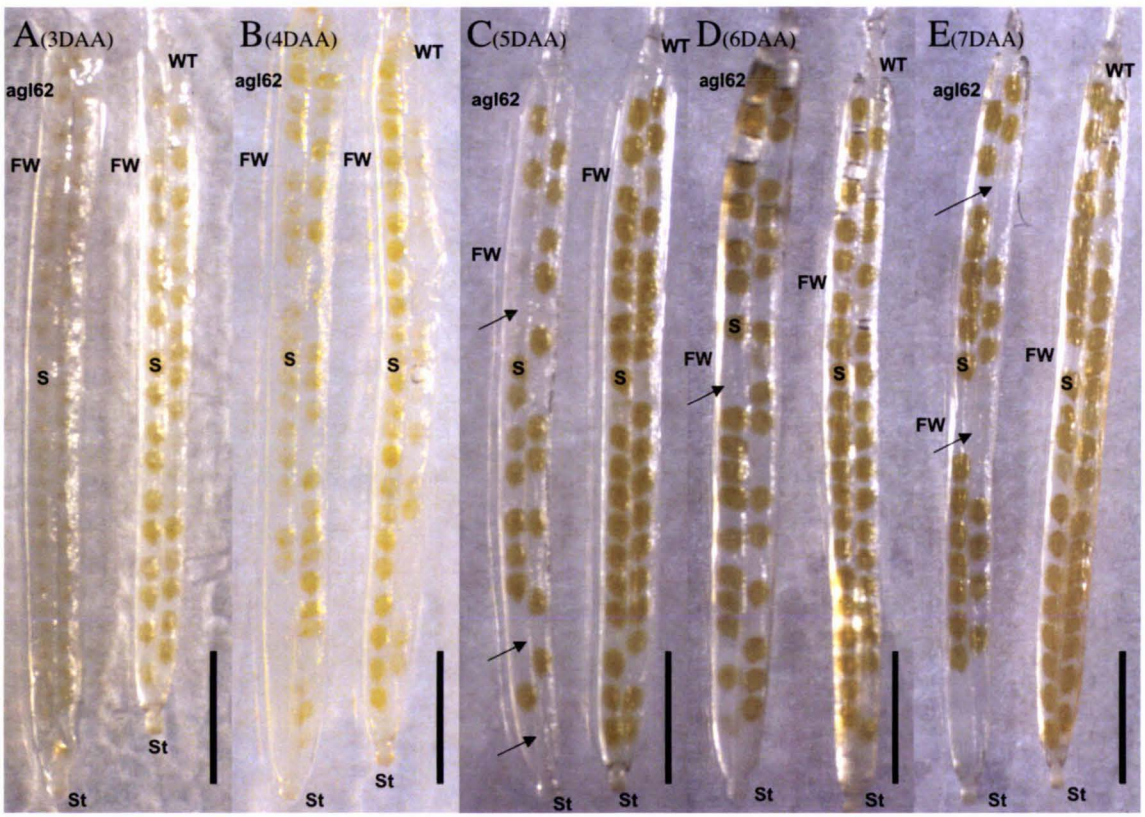


Figure 5.7. Comparison of seed numbers in Hoyer's cleared fruits from WT (WT) and Heterozygous *AGL62* T-DNA insertion mutants (agl62). In all pictures, WT is on the right hand side and mutant lines are on the left. Seeds in Heterozygous *AGL62* T-DNA insertion line fruits are segregating for WT, agl62/+ and agl62/agl62. Fruits were collected A) 3days after anthesis (DAA), B) 4DAA, C) 5DAA, D) 6DAF, and E) 7DAA. S - Seed, FW - Fruit Wall, St - Style, Arrows indicate aborted seed space. Scale bars represent 2mm.

Number of seeds in fruits from 3 to 7 days after anthesis was compared between WT and *agl62-1*/+ plants (Table 5.4) to further investigate whether the difference in seed numbers observed between WT and *agl62-1*/+ fruits were due to loss-of-function of AGL62. In 3 days and 4 days old fruits, little difference in seed number was observed between WT and *agl62-1*/+ plants (Figure 5.7 A and B and table 5.4). At 3days after anthesis, 37seeds on average was observed from WT fruits whereas the average number was 39seeds from the *agl62-1*/+ fruits (Table 5.4). Similarly at 4days after anthesis, average number of seeds in WT was 39 while the number was 36 in *agl62-1*/+ (Table 5.4). However, significant reduction in number of seeds in *agl62-1*/+ fruits were observed after 5 days of anthesis

(Figure 5.7 C, D, and E and Table 5.4) which correlated with the expression of *AGL62* in seeds 3 days after flowering. 5days after anthesis, WT fruit had average of 40seeds but *agl62-1*/+ fruits only had an average of 31seeds and by the 7days after anthesis, average number of seeds in fruits was 42 (WT) against 30 (*agl62-1*/+) indicating *agl62-1*/*agl62-1* embryos are aborting between 3days and 5days after anthesis.

Fruit age	7DAF	6DAF	5DAF	4DAF	3DAF
WT	42.25 ±2.66[a] (338)	41.50 ± 2.88 (249)	39.75 ± 2.12 (318)	39.88 ± 1.89 (319)	37.75 ± 2.60 (302)
Heterozygous AGL62	30.13 ±12.09 (241) *	34.50 ± 6.09 (207)	31.88 ± 6.73 (255)	36.25 ± 9.29 (290)	39.00 ± 8.59 (312)

Table 5.4. Number of seeds in Hoyer's cleared fruits of WT and Heterozygous *AGL62* mutant
[a] Mean ± SD. DAF – Days after flowering. Number in bracket represents the number of seeds observed. Number of fruits observed was as follows. 7DAF, 8fruits. 6DAF, 6fruits. 5DAF, 8fruits. 4DAF, 8fruits. 3DAF, 7fruits.

The significant reduction in number of seeds in *agl62-1*/+ fruits further defined a possible role of AGL62 play in seed development. Seeds were collected from WT and *agl62*-1/+ fruits 1 to 7 days after flowering and cleared with Hoyer's medium to see if there was any phenotypic defect in seeds (Figure 5.8, and 5.9).

In WT, embryo develops from 1-celled zygote stage at 0 days after pollination (DAP) to walking stick cotyledon stage at 7 DAP (Smyth et al., 1990). After walking stick stage, cotyledon enlarges and embryos reach maturity (Smyth et al., 1990). Seeds then desiccate (Mayer et al., 1993). Between 0 and 7 DAP, many changes occur in seeds to accommodate rapid embryo development (Smyth et al., 1990). 1DAP, zygote divides asymmetrically to give rise to 2-celled embryo (Figure 5.8A). 2DAP, embryo reaches 4-celled stage (Figure 5.8B) and embryo can reach globular stage that consists of up to 100 cells at the end of 3DAP (Figure 5.8C). Four days after pollination, apical-basal polarity is determined and cotyledon elongate to give heart stage embryo (Figure 5.8D). Cotyledons continue to elongate and gives torpedo stage embryo at 5DAP (Figure 5.8E). Embryo continues to grow 6DAP (Figure 5.8F) until the embryo reaches walking stick stage 7DAP.

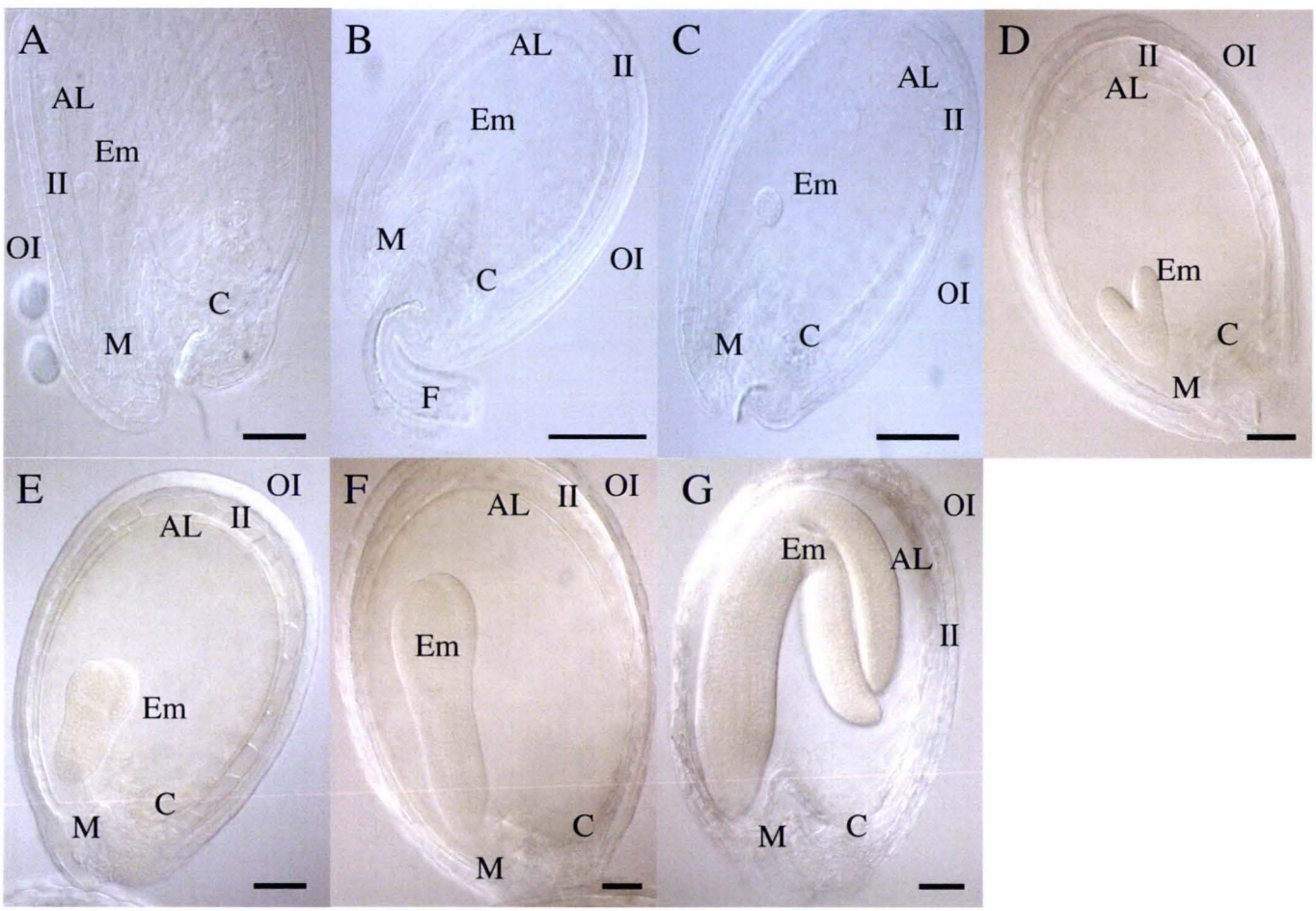


Figure 5.8. WT embryogenesis developmental stages. Fruits were dissected 1 to 7days after pollination and seeds were collected and cleared with Hoyer's medium to observe embryos. A) 1day after pollination (DAP), embryo is 2celled. B) 2DAP, embryo is 4 celled. C) 3DAP, embryo reaches globular stage. D) 4DAP,embryo is at heart stage. E) 5DAP,torpedo stage embryo. F) 6DAP, embryo is at transition between torpedo and walking stick stage. G) 7DAP, embryo is at walking stick stage. Bar represents 50um. Em - Embryo, Ch - Chalazal endosperm, Mi - Micropyle, II - Inner integument, OI - Outer integument, AL - Aleurone layer, F - Funicules

In comparison to WT embryos, some embryos from F1 segregating population of *agl62-1*/+ showed unusual traits (Figure 5.9). Until 2DAP, no difference was observed between WT embryos and embryos from F1 segregating population of *agl62-1*/+ (data not

shown). However 3DAP, some embryos from F1 segregating population of *agl62-1*/+ showed enlarged globular stage embryo that had rough appearance compared to WT's smooth and spherical globular embryo (Figure 5.9 B&C). Also the seeds with the enlarged embryos were smaller than the seeds, which had normal shaped embryos (Figure 5.9A). Four days after pollination, instead of cotyledon elongating to give heart shaped embryo, the globular embryo kept on enlarging (Figure 5.9 D&E). Five days after pollination, the enlarged embryo still remained as globular stage and did not proceed to the heart stage embryo (Figure 5.9F). After 6days of pollination, seeds with these enlarged embryos aborted (data not shown).

Identification of the unusually enlarged globular stage embryo between 3 to 5 days after pollination is in line with the spatial expression pattern of *AGL62* in seed (Figure 3.7) and the reduction of seed numbers in *agl62-1*/+ fruits after 4 days of flowering (Table 5.3). This identification further supports the hypothesis that AGL62 plays a role in seed development.

