2007

IDENTIFICATION AND CHARACTERIZATION OF PROTEINS THAT INTERACT WITH AGAMOUS-LIKE 15 (AGL15), A MADS-DOMAIN TRANSCRIPTION FACTOR THAT PREFERENTIALLY ACCUMULATES IN THE PLANT EMBRYO

Kristine Hill
University of Kentucky, khill0@uky.edu

ABSTRACT OF DISSERTATION

Kristine Hill

The Graduate School
University of Kentucky
2007

IDENTIFICATION AND CHARACTERIZATION OF PROTEINS THAT INTERACT WITH AGAMOUS-LIKE 15 (AGL15), A MADS-DOMAIN TRANSCRIPTION FACTOR THAT PREFERENTIALLY ACCUMULATES IN THE PLANT EMBRYO

ABSTRACT OF DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Agriculture at the University of Kentucky

By

Kristine Hill

Director: Dr. Sharyn E. Perry, Associate Professor of Plant and Soil Sciences

Lexington, Kentucky

2007

ABSTRACT OF DISSERTATION

IDENTIFICATION AND CHARACTERIZATION OF PROTEINS THAT INTERACT WITH AGAMOUS-LIKE 15 (AGL15), A MADS-DOMAIN TRANSCRIPTION FACTOR THAT PREFERENTIALLY ACCUMULATES IN THE PLANT EMBRYO

AGAMOUS-Like 15 (*AGL15*) encodes a MADS-domain transcription factor that is preferentially expressed in the plant embryo, and may function as a regulator in embryonic developmental programs. A number of direct downstream targets of AGL15 have been identified, and while some of these target genes are induced in response to AGL15, others are repressed. Additionally, direct target genes have been analyzed that exhibit strong association with AGL15 *in vivo*, yet *in vitro*, AGL15 binds only weakly. Taken together these data suggest that AGL15 may form heterodimers, or ternary complexes with other proteins, thus modulating the specificity and function of AGL15 *in planta*. Yeast two-hybrid screens were undertaken to identify novel proteins able to interact with AGL15, and a number of interesting and potentially biologically important AGL15-interacting partners are reported here. These include members of a histone deacetylase complex, a COLD SHOCK DOMAIN (CSD)-containing protein, a K-homology domain/CCCH type zinc finger containing protein, a bZIP transcription factor, a homeobox-leucine zipper protein, a LATERAL ORGAN BOUNDARIES (LOB) domain containing protein, and an Agenet domain containing protein. Interactions between AGL15 and other MADS domain factors that are expressed in embryonic tissue, including SEPALLATA 3 (SEP3) have also been indentified. The regions of AGL15 that mediate interactions with the aforementioned proteins were mapped, and the capacity of these proteins to interact with other plant MADS-domain proteins tested.

It is reported herein that AGL15 interacts with members of the SWI-INDEPENDENT 3/HISTONE DEACETYLASE (SIN3/HDAC) complex, and that AGL15 target genes are also responsive to an AGL15 interacting protein that is also a member of this complex, SIN3 ASSOCIATED POLYPEPTIDE OF 18 KD (SAP18). AGL15 can repress transcription *in vivo,* and a region essential to this repressive function contains an LxLxL motif that is conserved among putative orthologs of AGL15. What is more, the aforementioned motif mediates the association of AGL15 with SAP18 in yeast two-hybrid assays, thus providing a possible mechanism for explaining how role AGL15 regulates gene expression via recruitment of a histone deacetylase complex.

Key words: AGL15, MADS-domain transcription factor, Embryogenesis, Protein-protein interactions, Transcriptional repression

Krristine Hill

October 2007

IDENTIFICATION AND CHARACTERIZATION OF PROTEINS THAT INTERACT WITH AGAMOUS-LIKE 15 (AGL15), A MADS-DOMAIN TRANSCRIPTION FACTOR THAT PREFERENTIALLY ACCUMULATES IN THE PLANT EMBRYO

By
Kristine Hill

Dr. Sharyn E. Perry
(Director of Dissertation)

Dr. Arthur G. Hunt
(Director of Graduate Studies)

September 2007

RULES FOR THE USE OF DISSERTATIONS

Unpublished dissertations submitted for the Doctor's degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with the permission of the author, and with the usual scholarly acknowledgments.