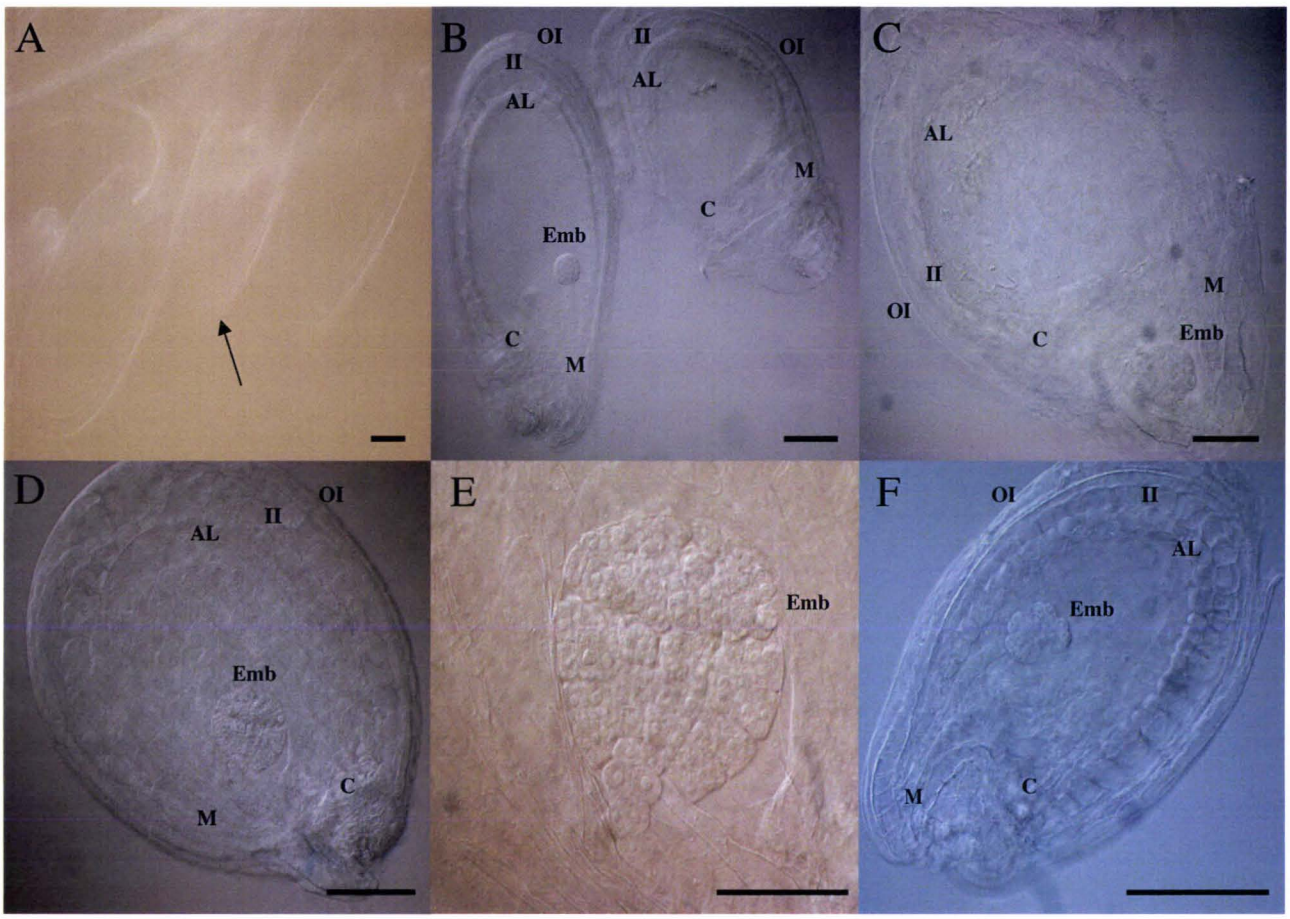


Figure 5.9. Seeds from Heterozygous *AGL62* T-DNA insertion line cleared with Hoyer's medium. Some seeds with defective embryos were observed from heterozygous *AGL62* T-DNA mutant fruits. A) 3DAP, arrow points to a seed that is considerably smaller than WT seeds. B) 3DAP, comparison between WT (left) and *agl62* mutant (right) seed. Mutant seed seems to have smaller starchy endosperm region and enlarged globular embryo. C) Close up photo of *agl62* mutant seed from B), large globular embryo was observed at the bottom. D) 4DAP, large globular embryo is observed. E) Close up photo of *agl62* mutant seed from D), enlarged globular embryo was observed. F) 5DAP, embryo was still remaining at globular stage. Emb - Embryo, OI - Outer Integument, II - Inner Integument, AL - Aleurone layer, M - Micropyle, C - Chalazal endosperm

Statistical analysis was then carried out to show the significance of the identification of enlarged globular embryo and study whether the enlarged globular embryo is caused by the loss-of-function of AGL62. A student's T-test was carried out with 3 and 4DAP fruits (Table 5.5). Up to 8 fruits were collected from WT and *agl62-1*/+ plants and their total seed contents were counted. Then the 3 and 4DAP seeds was Hoyer's cleared (Section 2.8.2.4) and observed for the enlarged globular embryo. Three days after pollination,

average of 16 seeds showed enlarged globular embryo in *agl62-1*/+ plants whereas the average number was only 3 seeds in WT (Table 5.5). Similarly 4DAP, *agl62-1*/+ plants showed average of 17 seeds with enlarged globular embryos while WT only showed average of 4 seeds (Table 5.5). P-value for the difference between WT and *agl62-1*/+ plants in number of seeds with enlarged embryos 3 and 4DAP were 0.0163 and 0.011 respectively confirming that the difference is statistically significant. In short, student's T-test showed that the difference in number of seeds with enlarged globular stage embryo observed in WT and *agl62-1*/+ lines were statistically significant and therefore phenotypic defects observed in putative *agl62-1/agl62-1* seeds was likely due to loss-of-function of AGL62.

3DAF

	1	2	3	4	5	6	7	8	Mean
WT	0/23	0/17	0/13	2/20	0/28	1/21	2/19		3.61
agl62-1/+	3/22	6/20	7/24	0/13	5/19	3/16	2/21	1/22	16.49

4DAF

	1	2	3	4	5	6	7	8	Mean
WT	1/14	0/22	1/18	2/23	3/27	0/12	0/16	1/20	4.68
agl62-1/+	3/17	1/22	2/19	2/20	3/19	5/20	4/13	2/17	17.79

Table 5.5. Number of seeds with defective phenotype in Heterozygous AGL62 T-DNA insertion lines. Seeds with defective phenotype described above were counted from WT and Heterozygous AGL62 T-DNA insertion lines after 3 and 4 days of flowering. Number of defective seeds is recorded over the total number of seeds in the fruit. P-value of samples 3DAF is 0.0163 and P-value of samples 4DAF is 0.011 respectively

5.4 Discussion

5.4.1 Possible role of AGL62 in plant development

PCR genotyping was carried out with over 100 plants from a F1 segregating population of *agl62-1*/+ selfed plants. The PCR genotyping was repeated twice but homozygous *agl62-1/agl62-1* plants were never recovered. Therefore we suspected that knocking out

the function of AGL62 results in a lethal phenotype. To understand function of AGL62, we investigated at which stage in development the loss of AGL62 function became lethal to plant by raising 3 hypotheses based on *AGL62* expression pattern shown by *in vitro* microarray (Table 3.1) and AGL62::GUS reporter gene system (Figures 3.4 to 3.7). Microarray and GUS staining assays together showed strong expression of *AGL62* in root meristem (Figure 3.4) and newly formed true leaves of the seedling (Figure 3.5), developing pollen at stage10 of flower development (Figure 3.6) (Smyth et al., 1990) and in developing seed at chalazal region 2 – 3DAP (Figure 3.7). All 3 hypotheses were based on these *AGL62* expression data. First possibility was that AGL62 plays a role in the transmission of T-DNA to the segregating population through male or female gametophyte and that loss-of-function of AGL62 resulted in T-DNA not being transmitted. The second hypothesis was that AGL62 regulates seed germination and knocking out its function resulted in seeds not germinating and hence mature *agl62-1/agl62-1* plant would never be recovered. The third hypothesis was that AGL62 regulates seed development and knocking out its function resulted in seed abortion. This hypothesis was favored because as well as the AGL62::GUS expression in the seed, the ratio of WT:*agl62-1*/+ in a segregating population of agl62-1/+ self supported this hypothesis (Table 4.2).

5.4.2 AGL62 and gametophyte development

To address the first possibility, Alexander staining and transmission assays that involved setting up controlled crosses between WT and *agl62-1*/+ were carried out to observe if mutation in AGL62 affected male or female gamete transmission. Initially resistance against the kanamycin conferred by T-DNA insertion was used as a marker to differentiate between WT and *agl62-1*/+ progeny in a F1 segregating population of the cross between WT and *agl62-1*/+ in tradition with the other research groups that study gametophytic mutants in plants (Howden et al., 1998). However, all the seeds plated on kanamycin supplemented MS media died. According to the SALK Institute T-DNA collections (http://signal.salk.edu/index.html) where the *AGL62* T-DNA insertion lines (SALK 107011) were obtained, T-DNA insertion genes can sometimes be silenced after several generations (http://signal.salk.edu/index.html). Therefore PCR genotyping was used as an alternative approach to differentiate between WT and *agl62-1*/+ in the F1

segregating population. In WT (F) and *agl62-1*/+ (M) crosses, the ratio between WT and *agl62-1*/+ in the F1 progeny was roughly 1:1 (Table 5.2) which suggested that transmission of *agl62-1* through male gamete is not affected by mutation in AGL62. On the other hand, in *agl62-1*/+ (F) and WT (M) crosses, the WT:*agl62-1*/+ ratio in the F1 progeny was roughly 1:2 (Table 5.2) which was not expected neither if AGL62 did or did not affect transmission through female gamete. The unusual WT:*agl62-1*/+ ratio may be explained if there was a contamination in seed samples by the *agl62-1*/+ seeds. In short the PCR genotyping of the segregating progenies showed that disruption of AGL62 function did not affect the transmission of *agl62-1* neither through the male nor female gametophyte (Table 5.2) but for *agl62-1*/+ (F) and WT (M) cross, more crosses may need to be set up to confirm the result.

In relation to expression of AGL62::GUS in developing pollen (Figure 3.6), it is possible that AGL62 has a role in pollen development. To address this question, Alexander staining was carried out to compare pollen viability between WT and *agl62-1*/+. The staining showed there was considerable variability in viability of pollen from *agl62-1*/+ plants ranging from 1.09% to 9.34% compared to WT's 7.68% (Table 5.1). Since pollen is a gametophytic (n) tissue in *agl62-1*/+ plants, half of the pollen should have WT AGL62 copy whereas the other half should have the non-functional *agl62* copy. Therefore 50% of pollen should be aborted in *agl62-1*/+ pollen. However, abortion rate of 2 *agl62-1*/+ plants (Ab435-1 and Ab435-3) were 1.47% and 1.09% which were much lower than WT's (7.68%) (Table 5.1). Also even the other 2 plants (Ab435-4 and Ab435-5) that showed higher pollen abortion rates of 9.34% and 8.22% were not much higher than that of WT's indicating that abortion of the pollen in *agl62-1*/+ plants were more likely caused by natural phenomenon rather than loss-of-function of AGL62. The Alexander staining showed that at least AGL62 alone does not affect the development of pollen. However, redundancy between closely related members is a common phenomenon in plant MADS-box genes (Liljegren et al., 2000) and AGL62 may well be playing a redundant role with other closely related members of MADS-box genes in transmission or development of pollen. Therefore production of multiple knock out lines may be required to address this issue.

5.4.3 AGL62 and germination

The second hypothesis was studied by comparing the rate of seed germination and seedling viability between WT seeds and F1 segregating seeds from *agl62-1*/+ self. In *Arabidopsis*, several seedling-lethal mutant lines have been identified by ethyl methanesulfonate (EMS) mutagenesis screen (Budziszewski et al., 2001) and interestingly it was suggested that often overlap existed in the embryo defective mutants and seedling lethal mutants (Budziszewski et al., 2001). Seed germination and seedling viability assay was designed based on Budziszewski's experiment (Budziszewski et al., 2001) in which they carried large-scale mutant screening to identify genes essential for seedling viability. Seed germination was observed after 3 days of plating the seeds on a MS plate and following that seedling viability was observed 5, 7 and 12 days after plating the seeds on MS plates. If AGL62 had an essential role in regulation of seed germination or seedling viability then, since one-quarter (25%) of F1 segregating population of *agl62-1*/+ self should be *agl62-1*/*agl62-1*, 25% of the F1 segregating seeds should die due to the loss-of-function of AGL62. However, no difference was observed between WT and F1 segregating population of *agl62-1*/+ self in neither germination rate (WT – 59.4% vs *agl62-1*/+ - 56.3) nor seedling viability (WT – 77.1% vs *agl62-1*/+ - 74.8%) (Table 5.3). The assay therefore indicated that either AGL62 has no function in seed germination and seedling viability or may have a redundant activity with other closely related members of MADS-box genes in regulation of seed germination and seedling development. To answer this question, production of multiple loss-of-function mutants of AGL62 and the other closely related type I MADS-box genes is required.

5.4.4 AGL62 and seed development

To investigate the possibility that knocking out the AGL62 function results in seed abortion, developing seeds from *agl62-1*/+ fruits were observed in detail. Seed consists of multi-layers of tissues including outer and inner integuments that are of maternal origin and endosperm and embryo that are formed as a result of double fertilization. In addition, seeds are enclosed in fruit walls making the organ not readily available for observation under microscope. Therefore developing seeds were dissected from WT and *agl62-1*/+ fruits and then cleared by Hoyer's medium (Section 2.8.2.4) and observed to

see if there was any defect in putative *agl62-1/agl62-1* developing seed. In line with expression of AGL62::GUS in seed chalazal region in 2 - 3 days old seeds (Figure 3.7), significant reduction in number of seeds were observed from *agl62-1/+* fruits after 4 days of pollination (Table 5.4). At 5days after pollination, compared to WT fruits that had roughly 40 seeds on average, *agl62-1/+* fruits only had average of 32 seeds per fruit. Since one-quarter (25%) of seeds from F1 segregating progeny of *agl62-1/+* plant should be *agl62-1/agl62-1*, if AGL62 had a role in seed development, then *agl62-1/agl62-1* seeds should abort due to loss-of-function of AGL62. The number of seeds in *agl62-1/+* fruits was roughly about 25% less than that of WT, which further supported the hypothesis that AGL62 played a role in seed development. Since embryo abortion is linked to seed abortion, Hoyer's cleared seeds (Section 2.8.2.4) were observed under microscope to see if there was any defective phenotype in developing embryo. Observation of putative *agl62-1/agl62-1* seeds under microscope showed enlarged globular stage embryo that never developed into heart stage embryo even in 5 days old seed (Figure 5.9) whereas in WT, embryo should be reaching late heart stage after 5 days of flowering (Figure 5.8) (Faure et al., 2002).

To conclude, AGL62::GUS expression (Figure 3.7), reduction in seed number (Table 5.4 and Figure 5.7) and identification of enlarged globular embryo (Figure 5.9) in *agl62-1/+* plants all indicated that AGL62 has a role in seed development.

5.4.5 Role of type I MADS-box genes in plant development

In plants, most functional characterization have been achieved with type II MADS-box genes that were shown to be especially important for flower development (Becker and Theissen, 2003). On the other hand, although type I MADS-box genes share the conserved DNA binding MADS domain with type II lineage, little is known about function of type I MADS-box genes in plant development (De Bodt et al., 2003). So far only 2 members of type I MADS-box genes were shown to have function in regulation of plant development (Kohler et al., 2003; Portereiko et al., 2006). It is interesting that the only 2 members of plant type I MADS-box genes that were functionally characterized – *AGL37* and *AGL80* – were also known to have function in seed development (Kohler et al., 2003; Portereiko et al., 2006). AGL37 is encoded by *PHERES1* and its expression is tightly regulated by Polycomb-group (PcG) proteins - MEDEA (MEA),

FERTILIZATION INDEPENDENT ENDOSPERM (FIE), and FERTILIZATION INDEPENDENT SEED2 (FIS2) by mechanism of chromatin remodeling (Kohler et al., 2003). These three proteins are called FIS-class proteins and are likely subunits of PcG protein complex (Makarevich et al., 2006). The single *fis*-class mutants all showed similar gametophytic maternal phenotype – all seeds with maternal mutant fis-class allele abort irrespective of the paternal allele (Reyes and Grossniklaus, 2003). In all the three *fis*-class mutants, *AGL37* is constitutively expressed and chromatin immunoprecipitation (ChIP) has shown that MEA and FIE can directly bind the *AGL37* promoter indicating that *AGL37* is a down stream target of the FIS protein complex (Kohler et al., 2005). However, the exact function of AGL37 is not yet characterized and production of AGL37 loss-of-function mutant and observation of seed development in the mutant would be necessary to show the function of AGL37 in the regulation of seed development (Kohler et al., 2005).

In these *fis*-class mutants cell division of the central cell is observed in the absence of fertilization and during embryogenesis, embryo development is arrested at enlarged heart stage and then seed abortion takes place (Luo et al., 2000). Because central cell division up to certain stage takes place without fertilization in *fis*-class mutants, fruit growth without pollination can be observed in *fis* mutants (Kohler and Makarevich, 2006). Fruit growth in the absence of pollination was observed in *agl62-1*/+ plants to see if AGL62 was a part of FIS-class PcG complex but no fruit growth was observed without fertilization (data not shown) indicating that AGL62 may not be a part of the FIS-class PcG complex. Several phenotypes such as delayed embryo development, enlarged globular stage embryo and seed abortion observed in *fis*-class mutants were also observed in developing putative *agl62-1*/*agl62-1* seeds indicating AGL62 could be a down stream target of FIS-class PcG complex and may be regulating seed development after being activated by the FIS-class PcG complex. However preliminary analyses of the *AGL62* promoter region did not detect any motif that is recognized by the FIS PcG complex (data not shown). Also microarray assays on *mea* mutant plants detected no change in level of *AGL62* expression in the *mea* background (Kohler et al., 2005). This indicated that there might be a parallel pathway that regulates seed development that does not involve FIS class complex. Their needs to be more detailed and refined analyses on the promoter region of *AGL62* to address whether *AGL62* could be a down-stream target of the FIS

PcG complex. Microarray assays on other *fis* class mutants may also be useful to study whether *AGL62* could be a downstream target of the FIS complex. In a future, it may also be interesting to cross *agl62-1/+* plants to *fis* class mutants and observe phenotype of the developing seed to study whether AGL62 and FIS-class proteins work in a same or different pathway to regulate seed development.

AGL80 is the other plant type I MADS-box gene that had its function characterized (Portereiko et al., 2006). *AGL80* was first identified in a mutant library screening for female gametophyte mutants (Portereiko et al., 2006). GFP reporter assays showed *AGL80* expression to be in central cell and the subsequent phenotypic analyses showed AGL80 was responsible for development of central cell into endosperm (Portereiko et al., 2006). In *agl80* mutants no endosperm formation was observed after fertilization and although egg cells become fertilized and forms zygote-like structure, the embryo will eventually abort due to the lack of endosperm.

MADS-box transcription factors belong to a large gene family and in *Arabidopsis*, 108 MADS-box genes have been identified so far (Martinez-Castilla and Alvarez-Buylla, 2004). There are over 60 members within type I lineage and Plant type I MADS box genes can be further divided into subclades - Mα, Mβ, and Mγ according to their MADS box DNA binding domain structure (Parenicova et al., 2003). Both AGL37 and AGL80 that are known to function in seed development belong to Mγ subclade (Parenicova et al., 2003; Portereiko et al., 2006). Furthermore 14 out of 16 Mγ type I MADS-box genes were shown to be expressed in fruits giving rise to the possibility that other members of the Mγ subclade may also play a role in seed development (Parenicova et al., 2003). Although *AGL62* belongs to Mα subclade, expression in seed and seed abortion phenotype in putative agl62-1/agl62-1 mutant seeds indicated AGL62 has a function in seed development and perhaps AGL62 may be interacting with the Mγ subclade type I MADS-box genes to regulate seed development. Yeast-two-hybrid assay showed AGL62 can physically interact with bothAGL37 and AGL80 (de Folter et al., 2005). Furthermore, apart from AGL97, AGL40 and AGL28 that belong to Mα subclade, all other interacting partners of AGL62 - AGL36, AGL38, AGL86 and AGL92 – belong to Mγ subclade (Parenicova et al., 2003). Also all these possible interacting partners are shown to be expressed in the fruit by RT-PCR (Parenicova et al., 2003). To address whether AGL62 may be interacting with these candidates or if *AGL62* expression is

affected by the candidates, 1.5kb *AGL62* upstream region was searched for CarG box – a conserved motif known to bind the MADS-box genes – but the CarG box sequence was not detected indicating AGL62 transcription may not be directly regulated by these candidate MADS-box genes. However, as suggested by the yeast-two-hybrid assay, they could be interacting with AGL62 through the MADS domain to regulate seed development in plants. Perhaps observation of spatial and temporal expression pattern of MADS-box genes by microarray to identify the genes that share the temporal and spatial expression domain with *AGL62* may be useful in a future to narrow down the candidates to identify the interacting partners of AGL62.

5.4.6 Embryo and endosperm development in putative agl62-1/agl62-1 mutant

In putative *agl62-1/agl62-1* mutant, defective embryo development was observed (Figure 5.9). In WT, 3 days after fertilization embryo reaches globular to transition state and 24 hours later the embryo should reach heart stage (Smyth et al., 1990). However, in putative *agl62-1/agl62-1* mutant, after 3 days of fertilization, the globular embryo is enlarged and does not proceed to heart stage but rather aborts (Figure 5.9). In plants, the close correlation between embryo and endosperm development was demonstrated in many cases confirming the need of endosperm on proper embryo development (Nickle and Meinke, 1998; Luo et al., 2000; Spillane et al., 2000; Reyes and Grossniklaus, 2003). In many mutants that show embryo defective phenotype such as *cyt1*, *kn*, and *hal* mutants, celluralization in endosperm development was shown to be absent (Nickle and Meinke, 1998). *AGL80* is a member of type I MADS-box gene that was shown to play a role in endosperm development (Portereiko et al., 2006). In *agl80* mutant, embryo development occurs in absence of the fertilization (Portereiko et al., 2006). However, the embryo only develops to zygote-like structure and then aborts due to absence of endosperm (Portereiko et al., 2006). Therefore it may also be possible that defective embryo observed in putative *agl62-1/agl62-1* mutant (Figure 5.9) is caused by absence of endosperm development in the putative *agl62-1/agl62-1* mutant.

Endosperm is formed as a result of double-fertilization where 2 pollen nuclei fuses with an egg cell (n) and central cell (2n) within ovule (Faure et al., 2002) (Figure 5.1). The fertilized egg cell develops into embryo (2n) while fertilized central cell (3n) develops into the endosperm (Faure et al., 2002). The endosperm is considered as the nutrient

source for the growing embryo and nourishes the embryo during embryogenesis until seed maturity (Dresselhaus, 2006). Endosperm development is mainly divided into 2 phases - synctium division where it produces a large mass of multinucleate cells and cellularization that divides the multinucleate cell into smaller cells with one nucleus each (Berger, 2003). During cellularization, endosperm is divided into 3 mitotic domains along anterior – posterior axis (Berger, 2003). These domains include micropylar region located adjacent to embryo, peripheral region that consist of most of the endosperm and posterior chalazal region that does not undergo cellularization and remain multinucleated (Berger, 2003). GUS staining assay showed AGL62::GUS expression in chalazal region of endosperm (Figure 3.7) and chalazal region was known to be responsible for transporting the nutrient from maternal tissue to the developing embryo (Berger, 2003). Therefore, in putative *agl62-1/agl62-1* mutant, perhaps the lack of nutrient delivered to developing embryo caused by defective chalazal endosperm may be the cause of defective phenotype observed (Figure 5.9) but confirming this hypothesis would require further phenotypic analyses on the putative *agl62-1/agl62-1* seeds.

Until recently only little was known about the molecular mechanisms underlying fertilization and subsequent endosperm and embryo development in *Arabidopsis*. However the sequencing of the entire genome of *Arabidopsis* coupled with the recent improvement in the identification of mutants started to give further insight into molecular processes during fertilization and subsequent endosperm and embryo development in *Arabidopsis* (Berger, 2003; Dresselhaus, 2006). When the pollen sperm cell (n) fertilizes the egg cell (n), 2 nuclei of these gametes need to fuse in a same cell-cycle phase in order to avoid aneuploidy and to allow proper subsequent embryo growth (Faure et al., 2002). Therefore, there must be some sort of cell cycle regulation in both the male and female gametes (Faure et al., 2002). Compared to animal gametes, regulation of the cell cycle in plant gametes remains largely unknown (Yadegari and Drews, 2004). Recently it was suggested that in *Arabidopsis*, sperm cells are likely to be in G2 phase when they are delivered to the ovule (McCormick, 2004). In female gametes, cell-cycle arrest is mediated by FIS PcG complex FIS, MEA, and FIE and loss-of-function mutations in any one of these genes result in the autonomous initiation of cell division in the central cell even without the fertilization (Luo et al., 2000). In *agl62-1/+* plants, fruits failed to form in absence of pollination (data not shown). Furthermore microarray assay on *mea* mutant

showed no change in expression level of *AGL62* (Kohler et al., 2005) indicating that *AGL62* may not be a downstream target nor interacting partner of FIS complex. Therefore there is a possibility that pathway that regulates seed development that does not involve FIS complex exists which AGL62 is part of.

Recent experimental evidences showed the importance of epigenetic control of gene expression on fertilization induced seed development (Chanvivattana et al., 2004; Ingouff et al., 2005; Kohler and Makarevich, 2006). Of special interest is the regulation of imprinted gene expression in developing endosperm by FIS PcG complex mediated methylation (Luo et al., 2000; Kohler and Makarevich, 2006). Pc-G genes initially identified in *Drosophila melanogastor* are a group of genes that maintain repression of target gene expression by methylating the target histone lysine residue and hence remodeling the chromatin structure (Muller et al., 2002). In *Drosophila*, *pc-g* mutants failed to maintain transcriptional repression of homeobox genes (Muller et al., 2002). Pc-G genes assemble in 2 complexes to carry out their function of repressing expression of target genes (Muller et al., 2002). First, Polycomb Repressive Complex 2 (PRC2) or E(Z)/ESC complex that consists of 4 Pc-G proteins – Enhancer of Zeste (E(z)), Extra Sex Comb (ESC), Supressor of Zeste 12 (Su(z)12) and NURF-55 – bind to histone H3 and methylate lysine27 residue (H3K27) to create an epigenetic mark (Muller et al., 2002). The second complex PRC1 is then recruited to H3K27 via interaction of its component with the methylated H3K27 (Muller et al., 2002). PRC1 complex also consists of 4 proteins – Polycomb (PC), Polyhomeotic (PH), Posterior Sex Combs (PSC), and dRing (Muller et al., 2002). After binding to a target gene via methylated H3K27, PRC1 interacts with SWI/SNF chromatin remodeling complex to modify chromatin structure in a way that transcription initiation is blocked (Muller et al., 2002).

In plants, Pc-G genes were first identified in a genetic screen for mutants with defective seed development (Spillane et al., 2000). Several Pc-G genes have been isolated since and the presence of a PRC2 like complex was confirmed (Schubert et al., 2005). However, there is still no evidence for the existence of PRC1-like complex in plants yet (Schubert et al., 2005). To date, 3 PRC2 like complexes have been characterized. These include MEA-FIE complex, CLF complex and VRN complex (Yadegari et al., 2000; Michaels et al., 2003; Reyes and Grossniklaus, 2003; Schonrock et al., 2006). Of these 3 complexes, MEA-FIE complex has been shown to play an important role in gametophyte

and early seed development (Luo et al., 2000; Dresselhaus, 2006). MEA-FIE complex consists of 3 proteins – MEA, FIS2, and FIE – and these 3 proteins are homologues of E(z), Su(z)12, and ESC respectively (Pien and Grossniklaus, 2007). The *mea, fis2,* and *fie* mutants all show similar phenotypes – autonomous endosperm development in absence of fertilization, formation of zygote-like structure and failure of mutant seeds to develop beyond endosperm cellulalization (Wang and Ma, 2007). These common phenotypes indicated that MEA, FIE, and FIS2 work in the same pathway to suppress various aspects of seed development in the absence of fertilization and that fertilization inactivates MEA-FIE complex to initiate seed development (Makarevich et al., 2006). Since MEA and FIE are homologues of E(z) and ESC respectively and E(z) and ESC physically interact in Drosophila to repress target genes' expression, one can postulate that MEA and FIE do the same in Arabidopsis as well (Pien and Grossniklaus, 2007). A yeast-2-hybrid assay was carried out to test this hypothesis which showed MEA and FIE do indeed interact with each other (Kohler et al., 2003). The assay also showed FIS2 does not physically interact with either MEA or FIE. This implies a still unidentified component in MEA-FIE complex that attaches FIS2 to the rest of complex (Kohler et al., 2003). In a future, to study whether AGL62 may form complex with the MEA-FIE complex, protein-protein interaction assay such as yeast-2-hybrid assay may be done to see whether AGL62 can physically interact with them.