Extensive copying or publication of the dissertation in whole or in part also requires the consent of the Dean of the Graduate School of the University of Kentucky.

A library that borrows this dissertation for use by its patrons is expected to secure the signature of each user.

Name	Date

DISSERTATION

Kristine Hill

The Graduate School
University of Kentucky
2007

IDENTIFICATION AND CHARACTERIZATION OF PROTEINS THAT INTERACT WITH AGAMOUS-LIKE 15 (AGL15), A MADS-DOMAIN TRANSCRIPTION FACTOR THAT PREFERENTIALLY ACCUMULATES IN THE PLANT EMBRYO

DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Agriculture at the University of Kentucky

By

Kristine Hill

Lexington, Kentucky

Director: Dr. Sharyn E. Perry, Associate Professor of Plant and Soil Sciences
Lexington, Kentucky

2007

DEDICATION

I dedicate this dissertation to the memory of my grandmother, Dorothy Patricia Edwards (1925-2004)

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my mentor Dr. Sharyn Perry for her continuous support, advice and guidance over the years. I thank Drs. Arthur G. Hunt, Robert L. Houtz, L. Edward Demoll, for serving on my graduate committee, and Dr. Brian Rymond for agreeing to be my outside examiner.

I would like to acknowledge all members of the Perry lab, past and present for their advice, friendship, and collaboration. I would particularly like to express my gratitude to former lab members Weining Tang, Cong Zhu, Huai Wang and Ellen Harding who patiently answered questions and explained procedures when I was a new student. Yumei Zheng is a newer addition to the Perry Lab, who has generously supplied additional "ChIP" population for my experiments, and participated in invaluable discussions. Kate Mitchell joined the lab early this year, and has since provided valuable technical assistance especially with additional yeast two-hybrid experiments. I specifically acknowledge Huai Wang's contribution to the AGL15 transient repression studies, Weining Tang's contribution to the cloning of *GmAGL15*, and Cong Zhu for providing the microarray expression data from which two of "my" AGL15 direct target genes were first identified.

It was my time spent as an undergraduate in the lab of Dr. Hildebrand that led me to pursue a doctoral degree in plant biology, and I thank Dr. Hildebrand for advising me early on in my career and allowing me to gain valuable experience. I would like to extend my gratitude to all members of the department for creating an environment where the sharing of resources, expertise, and ideas is common place. My participation in, and interactions with the members of the UK Seed group has also greatly enhanced my academic journey here at the University of Kentucky.

I thank Dr. Michael Goodin (University of Kentucky) for his time and expertise in the bimolecular fluorescence complementation (BiFC) studies, the Goodin lab for donation of *N. benthamiana* plants, and Dr. Stanton Gelvin (University of Purdue) for providing the pSAT vectors. I also thank Dr. Gilberto Sachetto—Martins (University of Rio de Janeiro, Brazil) who generously provided the anti-GRP2 antibody used in co-immunoprecipitation

studies and Dr. Michael Thomashow (Michigan State University) who provided the p*CBF2: GUS* transgenic seed.

I thank my family, especially my mum and my grandparents for believing in me. Last but not least, I thank Ryan for his love and support, his patience, and for simply being there.