Microarray assay was carried out in *Arabidopsis fis* class mutants to identify the down-stream targets of the MEA-FIE complex (Kohler et al., 2005). In the assay, change in *AGL62* expression level was not detected (Kohler et al., 2005). However, *PHERES1 (PHE1)* expression was up-regulated in *fis*-class mutants indicating PHE1 as a target of MEA-FIE complex (Kohler et al., 2005). *PHE1* encodes a MADS-box transcription factor AGL37 and appears to play an important role in seed development (Kohler et al., 2005). In WT, *PHE1* is not expressed in the central cell before pollination but the expression is induced 1 – 2 DAP in seeds containing pre-globular stage embryo (Kohler et al., 2005). In *fis* class mutants, *PHE1* expression is observed much earlier and is expressed directly after pollination (0DAP) because repression from MEA-FIE complex is released in the *fis*-class mutants (Kohler et al., 2005). Subsequent ChIP assays showed MEA and FIE can bind *PHE1* promoter and confirmed *PHE1* as direct target of MEA-FIE Pc-G complex (Kohler et al., 2005). When *PHE1* expression was down-regulated in

a *mea* mutant back-ground, the seed abortion phenotype of *mea* mutant was rescued confirming the role PHE1 plays in seed development (Kohler et al., 2005). Following the discovery of the Pc-G complex genes in *Arabidopsis*, molecular mechanisms underlying regulation of imprinting gene expression in seed was studied (Kohler and Makarevich, 2006). *MEA* was the first gene in *Arabidopsis* to be identified as the imprinted gene (Vinkenoog et al., 2003; Wang and Ma, 2007). In developing endosperm (3n), there are 2 maternal alleles of *MEA* and 1 paternal *MEA* allele. However, genetic analysis showed that maternal and not paternal WT *MEA* allele is required for embryo and endosperm development and that while maternal alleles are expressed, paternal allele expression is repressed in developing endosperm (Gehring et al., 2006). This MEA imprinting mechanism involves 2 proteins - DEMETER (DME) and MET1(Gehring et al., 2006). DME encodes a DNA glycosylase and is required for expression of MEA in endosperm (Gehring et al., 2006). In *dme* mutants, a parent-of-origin seed viability phenotype was observed (Gehring et al., 2006). In *dme* mutants, seed viability depended on the maternal allele of *DME* (Gehring et al., 2006). In other words, when *dme* mutant flower was pollinated with WT pollen, an aborted seed phenotype developed whereas when a WT flower was pollinated with *dme* mutant pollen, seed development was rescued and seeds developed normally (Gehring et al., 2006). While DME induces *MEA* expression, MET1 has been shown to repress *MEA* expression (Kohler et al., 2005). Maternal *dme* or *mea* mutants resulted in seed abortion but seeds with maternal *dme;met1* mutant were able to develop as long as maternal *MEA* allele was WT (Kohler et al., 2005). This confirms that MET1 counteracts the activation of *MEA* by DME. Subsequent experiments have shown that before fertilization, the *MEA* gene is methylated by the action of MET1 in both the central cell and pollen sperm cell (Kohler and Makarevich, 2006). In the sperm cell, the Pc-G complex is therefore not formed and *PHE1* is expressed in the sperm cell (Kohler and Makarevich, 2006). In the central cell however, DME activity removes methylation from the promoter of *MEA* releasing *MEA* expression from repression (Kohler and Makarevich, 2006). Hence MEA protein is expressed and can form MEA-FIE complex leading to repression of *PHE1* expression in the central cell (Spillane et al., 2000). After fertilization, methylated maternal MEA continues to be expressed and MEA-FIE complex will target paternal MEA allele in repressed state (Spillane et al., 2000). This novel self-imprinting mechanism of maternal MEA allele repressing paternal MEA allele

can give further insight into the molecular mechanism of imprinting in plants. *PHE1* is another example of an imprinted gene but unlike *MEA*, *PHE1* shows paternal allele expression and maternal allele repression (Kohler et al., 2005). In the central cell before fertilization, *PHE1* expression is repressed by MEA-FIE complex and after fertilization in the endosperm, maternal *PHE1* expression remains to be repressed by MEA-FIE complex while paternal *PHE1* is expressed (Kohler et al., 2005). In our study, it appears that AGL62 may play a role in seed development as putative *agl62-1/agl62-1* developed a seed abortion phenotype due to arrest of the embryo growth at enlarged globular stage. There is no data however, to support that expression of *AGL62* is regulated by imprinting or is AGL62 is a target of MEA-FIE complex. In fact, microarray assay (Kohler et al., 2005) and preliminary promoter analyses on *AGL62* promoter showed that *AGL62* is unlikely to be a downstream target of MEA-FIE complex. Therefore it is possible that a parallel pathway exists that regulates seed development that involves AGL62 but not the MEA-FIE complex.

5.4.7 AGL62 and communication between seed components

Since embryo lethal phenotype observed in putative *agl62-1/agl62-1* seeds are unlikely to be caused by the FIS-class complex, there may be a different pathway that regulate seed development via AGL62. A strong AGL62::GUS expression was identified in chalazal endosperm (Figure 3.7) between 2 - 3 days of pollination. By this stage, embryos should be reaching globular stage but cellularization have not started at the peripheral endosperm region (Berger, 2003). The Seed consists of embryo, endosperm, and seed integument, which is of maternal origin. During seed growth, these 3 tissues must communicate and co-ordinate with each other for proper seed development (Berger, 2003). Recently quite a few mutants have been identified that supports the signaling between embryo, endosperm and integument for a proper seed development (Nickle and Meinke, 1998; Garcia et al., 2003; Portereiko et al., 2006). For example, in a mutagenesis assay screening for seeds with reduced size, 2 new genes *HAIKU1 (IKU1)* and *HAIKU2 (IKU2)* have been identified (Garcia et al., 2003). Morphologically *haiku (iku)* mutant seeds are more spherical and reduced in size compared to the oblong structure of WT seeds (Garcia et al., 2003). Further phenotypic analysis showed that *iku* mutant seeds have reduced endosperm size, and embryo growth (Garcia et al., 2003). Most *iku* mutant

seeds are still capable of reaching the maturity and germinate although a small portion of them (< 10%) are so reduced in size that they abort before maturity (Garcia et al., 2003). When these *iku* mutant seeds were planted, the mutant developed into morphologically normal plants suggesting *iku* mutations affect only seed size and not plant morphogenesis (Garcia et al., 2003). This is in contrast to other classes of mutants that gives reduced seed size (Berger, 2003). In these mutants, other aspects of plant development are affected as well including *exs* mutant with male-sterility, and *ctr1* mutant with defect in cell elongation and ethylene signal transduction (Berger, 2003). Therefore *iku* mutants represent a new class of mutant that specifically affects seed size (Garcia et al., 2003). In *iku* mutants, pre-mature cellularization occurs in the endosperm and the endosperm is completely cellularized when embryo is still at triangular stage while in WT, cellularization does not initiate until embryo reaches heart stage (Smyth et al., 1990; Garcia et al., 2003). The premature cellularization in endosperm of *iku* mutant affects subsequent proliferation of cellularized endosperm and in turn reduces embryo proliferation after early torpedo stage resulting in reduced embryo growth at later stages of embryogenesis (Garcia et al., 2003). In WT endosperm, cellulalization is absent in the posterior chalazal region (Berger, 2003). Instead, the posterior region of WT endosperm is characterized by the presence of 3 multinucleate structures – nucleocytoplasmic domain (NCD) consisting of a single nuclei surrounded by cytoplasmic units, nodule which is a fusion product of NCD and posterior-most cyst formed by a fusion of nodules (Berger, 2003). Using a marker line that expresses GFP in posterior region of endosperm crossed to *iku* mutants, it was found that the overall size of posterior region is reduced in *iku* mutants (Garcia et al., 2003). Also, in some cases cellularization was observed in the posterior region of the endosperm in *iku* mutants suggesting that the *iku* mutation may cause displacement of the boundary between peripheral and posterior region of the endosperm (Garcia et al., 2003). When the *iku* mutant was pollinated by WT pollen, seed development was normal indicating that the *iku* mutation does not have a maternal sporophytic effect and hence the growth of the integuments can not be directly affected by *iku* mutation (Garcia et al., 2003). However, to accommodate reduced seed size, integument size must be affected as well in *iku* mutants (Garcia et al., 2003). In *iku* mutants, it was found that cell elongation was reduced in integuments and resulted in spherical rather than oblong seeds (Garcia et al., 2003). This indicated an as yet

unidentified signal from the endosperm that initiates cell elongation in the integuments (Garcia et al., 2003). In summary, mutations in *HAIKU* genes *IKU1* and *IKU2* both resulted in reduced endosperm size due to pre-mature completion of cellularization. This endosperm reduced in size in turn sends an as yet unidentified signal to seed integuments and reduces cell elongation in integuments giving rise to seeds with a reduced size by roughly a quarter. Reduced endosperm size also causes reduction of embryo size after heart stage due to decreased cell proliferation in the embryo after the heart stage. Decreased cell proliferation is likely to be caused by defective endosperm in the HAIKU mutants.

In WT endosperm, the ratio of maternal:paternal copy of genome is 2:1 and when this dosage balance is impaired, it has been reported that timing of celluralization and degree of proliferation in endosperm are primarily affected (Luo et al., 2000). In Arabidopsis, excess paternal dosage in endosperm results in increased seed size similar to the phenotype observed in *fis* class mutants (Spillane et al., 2000). On the other hand, excess maternal dosage in endosperm reduces the seed size similar to the *iku* mutants (Berger, 2003). These results suggest that endosperm plays a key role in determining seed size. In a segregating population of *agl62-1*/+ plant, some seeds showed a reduction in seed size (data not shown) but when these smaller seeds were germinated on Whattman paper, their germination rates were comparable to that of WT. Also the number of these small seeds in segregating population of *agl62-1*/+ was not statistically significant and no phenotypic defect was observed in their endosperm (data not shown) indicating that the reduction in seed size or unbalanced parental dosage in endosperm is not caused by loss-of-function mutation in AGL62. Also although the delay in embryo growth and arresting of the embryo growth was observed at enlarged globular stage (Figure 5.9), no defect in the endosperm was observed in their seeds which is unusual compared to other known mutants with defect in endosperm cytokinesis in a sense that most mutants that show defective cytokinesis in endosperm also show the arresting of embryo growth at some stage. However, there have been reports that showed there is a mutant with defective embryo without any phenotype in endosperm (Nickle and Meinke, 1998). Perhaps loss-of-function mutation in AGL62 may result in a similar manner and AGL62 may be specifically affecting embryo development but does not function to regulate endosperm or other components of the developing seeds.

Our phenotypic analyses in this project mainly focused on embryo development and only little analyses were carried out for endosperm (data not shown). Therefore in a future, perhaps phenotypic analyses on endosperm development in putative *agl62-1/agl62-1* mutant seed can be carried out that may give further insight on timing of spatial organization of endosperm and may give us more information about function of AGL62. It is particularly interesting that very recently a paper was published that reported the function of AGL62 in regulation of cellularization in *Arabidopsis* endosperm (Kang et al 2008). Their data on defects observed in *agl62* mutant seeds (Kang et al 2008) correspond to our results which further confirms that AGL62 has a role in seed development particularly at the cellularization stage. Since AGL62 functions as a transcription factor, it would be interesting in a future to identify down-stream target of AGL62 probably using microarray to further investigate on AGL62 function and identify the complex gene network that regulate endosperm development in plants.

Chapter 6 General conclusion and future direction

Over 100 MADS-box genes have been identified in *Arabidopsis thaliana* of which roughly 60 of them belong to type I lineage. However, most of the functionally characterized MADS-box genes belong to type II lineage that contains about 40 MADS-box genes and only a few type I MADS-box genes have been functionally characterized such as *AGL37* and *AGL80* (Kohler et al., 2003; Portereiko et al., 2006). Although mammalian and fungal type I MADS-box genes are well studied and known to play various roles in development (Treisman and Ammerer, 1992), in plants little is known about the role type I MADS-box genes play in development. Therefore in this project, we aimed to functionally characterize AGL40 and AGL62 by analyzing the T-DNA insertion mutant lines and give further insight into whether type I MADS-box genes play a role in plant development.

We have managed to identify loss-of-function mutant of AGL40 (*agl40-1*) through PCR genotyping and RT-PCR (Figures 4.6 & 4.7). However, subsequent phenotypic analyses on tissues that express *AGL40* that include inflorescence and silliques did not detect any significant difference from that of WT indicating the possibility that AGL40 has a redundant activity with other members of MADS-box gene family – a phenomenon commonly observed in MADS-box gene family (Martinez-Castilla and Alvarez-Buylla, 2004).

We therefore produced double *agl40;agl62* knock out mutant by crossing *agl40-1* and heterozygous *AGL62* T-DNA insertion mutant (*agl62-1/+*) because *AGL62* is the most closely related MADS-box gene to *AGL40*. Subsequent PCR genotyping of the F1 and F2 generation managed to identify the *agl40-1/agl40-1;agl62-1/+* mutant (Figure 4.12). However, the double *agl40;agl62* mutant had a phenotype similar to that of single *agl62-1* mutant indicating that AGL40 and AGL62 are not redundant to each other and that AGL40 has another genes which it works redundantly with. This experiment also showed that *AGL62*, although it shared the expression domain with *AGL40*, has its own function especially in seed development.

We have also carried out PCR genotyping with the *AGL62* T-DNA insertion lines to identify the AGL62 loss-of-function mutant. However, we did not detect any homozygous T-DNA insertion in *AGL62* even after the large-scale PCR genotyping involving over 100 plants from *AGL62* T-DNA insertion lines (Figure 4.11 and Table 4.2). We therefore suspected that loss of AGL62 function result in lethal phenotype. From the *AGL62::GUS* expression assay and web-based microarray, we confirmed that *AGL62* is expressed in developing pollen (Figure 3.6), developing seed in a chalazal endosperm region (Figure 3.7) and root meristem and leaves of the developing seedling (Figure 3.4 & 3.5). According to the expression profile of *AGL62*, we established 3 hypotheses as to what stage the loss of function in AGL62 becomes lethal.

The first possibility is that AGL62 plays a role in gametophyte development and knocking out AGL62 function results in aborted fertilization subsequently. *AGL62::GUS* expression assay did not detect *AGL62* expression in female gametophyte so we have focused on the male gametophyte (pollen). We have used Alexander staining to differentiate between the viable and aborted pollen in *agl62-1* plants and compared the number of aborted pollen to that of WT. We discovered that number of dead pollen in *agl62-1* vary quite a bit and are not necessarily that different from that of WT (Table 5.1). We have also looked at the transmission assay where we crossed *agl62-1* to WT and investigated whether *agl62-1* can be transmitted through male or female gametophytes. In *agl62-1* (Male) x WT (Female) cross, the ratio of WT:*agl62-1* in a progeny was roughly 1:1 (Table 5.2) indicating *agl62-1* can be transmitted through pollen and that loss of function of AGL62 does not result in male gametophyte lethality.

Second hypothesis we tested was that because *AGL62* was expressed in root meristem and new leaves of developing seedlings, perhaps the loss-of-function of AGL62 result in aborted seed germination. To this end, we plated roughly 300 seeds each from WT and *agl62-1* lines and compared their germination rates (Table 5.3). We compared the germination and subsequent seedling survival rate of WT and *agl62-1* lines 3,5,7 and 12 days after plating the seeds but did not detect any difference between WT and *agl62-1* indicating loss-of-function of AGL62 is unlikely to affect seed germination.

The last hypothesis we tested was that loss of function of AGL62 result in embryonic lethal phenotype. Embryonic lethal mutants typically have reduced seed set in a sillique

and distorted segregation ratio (Drews and Yadegari, 2002). When large scale PCR genotyping was carried out with segregating population of *agl62-1* (Table 4.2), the ratio between WT:*agl62-1* in a progeny was 1:2 which also supports that loss of function of AGL62 result in embryonic lethality (Figure 5.3). We first compared the number of seeds in WT and *agl62-1* fruits and identified that *agl62-1* lines have significantly less seeds compared to WT after 4 days of flowering (Figure 5.7 and Table 5.4). After 4 days of flowering, embryo in the developing seed reaches a triangular to heart stage (Figure 5.8). In *AGL62::GUS* expression assay, *AGL62* expression was evident in chalazal endosperm region of the developing seed after 3 days of flowering (Figure 3.7) and coincide with the timing when we started to observe the reduction in seed number after 4 days of flowering (Table 5.4). We therefore had a more detailed observation at the embryos 3 to 5 days after flowering in *agl62-1* line and discovered that some seeds in *agl62-1* lines were much smaller and some of the embryos had enlarged embryos that never proceeded to heart stage even after 5 days of flowering and then aborted (Figure 5.9). These results indicated that AGL62 functions in chalazal endosperm region of developing seed after 3days of flowering and aids the proper development of embryo in someway.

In WT plants, endosperm (3n) is formed as a result of fertilization of central cell (2n) by pollen nuclei (Faure et al., 2002). Endosperm is very important component of the developing seed as it provides the essential nutrients to the growing embryo and developing seedling during germination (Faure et al., 2002). In *Arabidopsis* seeds, endosperm development is divided into 2 main phases – initial syncytial phase followed by the cellularization phase (Faure et al., 2002). In syncytial phase, endosperm nuclei undergo mitosis without cytokinesis resulting in a multinucleate cell containing as high as over 100 nuclei (Faure et al., 2002). Then at a specific stage during development endosperm becomes cellularized and cellularized endosperm cells differentiate into several cell types and perform specific roles for the further endosperm and embryo growth (Faure et al., 2002) In *Arabidopsis*, cellularization starts roughly 3 days after flowering when embryo reaches transition stage (Faure et al., 2002). Many reports have shown that timing of endosperm cellularization is an important control point for seed development as it correlates with extent of nuclear proliferation, seed size, and grain weight (Faure et al., 2002). For example premature endosperm cellularization result in

reduced nuclear proliferation and reduced seed size and delayed cellularization leads to increased nuclear proliferation and increased seed size (Wang and Ma, 2007).

Molecular mechanism underlying endosperm cellularization is not yet fully understood but many reports have shown that FIS class PcG complex play a central role in endosperm development (Kohler et al., 2003; Schubert et al., 2005; Wang and Ma, 2007). Interestingly 2 papers were published recently that reported on roles of 2 new type I MADS-box genes - *AGL23* (Colombo et al., 2008) and *AGL62* (Kang et al., 2008) in seed development. Particularly report on *AGL62* by Kang was interesting as their expression analyses of *AGL62* was similar to that of ours' and showed expression of *AGL62* in endosperm at syncytial phase but never in embryo (Kang et al., 2008). They have shown that their *agl62* seeds had premature endosperm cellularization event with reduced seed size and numbers and abnormal embryo development (Kang et al., 2008). This result was similar to what we identified in our phenotypic analyses where *agl62-1* seeds had enlarged globular embryo (Figure 5.9) that never proceeded to the heart stage and the subsequent reduction in seed size as well as the reduction in seed number in the *agl62-1/+* plants (Figure 5.7 and Table 5.4) further indicating that AGL62 plays a role in seed development through regulation of endosperm cellularization. In their paper, Kang reported that *AGL62* expression is not down-regulated in *fis-class* mutants after embryo reaches transition stage and delayed endosperm cellularization associated with it indicating *AGL62* as a down stream target of the FIS class PcG complex (Kang et al., 2008). In our preliminary promoter region analysis (data not shown), we did not detect any target motif for PcG complex in *AGL62* promoter but more detailed search would be required now to identify the target site for PcG complex in *AGL62*.

To date, 4 type I MADS-box genes have been shown to play a role in plant development – *AGL37* (*PHERES1*), *AGL80*, *AGL23*, and *AGL62* (Kohler et al., 2005; Portereiko et al., 2006; Colombo et al., 2008; Kang et al., 2008). Interestingly all these 4 type I MADS-box genes have a regulatory role in seed development and all are shown to be down stream target of FIS class PcG complex. Identification of genetic network regulating seed development would be a significant importance both for biological and agricultural purposes. Perhaps plant type I MADS-box genes are generally involved in seed development much like type II MADS-box genes that are well known for their regulatory

role in floral organ development. However, more type I MADS-box genes need to be functionally characterized to confirm whether the statement is correct or not. Kang et al mentioned physical interaction of AGL62 and AGL80 in syncytial endosperm (Kang et al., 2008) and that may be in a future, knocking out other interacting partners of AGL62 may be an interesting approach to further studying the function of AGL62. Also since AGL62 is a transcription factor, we may also try transcription profiling experiment such as ChIP assay to identify the down-stream target of AGL62 transcription factor to give an insight into a complex genetic net-work that regulate endosperm cellularization.

Chapter 7 References

ABRC (www.biosci.ohio-state.edu/pcmb/Facilities/abrc/abrchome.htm)

Alexander, M. P. (1969). "Differential staining of aborted and nonaborted pollen." Stain technology 44(3): 117 - 122.

Alvarez-Buylla, E.R. (2001). MADS-box gene evolution beyond flower: expression in pollen, endosperm, guard cells, roots and trichomes (vol 24, pg 457, 2000). Plant Journal 25, 593-593.

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 between eudicots and monocots. Molecular Cell 5, 569-579.

Arabidopsis Genome, I. (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 408, 796-815.

Arsenian, S., Weinhold, B., Oelgeschlager, M., Ruther, U., and Nordheim, A. (1998). Serum response factor is essential for mesoderm formation during mouse embryogenesis. Embo Journal 17, 6289-6299.

Baum, D.A. (1998). The evolution of plant development. Current Opinion in Plant Biology 1, 79-86.

Bebien, M., Salinas, S., Becamel, C., Richard, W., Linares, L., and Hipskind, R.A. (2003). Immediate-early gene induction by the stresses anisomycin and arsenite in human osteosarcoma cells involves MAPK cascade signaling to Elk-1, CREB and SRF. Oncogene 22, 1836-1847.

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, A., Winter, K.U., Meyer, B., Saedler, H., and Theissen, G. (2000). MADS-box gene diversity in seed plants 300 million years ago. Molecular Biology and Evolution 17, 1425-1434.

Bedinger, P. (1992). The Remarkable Biology of Pollen. Plant Cell 4, 879-887.

Bent, S. J. C. a. A. F. (1998). "Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana." Plant journal 16(6): 735-43.

Berger, F. (2003). Endosperm: the crossroad of seed development. Current Opinion in Plant Biology 6, 42-50.

Budziszewski, G. J., S. P. Lewis, et al. (2001). "Arabidopsis genes essential for seedling viability: Isolation of Insertional mutants and molecular cloning." Genetics 159(4): 1765-1778.

Chanvivattana, Y., A. Bishopp, et al. (2004). "Interaction of polycomb-group proteins controlling flowering in Arabidopsis." Development 131(21): 5263-5276.

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

Colombo, M., Masiero, S., Vanzulli, S., Lardelli, P., Kater, M.M., and Colombo, L. (2008). AGL23, a type I MADS-box gene that controls female gametophyte and embryo development in Arabidopsis. Plant Journal **54,** 1037-1048.

David R. Smyth, J. L. B., and Elliot M. Meyerowitz (1990). "Early flower development in Arabidopsis." The Plant Cell 2: 755 - 767.

Davis, F.J., Gupta, M., Pogwizd, S.M., Bacha, E., Jeevanandam, V., and Gupta, M.P. (2002). Increased expression of alternatively spliced dominant-negative isoform of SRF in human failing hearts. American Journal of Physiology-Heart and Circulatory Physiology 282, H1521-H1533.

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

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

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., Davies, B., and Angenent, G.C. (2005). Comprehensive interaction map of the Arabidopsis MADS box transcription factors. Plant Cell 17, 1424-1433.

Diethard Mattanovich, F. R., Artur da Camara Machado, Margit Laimer, Ferdinand Regner, Herta Steinkellner, Gottfried Himmler and Herman Katinger (1989). "Efficient transformation of Agrobacterium spp. by electroporation " Nucleic Acids Research 17: 6747.

Dresselhaus, T. (2006). Cell-cell communication during double fertilization. Current Opinion in Plant Biology 9, 41-47.

Drews, G.N., and Yadegari, R. (2002). Development and function of the angiosperm female gametophyte. Annual Review of Genetics 36, 99-124.

Edwards, K. J., C; Thompson, C (1991). "A SIMPLE AND RAPID METHOD FOR THE PREPARATION OF PLANT GENOMIC DNA FOR PCR ANALYSIS." NUCLEIC ACIDS RESEARCH 19((6)): 1349-1349.

Escalante, R., and Sastre, L. (1998). A serum response factor homolog is required for spore differentiation in Dictyostelium. Development 125, 3801-3808.

Faure, J.E., Rotman, N., Fortune, P., and Dumas, C. (2002). Fertilization in Arabidopsis thaliana wild type: Developmental stages and time course. Plant Journal 30, 481-488.

Garcia, D., V. Saingery, et al. (2003). "Arabidopsis haiku mutants reveal new controls of seed size by endosperm." Plant Physiology 131(4): 1661-1670.

Gehring, M., J. H. Huh, et al. (2006). "DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-specific demethylation." Cell 124(3): 495-506.

Hellens, R., Mullineaux, P., and Klee, H. (2000). A guide to Agrobacterium binary Ti vectors. Trends in Plant Science 5, 446-451.

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.

Howden, R., S. K. Park, et al. (1998). "Selection of T-DNA-tagged male and female gametophytic mutants by segregation distortion in Arabidopsis." Genetics 149(2): 621-631.

http://www.arabidopsis.org/ The Arabidopsis information resource

ICHIRO KASAJIMA, Y. I., NAOKO OHKAMA-OHTSU, HIROAKI HAYASHI, TADAKATSU YONEYAMA and TORU FUJIWARA (2004). "A Protocol for Rapid DNA Extraction From Arabidopsis thaliana for PCR Analysis " Plant Molecular BiologyReporter 22: 49–52.

Ingouff, M., J. Haseloff, et al. (2005). "Polycomb group genes control developmental timing of endosperm." Plant Journal 42(5): 663-674.

Irish, V.F. (2003). The evolution of floral homeotic gene function. Bioessays 25, 637-646.

Iyer, D., Belaguli, N., Fluck, M., Rowan, B.G., Wei, L., Weigel, N.L., Booth, F.W., Epstein, H.F., Schwartz, R.J., and Balasubramanyam, A. (2003). Novel phosphorylation target in the serum response factor MADS box regulates alpha-actin transcription. Biochemistry 42, 7477-7486.

Jefferson, R.A., Kavanagh, T.A., and Bevan, M.W. (1987). Gus Fusions - Beta-Glucuronidase as a Sensitive and Versatile Gene Fusion Marker in Higher-Plants. Embo Journal 6, 3901-3907.

Jia, H.W., Chen, R., Cong, B., Cao, K.M., Sun, C.R., and Luo, D. (2000a). Characterization and transcriptional profiles of two rice MADS-box genes. Plant Science 155, 115-122.

Jia, H.W., Luo, D., Cong, B., Gao, Z.Z., Shao, J.Y., Cao, K.M., and Sun, C.R. (2000b). Cloning and expression of a new MADS-box gene from rice. Acta Botanica Sinica 42, 490-495.

Johansen, B., Pedersen, L.B., Skipper, M., and Frederiksen, S. (2002). MADS-box gene evolution-structure and transcription patterns. Molecular Phylogenetics and Evolution 23, 458-480.

Kang, I.H., Steffen, J.G., Portereiko, M.F., Lloyd, A., and Drews, G.N. (2008). The AGL62 MADS domain protein regulates cellularization during endosperm development in Arabidopsis. Plant Cell **20,** 635-647.

Kellogg, E.A. (2004). Evolution of developmental traits. Current Opinion in Plant Biology 7, 92-98.

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.

Kohler, C., and Makarevich, G. (2006). Epigenetic mechanisms governing seed development in plants. Embo Reports 7, 1223-1227.

Kohler, C., Page, D.R., Gagliardini, V., and Grossniklaus, U. (2005). The Arabidopsis thaliana MEDEA Polycomb group protein controls expression of PHERES1 by parental imprinting. Nature Genetics 37, 28-30.

Kohler, C., Hennig, L., Spillane, C., Pien, S., Gruissem, W., and Grossniklaus, U. (2003). The Polycomb-group protein MEDEA regulates seed development by controlling expression of the MADS-box gene PHERES1. Genes & Development 17, 1540-1553.

Lawton-Rauh, A.L., Alvarez-Buylla, E.R., and Purugganan, M.D. (2000). Molecular evolution of flower development. Trends in Ecology & Evolution 15, 144-149.
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.