Table of Contents

List of Figures

List of Tables

Prologue

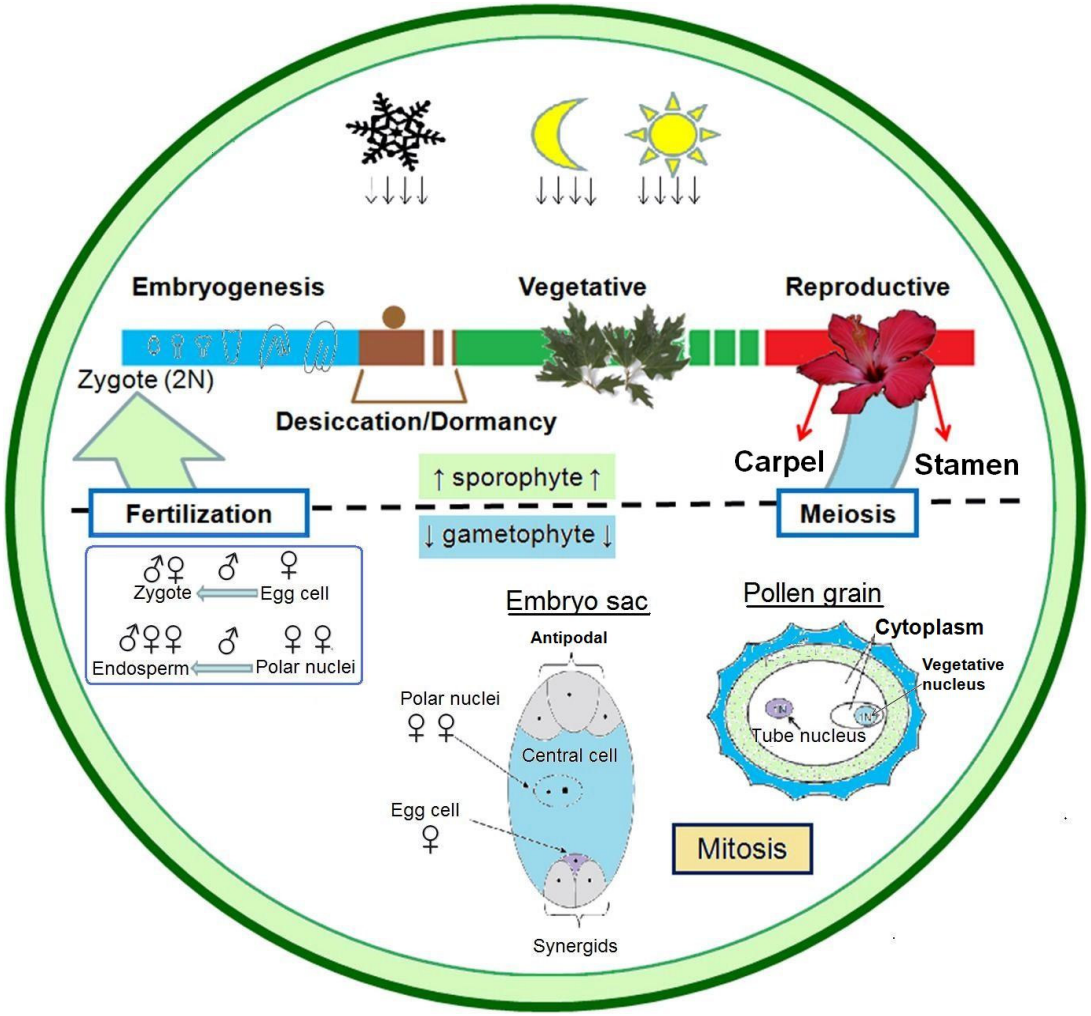
Research comprising this dissertation primarily concerns understanding how the MADS-domain factor, AGAMOUS-like 15 (AGL15), functions to regulate the expression of downstream target genes during embryogenesis, via its association with other factors. The purpose of this literature review is to introduce the reader to the morphology, physiology, and molecular biology of plant embryogenesis and to emphasize the importance of understanding the processes governing both zygotic and somatic embryogenesis. This chapter begins by providing an overview of the plant life-cycle, and presents other important developmental events, such as germination, flowering and gametogenesis. This review endeavors to provide a broad, but by no means comprehensive picture of key events in the life of the plant, and to highlight some choice research. The bias is towards providing examples involving MADS-domain proteins, because they are most relevant to this dissertation. The MADS-box family is a fascinating group of genes that have been monopolized, evolutionarily speaking, by plant life and which play fundamental roles in the developmental programs of these organisms. Flowering and floral organ identity comprises the larger portion of the current literature pertaining to MADS-domain proteins in plants. Not only must flowering pathways be repressed during the embryonic phase, but many of the mechanisms assigned to MADS-domain proteins in floral transition or identity provide insight into how a MADS-domain factor might regulate gene expression during the embryonic phase. Mechanisms of gene regulation and epigenetic memory are also discussed because they are relevant to how AGL15 might function in conjunction with other proteins to regulate the expression of specific downstream target genes. The chapter will conclude with a brief introduction to, and rational behind the research undertaken in the subsequent chapters.