Luo, M., Bilodeau, P., Dennis, E.S., Peacock, W.J., and Chaudhury, A. (2000). Expression and parent-of-origin effects for FIS2, MEA, and FIE in the endosperm and

embryo of developing Arabidopsis seeds. Proceedings of the National Academy of Sciences of the United States of America 97, 10637-10642.

Makarevich, G., Leroy, O., Akinci, U., Schubert, D., Clarenz, O., Goodrich, J., Grossniklaus, U., and Kohler, C. (2006). Different Polycomb group complexes regulate common target genes in Arabidopsis. Embo Reports 7, 947-952.

Marais, R., Wynne, J., and Treisman, R. (1993). The Srf Accessory Protein Elk-1 Contains a Growth Factor-Regulated Transcriptional Activation Domain. Cell 73, 381-393.

Marais, R.M., Hsuan, J.J., McGuigan, C., Wynne, J., and Treisman, R. (1992). Casein Kinase-Ii Phosphorylation Increases the Rate of Serum Response Factor-Binding Site Exchange. Embo Journal 11, 97-105.

Martinez-Castilla, L.P., and Alvarez-Buylla, E.R. (2004). Adaptive evolution in the Arabidopsis MADS-box gene family inferred from its complete resolved phylogeny (vol 100, pg 13407, 2003). Proceedings of the National Academy of Sciences of the United States of America 101, 1110-1110.

Mayer, U., and Jurgens, G. (1998). Pattern formation in plant embryogenesis: A reassessment. Seminars in Cell & Developmental Biology 9, 187-193.

Mayer, U., Buttner, G., and Jurgens, G. (1993). Apical-Basal Pattern-Formation in the Arabidopsis Embryo - Studies on the Role of the Gnom Gene. Development 117, 149-162.

McCormick, S. (1993). Male Gametophyte Development. Plant Cell 5, 1265-1275.

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

Mena, M., Mandel, M.A., Lerner, D.R., Yanofsky, M.F., and Schmidt, R.J. (1995). A characterization of the MADS-box gene family in maize. Plant Journal 8, 845-854.

Michaels, S.D., Ditta, G., Gustafson-Brown, C., Pelaz, S., Yanofsky, M., and Amasino, R.M. (2003). AGL24 acts as a promoter of flowering in Arabidopsis and is positively regulated by vernalization. Plant Journal 33, 867-874.

Muller, J., Hart, C.M., Francis, N.J., Vargas, M.L., Sengupta, A., Wild, B., Miller, E.L., O'Connor, M.B., Kingston, R.E., and Simon, J.A. (2002). Histone methyltransferase activity of a Drosophila polycomb group repressor complex. Cell 111, 197-208.

Nam, J., dePamphilis, C.W., Ma, H., and Nei, M. (2003). Antiquity and evolution of the MADS-box gene family controlling flower development in plants. Molecular Biology and Evolution 20, 1435-1447.

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

Nickle, T. C. and D. W. Meinke (1998). "A cytokinesis-defective mutant of Arabidopsis (cyt1) characterized by embryonic lethality, incomplete cell walls, and excessive callose accumulation." Plant Journal 15(3): 321-332.

Ostergaard, L., and Yanofsky, M.F. (2004). Establishing gene function by mutagenesis in Arabidopsis thaliana. Plant Journal 39, 682-696.

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., Angenent, G.C., and Colombo, L. (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.

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., Gustafson-Brown, C., Kohalmi, S.E., Crosby, W.L., and Yanofsky, M.F. (2001). APETALA1 and SEPALLATA3 interact to promote flower development. Plant Journal 26, 385-394.

Pien, S., and Grossniklaus, U. (2007). Polycomb group and trithorax group proteins in Arabidopsis. Biochimica Et Biophysica Acta-Gene Structure and Expression 1769, 375-382.

Pinyopich, A., Ditta, G.S., Savidge, B., Liljegren, S.J., Baumann, E., Wisman, E., and Yanofsky, M.F. (2003). Assessing the redundancy of MADS-box genes during carpel and ovule development. Nature 424, 85-88.

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., Rounsley, S.D., Schmidt, R.J., and Yanofsky, M.F. (1995). Molecular Evolution of Flower Development - Diversification of the Plant Mads-Box Regulatory Gene Family. Genetics 140, 345-356.

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

Reyes, J.C., and Grossniklaus, U. (2003). Diverse functions of Polycomb group proteins during plant development. Seminars in Cell & Developmental Biology 14, 77-84.

Richard A. Jefferson, t. A. K., and Michael W. Bevan (1987). "GUS fusions: B-glucuronidase as a sensitive and versatile gene fusion marker and higher plants." EMBO journal 6(13): 3901 - 3907.

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.

Sambrook J., R. D. W. (2001). Molecular Cloning A LABORATORY MANUAL. New York, COLD SPRING HARBOR LABORATORY PRESS.

Schonrock, N., R. Bouveret, et al. (2006). "Polycomb-group proteins repress the floral activator AGL19 in the FLC-independent vernalization pathway." Genes & Development 20(12): 1667-1678.

Schubert, D., O. Clarenz, et al. (2005). "Epigenetic control of plant development by Polycomb-group proteins." Current Opinion in Plant Biology 8(5): 553-561.

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. Plant Cell 9, 355-365.

Simon, K.J., Grueneberg, D.A., and Gilman, M. (1997). Protein and DNA contact surfaces that mediate the selective action of the Phox1 homeodomain at the c-fos serum response element. Molecular and Cellular Biology 17, 6653-6662.

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

Spillane, C., C. MacDougall, et al. (2000). "Interaction of the Arabidopsis Polycomb group proteins FIE and MEA mediates their common phenotypes." Current Biology 10(23): 1535-1538.

Theissen, G., and Becker, A. (2004). Gymnosperm orthologues of class B floral homeotic genes and their impact on understanding flower origin. Critical Reviews in Plant Sciences 23, 129-148.

Treisman, R. (1992). The Serum Response Element. Trends in Biochemical Sciences 17, 423-426.

Treisman, R., and Ammerer, G. (1992). The SRF and MCM1 transcription factors. Curr Opin Genet Dev 2, 221-226.

Vergara-Silva, F., Martinez-Castilla, L., and Alvarez-Buylla, E.R. (2000). MADS-box genes: Development and evolution of plant body plans. Journal of Phycology 36, 803-812.

Vinkenoog, R., C. Bushell, et al. (2003). "Genomic imprinting and endosperm development in flowering plants." Molecular Biotechnology 25(2): 149-184.

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.

www.signal.salk.edu/. "SALK institute." from www.signal.salk.edu/.

Wang, X. X. and L. G. Ma (2007). "Polycomb-group (Pc-G) proteins control seed development in Arabidopsis thaliana L." Journal of Integrative Plant Biology 49(1): 52-59.

Wynne, J., and Treisman, R. (1992). Srf and Mcm1 Have Related but Distinct DNA-Binding Specificities. Nucleic Acids Research 20, 3297-3303.

Xi, H.K., and Kersh, G.J. (2002). Elk-1 phosphorylation dependent induction of Egr-1 in T cells in response to agonist and partial agonist TCR ligands. Faseb Journal 16, A701-A701.

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

Yadegari, R., T. Kinoshita, et al. (2000). "Mutations in the FIE and MEA genes that encode interacting polycomb proteins cause parent-of-origin effects on seed development by distinct mechanisms." Plant Cell 12(12): 2367-2381.

Yang, Y.Z., and Jack, T. (2004). Defining subdomains of the K domain important for protein-protein interactions of plant MADS proteins. Plant Molecular Biology 55, 45-59.

Yang, Y.Z., Fanning, L., and Jack, T. (2003). The K domain mediates heterodimerization of the Arabidopsis floral organ identity proteins, APETALA3 and PISTILLATA. Plant Journal 33, 47-59.

Zhian, S. B. a. S. (2002). "An Improved Clearing method for GUS assay in Arabidopsis endosperm and seeds." Plant Molecular Biology Reporter 20: 107 - 114.

Appendix 1 AGL40 & AGL62 DNA sequences and maps of vectors used

747bp *AGL40* (At4g36590) DNA sequence used in this project is shown below. The sequence was obtained from NCBI home page (http://www.ncbi.nlm.nih.gov/) (NCBI reference ID number 829811) and used to align the sequencing product and confirm that the sequenced *AGL40* T-DNA fragment had no random PCR mutation.

AGL40 (At4g36590 747bp)

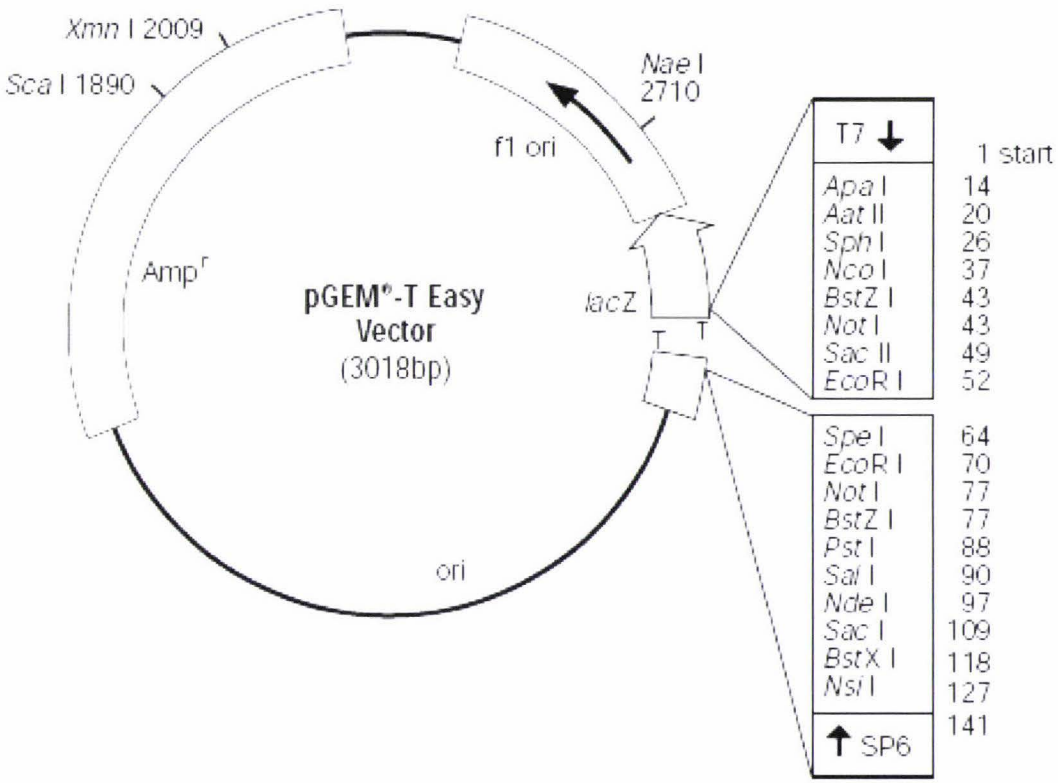
ATGGTGAGAAGTACCAAAGGTCGTCAGAAAATAGAGATGAAAAAAATGGAAAACG
AAAGCAACCTTCAGGTTACTTTCTCAAAAAGAAGATTCGGTCTTTTCAAAAAAGC
TAGTGAACTTTGCACATTAAGTGGTGCAGAGATTCTGTTGATTGTGTTCTCTCCT
GGTGGGAAAGTGTTTTCTTTTGGCCATCCAAGTGTTCAAGAACTCATTCATCGCT
TTTCGAATCCTAACCATAATTCTGCCATTGTCCATCATCAGAACAACAATCTCCA
ACTTGTTGAAACCCGTCCGGATAGAAATATCCAATATCTCAACAATATACTCACT
GAGGTGCTGGCAAACCAGGAAAAGGAGAAACAGAAGAGAATGGTTTTGGACCTAT
TGAAAGAATCCAGAGAACAAGTAGGAAACTGGTATGAAAAAGATGTGAAAGATCT
CGACATGAATGAAACCAACCAGCTGATATCTGCTCTTCAAGATGTGAAAAAGAAA
CTGGTAAGAGAAATGTCTCAATATTCTCAAGTAAATGTTTCGCAGAATTACTTTG
GTCAAAGTTCTGGCGTGATTGGTGGTGGTAATGTTGGCATTGATCTTTTTGATCA
AAGAAGAAATGCATTCAACTATAATCCAAACATGGTGTTTCCCAATCATACACCA
CCAATGTTTGGATACAACAATGATGGAGTTCTCGTTCCGATATCCAACATGAACT
ACATGTCAAGTTACAACTTCAACCAGAGCTAG

900bp *AGL62* (At5g60440) DNA sequence used in this project is shown below. The sequence was obtained from NCBI home page (http://www.ncbi.nlm.nih.gov/) (NCBI reference ID number 836165) and used to align the sequencing product and confirm that the sequenced *AGL62* T-DNA fragment had no random PCR mutation.

***AGL62* (At5g60440 900bp)**

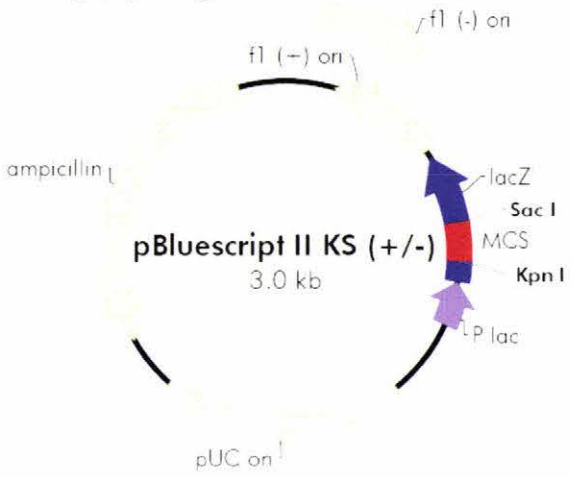
ATGGTGAAAAAAAGCAAAGGTCGTCAAAAAATAGAGATGGTCAAAATGAAAAATG
AAAGTAACCTTCAAGTTACTTTTTCAAAAAGAAGATCTGGACTTTTCAAAAAGGC
TAGTGAGCTTTGCACACTTTGTGGTGCAGAAGTGGCCATAGTTGTGTTCTCACCT
GGTCGAAAAGTCTTTTCTTTTGGTCATCCAAATGTTGATTCTGTAATTGATCGAT
TCATAAACAATAACCCTCTACCTCCTCACCAACACAACAACATGCAACTTAGAGA
AACTCGTCGAAATTCGATTGTTCAGGATCTGAATAATCATCTTACTCAGGTGTTG
AGTCAATTAGAAACAGAGAAAAAGAAGTACGACGAGTTAAAGAAAATAAGAGAAA
AGACAAAAGCCCTTGGGAATTGGTGGGAAGATCCCGTTGAGGAACTTGCGTTATC
TCAACTCGAGGGGTTCAAAGGTAATCTTGAAAATTTGAAGAAAGTAGTTACAGTC
GAAGCTTCCAGATTTTTTCAGGCAAATGTTCCAAACTTCTATGTGGGAAGTTCTA
GTAATAATGCTGCTTTTGGGATTGATGATGGTAGTCATATCAACCCTGATATGGA
TCTCTTTAGCCAAAGAAGAATGATGGACATAAATGCCTTCAACTACAACCAGAAC
CAAATTCACCCTAATCATGCATTACCACCCTTTGGAAACAATGCTTATGGTATTA
ATGAAGGGTTTGTTCCAGAATACAATGTGAACTTCAGACCAGAGTATAACCCAAA
CCAAAACCAAATCCAAAACCAAAATCAAGTTCAAATCCAAATCCAAAACCAGAGT
TTTAAGAGAGAAAACATCTCTGAATATGAACATCATCATGGTTATCCTCCCCAGT
CTAGATCTGATTACTATTAA

pGEM T-EASY vector used in the project was purchased from Promega (www.promega.com) and pBSKS vector was purchased from Stratagene (www.stratagene.com) respectively for cloning of both *AGL40* T-DNA fragment and *AGL62* T-DNA fragment.

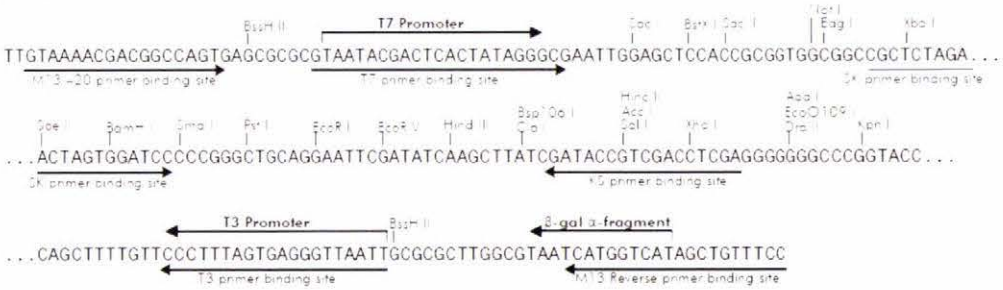


www.promega.com

pBluescript® II KS (+/–) Phagemids



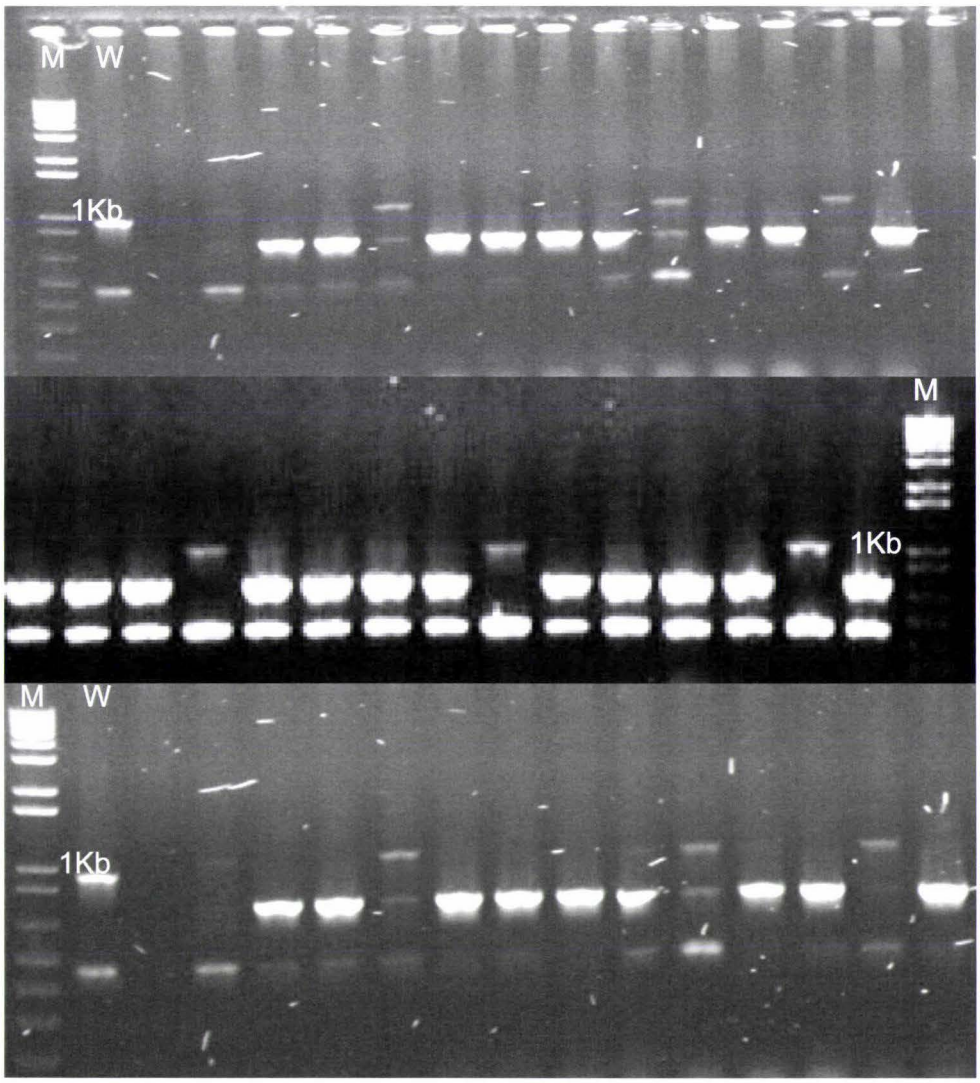
pBluescript II KS (+/–) Multiple Cloning Site Region
(sequence shown 598–826)

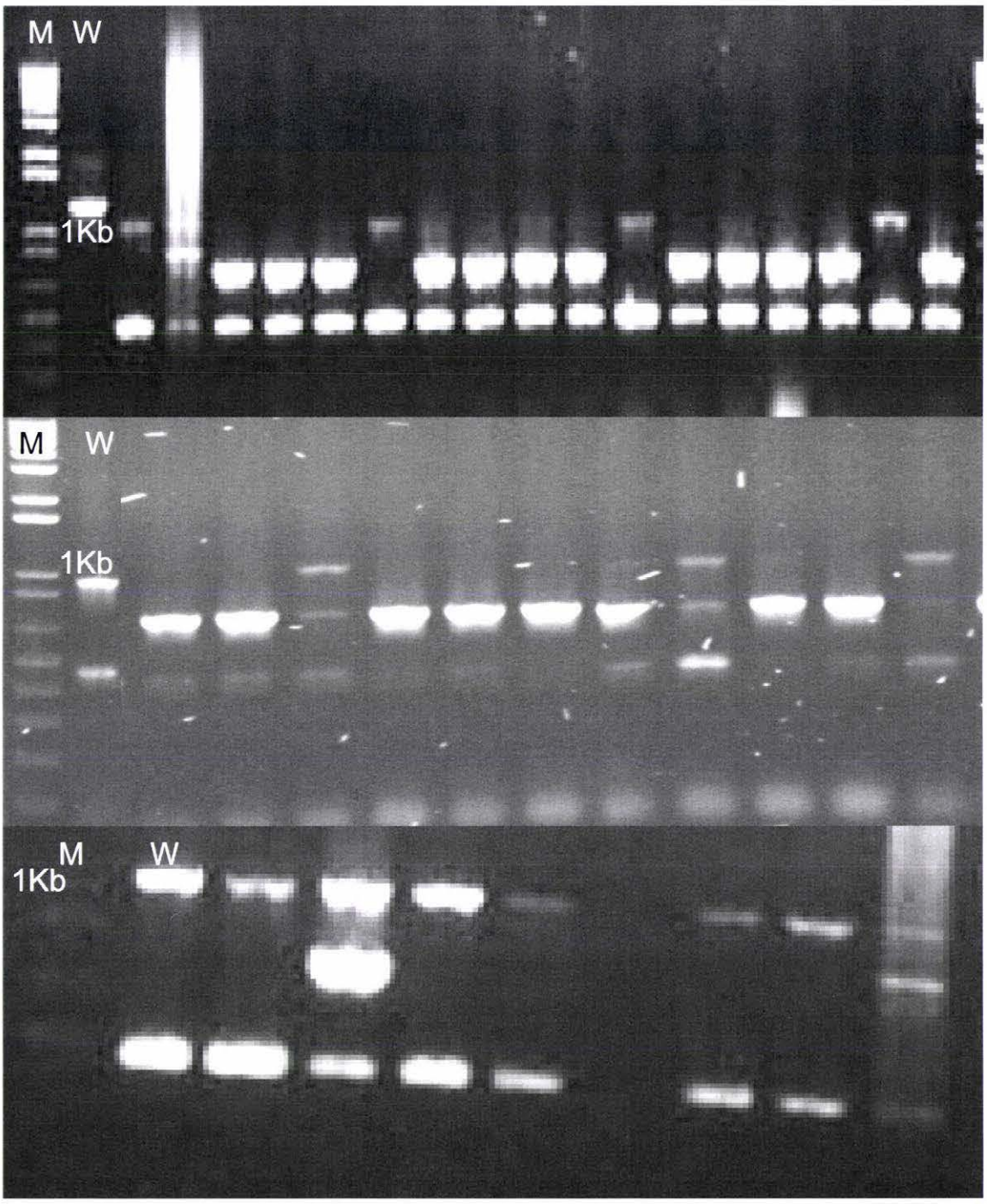


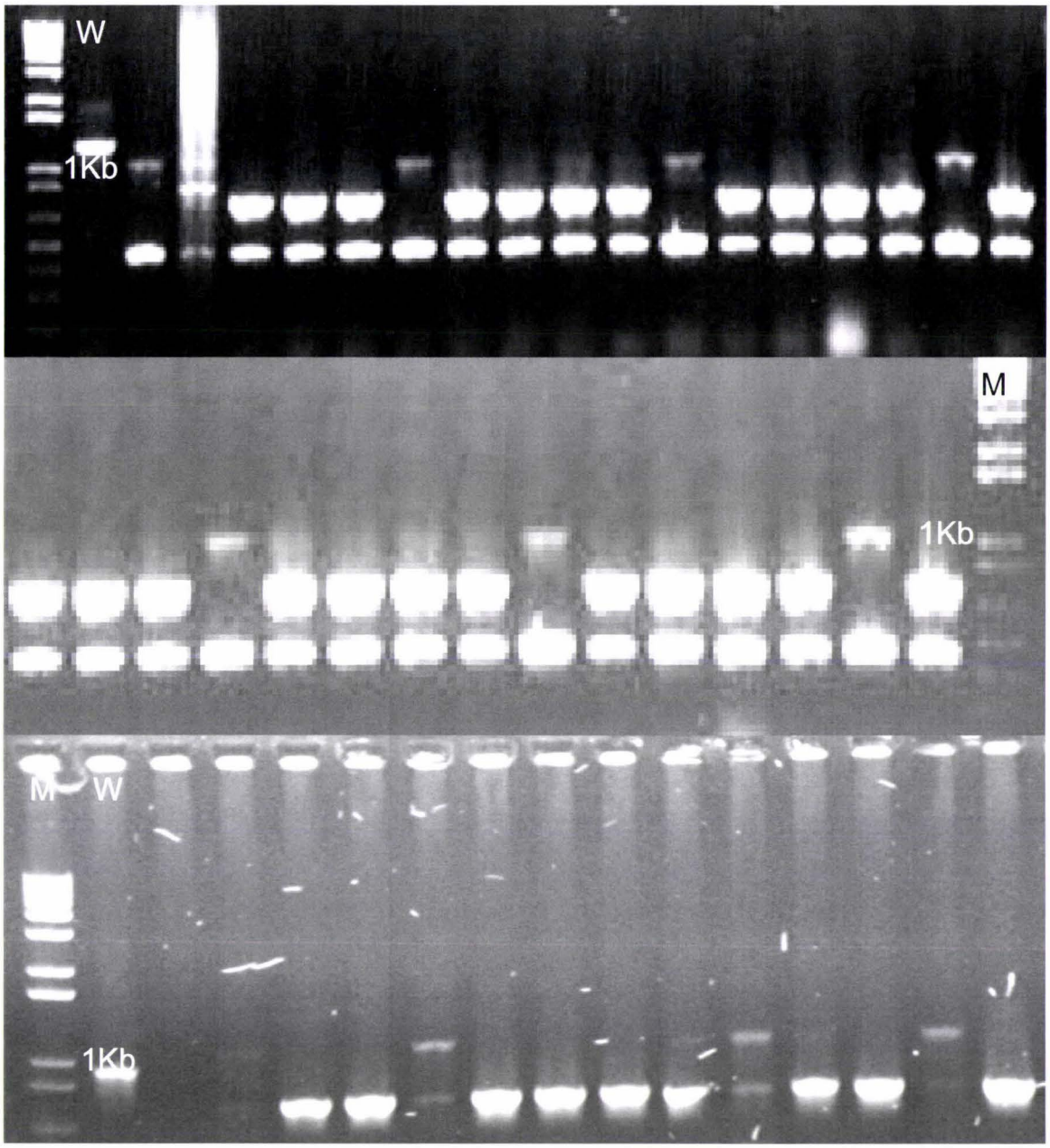
www.stratagene.com

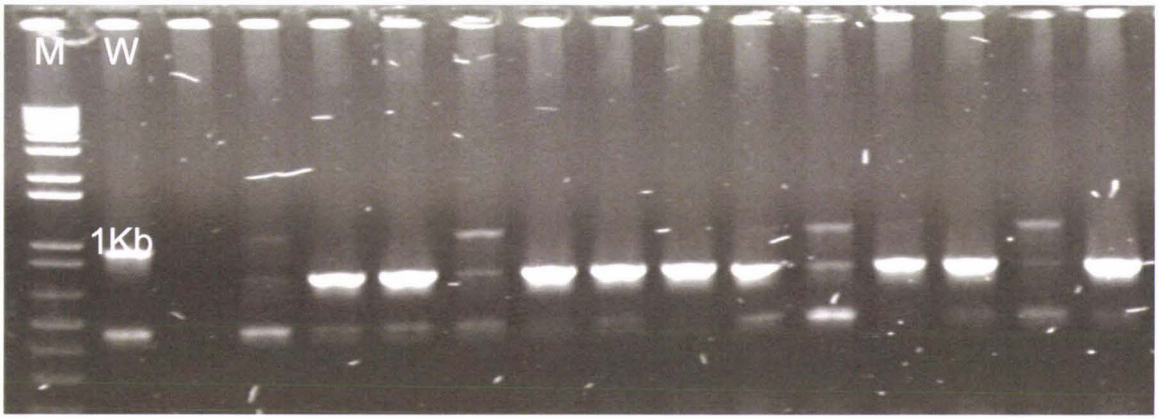
Appendix2.1 Heterozygous PCR genotyping gel photos

151 PCR based genotyping reactions were set up to identify the T-DNA insertion line for *AGL62* using 02AGL62reverse and LBa1 T-DNA primers. The PCR products were run on 1% agarose gel with 1kb plus DNA marker (M) and WT control (W). Combination of 02AGL62 and LBa1 primers was shown to amplify roughly 750bp band only if the line contained a T-DNA fragment. WT as well as some T-DNA containing lines showed 1kb band when run on 1% agarose TBE gel which corresponds to a primer match elsewhere in the genome.