1.1 Plant developmental programs

1.1.1 Overview of the life cycle of flowering plants

Greater than 1 billion years have past since the last shared progenitor of plants and animals lived, and although many basic cellular functions are highly conserved even between plants and animals, fundamental differences exist between these two kingdoms. The presence of genetically active haploid (gametophyte) and diploid (sporophyte) phases of the life cycle (alternation of generations), the absence of a germline established during embryogenesis, and a double fertilization event, which produces both an embryo and a nutritive tissue (endosperm), are unique features of flowering plants. In higher plants the sporophytic (diploid) phase consists of a series of major transitions from the embryo, through the vegetative state and eventually to the reproductive state (Figure 1.1). The sporophytic phase ends with meiosis and leads in to the reduced gametophytic (haploid) phase which produces the pollen (male) and the embryo sac (female).

Figure 1.1 Overview of the life cycle of the flowering plant



The life cycle of flowering plants includes a gametophytic and sporophytic phase. Double fertilization produces a nutritive tissue (endosperm) and an embryo, which marks the beginning of the sporophytic phase. In higher plants the sporophytic phase consists of a series of major transitions from the embryo, through the vegetative state, to the reproductive state. The sporophytic phase ends with meiosis producing the reduced gametophytic phase that produces the pollen and the embryo sac.

1.1.1.1 Sporophytic (diploid) phase

The haploid phase ends with double fertilization, whereby the fusion of the egg and one sperm cell leads to the formation of a zygote and the fusion of the central nuclei with the other sperm nucleus leads to a triploid endosperm (Figure 1.1) that serves as a source of nutrition for the developing embryo. The sporophytic generation of the angiosperm can be divided into three distinct developmental phases; embryo development (discussed in further detail in 1.1.2), vegetative development, and floral development (which gives rise to the gametes). Postembryonic plant development undergoes phase changes regulated by both intrinsic and extrinsic factors (Poethig, 1990, 2003). Transitions between developmental phases follow the same order within each life cycle. However, unlike their higher animal counterparts the plants' progression to the next stage has far greater plasticity. The plasticity of plant development allows greater integration of environmental cues (periods of cold, length of day) and developmental timing. Plants maintain a reservoir of stem cells in their shoot and root apical meristems throughout their life spans (reviewed by Sharma, *et al.,* 2003). Stem cells in the shoot apical meristem (SAM) are the progenitors of all cells that make up stems, leaves, branches, and flowers. The root apical meristem (RAM) is the source of all of the cells of the primary and lateral root system. The maintenance of gene expression patterns and the transition from one pattern to another are central to plant development and are reset at beginning of each new life cycle. Epigenetic control is one means by which developmental transitions are regulated (for review, see Berger and Gaudin, 2003 Henderson and Dean, 2004, Boss *et al.,* 2004). The concept of "resetting" describes how epigenetic memory is wiped clean with each generation and is discussed further in 1.2.3

1.1.1.2 The floral transition

Flowering produces the floral organs. The stamen, comprised of filament and anther, are structures that produce the pollen grain, which houses the male gametophyte. The carpel is comprised of the stigma, style, and ovary, which houses the female gametophyte and is the site of fertilization. The length of the vegetative phase and the transition to a reproductive mode (flowering) is influenced by environmental cues that trigger the switch. For example many plants require vernalization, a period of cold,

before they are able to flower. The hormone gibberellic acid (GA), photoperiod, vernalization, and the genes of the autonomous pathway regulate the expression of the flowering time genes, and these flowering time genes affect the expression of the meristem identity genes, which in turn regulate the organ identity genes and ultimately the development of floral organs. Many excellent reviews have been written on this subject, including Kaufmann *et al.,* 2005, whose main focus is on higher-order MADS-domain factor complexes and highlights how many genes involved in flowering and organ identity are MADS-domain transcription factors (discussed further in Chapter 1.3).