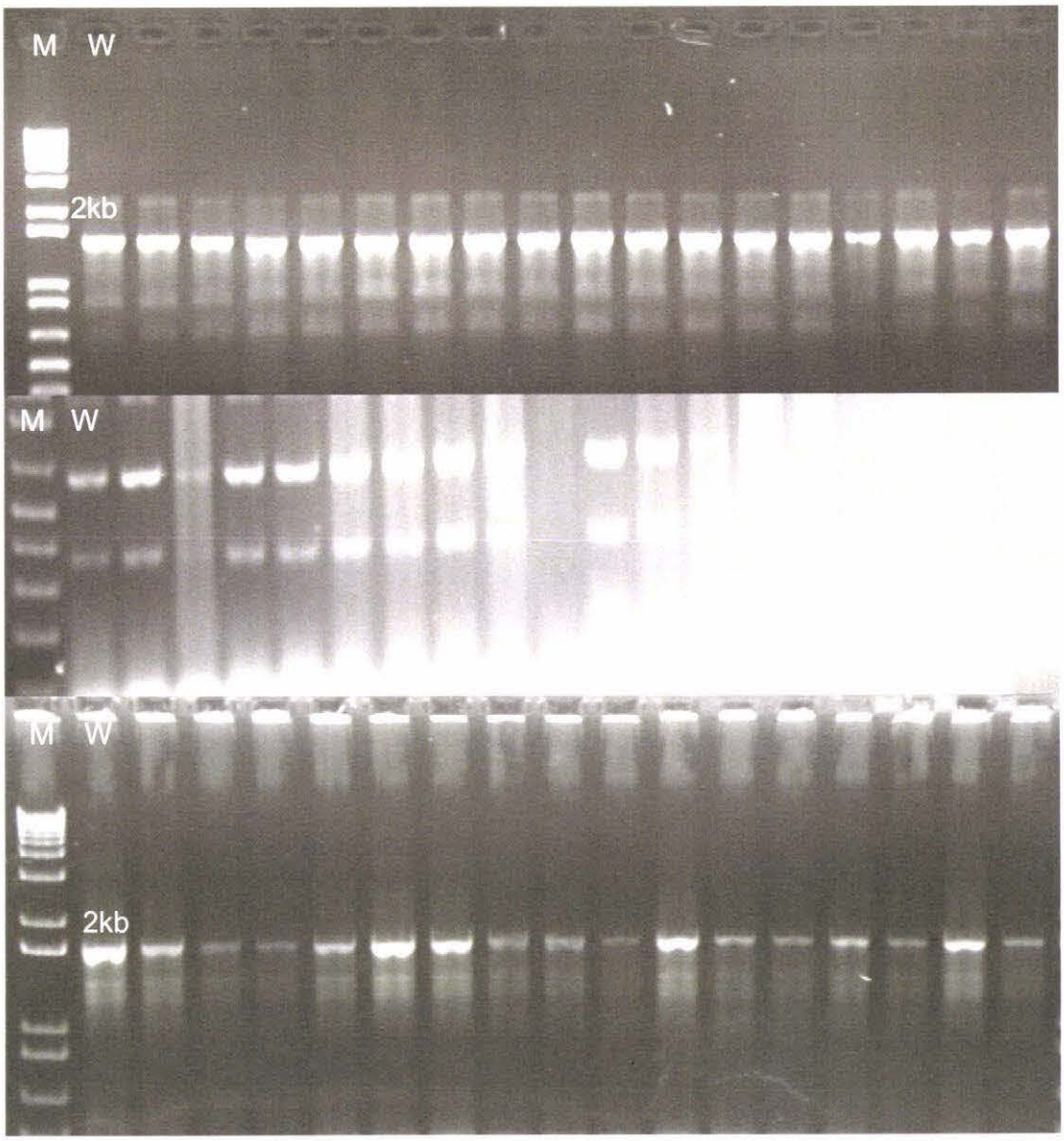

M W
1Kb
M W
1Kb
M W
1Kb

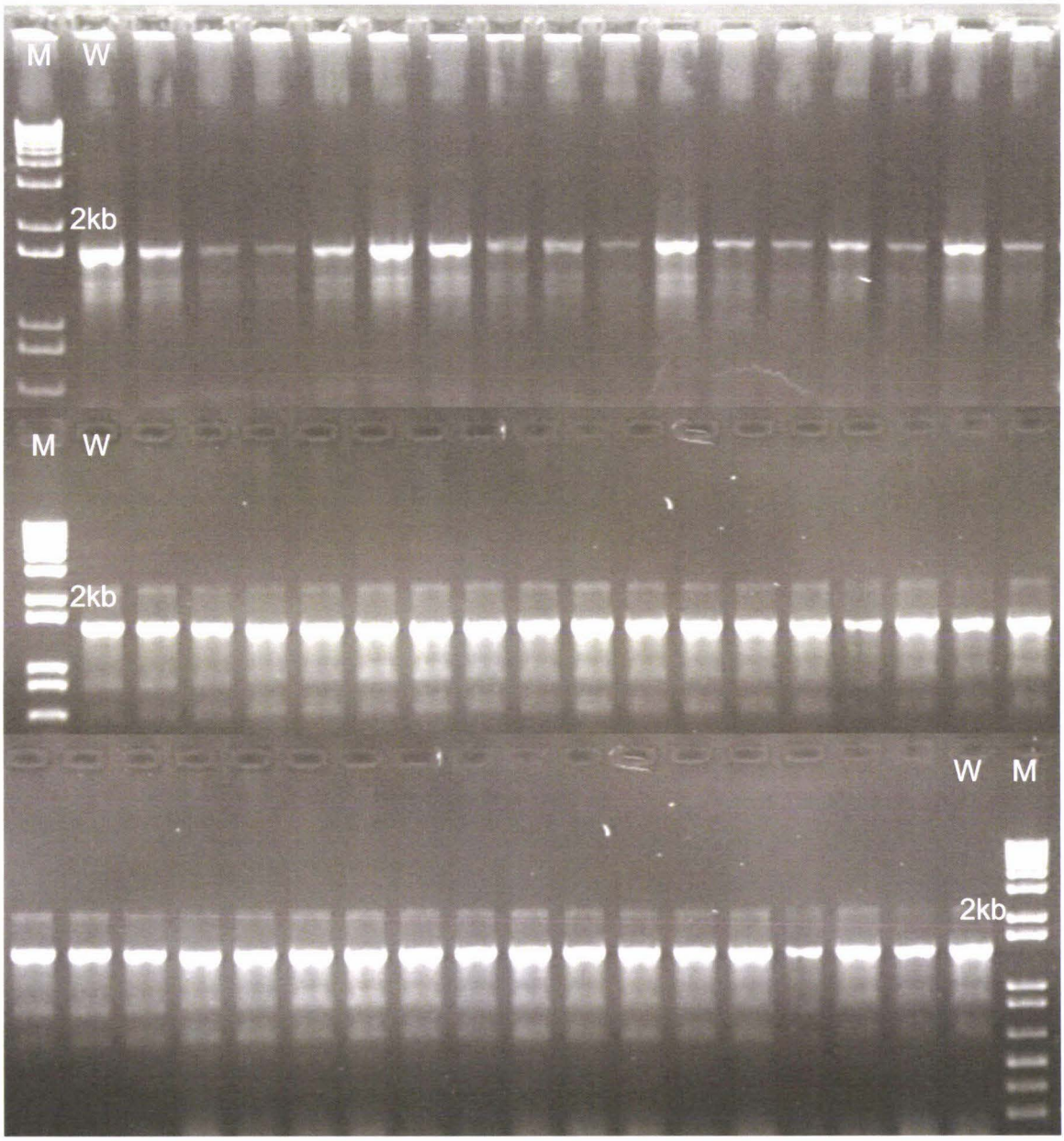

W
1Kb
M
1Kb
M
W
1Kb

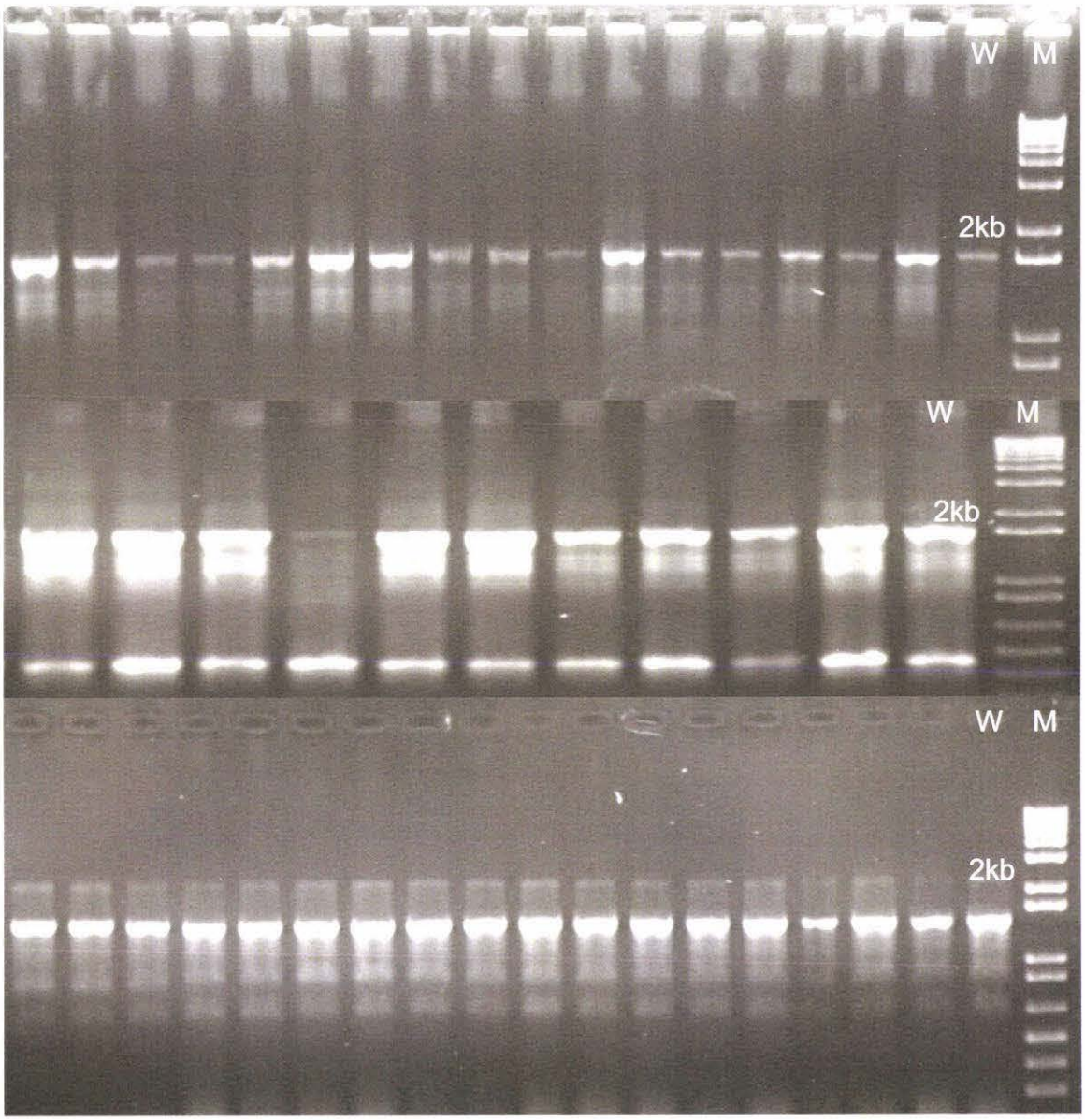

M
W
1Kb

Appendix2.2 Homozygous PCR genotyping gel photos

151 PCR based genotyping reactions were set up to identify the homozygous T-DNA insertion line for *AGL62* using 01AGL62forward and 06AGL62reverse as gene specific primers. The PCR products were run on 1% agarose gel with 1kb plus DNA marker (M) and WT control (W). Combination of 01AGL62 and 06AGL62 primers was shown to amplify roughly 1.5kb band in WT when run on a gel. All the T-DNA insertion lines gave exactly the same banding pattern as WT indicating no homozygous T-DNA insertion line was present.




M W
2kb
M W
2kb
W M
2kb


W
M
2kb
W
M
2kb
W
M
2kb