Winter temperatures promote flowering by silencing the MADS-domain transcription factor, *FLOWERING LOCUS C* (*FLC*), in a quantitative manner, thus permitting flowering following an extended period of cold (reviewed by Henderson and Dean, 2004). *FLC* repression must be 'remembered' through mitotic proliferation until flowering occurs. Post-flowering, either during the formation of the gametes or during embryo development, genes involved in a "resetting" pathway appear to be necessary to reset the expression states of the floral genes. Perturbation of their function results in ectopic expression of floral genes and premature flowering (Moon *et al*., 2003a). *CONSTANS* (*CO*) is a B-box transcription factor that functions as a promoter of flowering. *CO* mRNA expression is subject to the circadian clock and two groups of photoreceptors, phytochromes and cryptochromes have antagonistic effects on protein stability (reviewed by Henderson and Dean, 2004). Down stream targets of GA signaling include the floral timing gene, *AGAMOUS-LIKE 20/SUPRESSOR OF CONSTANS 1* (*AGL20/SOC1*; Moon *et al.,* 2003b), and the floral meristem identity gene *LEAFY* (*LFY*; Blazquez and Weigel, 2000), floral homeotic genes *APETALA 3* (*AP3*), *PISTILLATA* (*PI*), and *AGAMOUS* (*AG*; Yu *et al.,* 2004). AP3, PI, and AG are targets of transcriptional repression by the DELLA protein, *REPRESSOR OF ga1-3* (*RGA*), and GA likely promotes the expression of floral homeotic genes by antagonizing the effects of DELLA proteins (Yu *et al.,* 2004). The hormone GA promotes flowering by up regulation of the floral promoter genes *SOC1/AGL20* (Moon *et al.,* 2003b) and *LFY* (Blazquez and Weigel, 2000). Plants ectopically expressing *AGL15* flower late (Fernandez *et al.,* 2000). AGL15 binds and induces the expression of *AtGA2ox6*, which encodes for an enzyme that converts bioactive GA into inactive GA and is consistent with AGL15 being a repressor of flowering (Wang *et al.,* 2002, 2004).

1.1.1.3 Gametophytic (haploid) phase

In haploid organisms a recessive deleterious allele that would be masked in diploid organisms is subjected to natural selection. The existence of the multi-cellular haploid (gametophyte) allows for the "haploid sufficiency test" a process describing how certain mutations not compatible with multi-cellular life can be eliminated, thus reducing the genetic load. In the basal green plants such as the liverworts and mosses the haploid phase is dominant, free living, and expresses the majority of the genome. In contrast the angiosperms life cycle is completely reversed and the haploid phase usually consists of 2–7 cells that depend on the diploid parent for nutrient acquisition. However, pollen does express a large number of genes (Pina *et al.,* 2005).

Male gametophyte

The male gametophyte develops within the pollen grain. In bi-cellular pollen there are two haploid nuclei (the germ nucleus and tube nucleus) contained within the exine of the pollen grain. In tricellular pollen the generative cell divides to produce two sperm cells inside the vegetative cell. Among genes preferentially expressed during pollen development are a number of type I and non-classical MIKC* MADS-box genes (discussed in Chapter 1.3), which may play unique roles in pollen development (Pina *et al.,* 2005). Many genes essential to male gametophytic viability have been identified. These include *SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASES 1* and *2* (*SERK1* and *SERK2*), which are essential for tapetum development and microspore maturation (Colcombet *et al.,* 2005). Calcium levels and calmodulin are known to be important for pollen function, and the gene *NO POLLEN GERMINATION1* (NPG1) is a pollen-specific calmodulin binding protein that is essential for pollen germination (Golovkin and Reddy, 2003). A conditional mutant, *pop1* (for *defective pollen–pistil interactions*), is male-sterile mutant only under low humidity, thus allowing the desired phenotype to be selected for under the non-permissive condition (Preuss *et al.,* 1993). A more comprehensive review of gametophyte lethal mutants can be found in a review written by Wilson *et al.,* 2004. Upon pollen germination, a tube is formed that elongates dramatically through female tissues to reach and fertilize the ovule. Pollen tube biology is complex, and beyond the scope of this review, and for a more comprehensive review of this subject the reader is referred to recent reviews written by Boavida *et al.,* 2005a, 2005b or McCormick, 2004. However, to achieve this explosive growth the pollen grain

accumulates and stores large amounts of both protein and RNA. Thus most mRNAs are pre-synthesized before pollen maturation and the long held assumption was that pollen mRNAs were long lived. Ylstra and McCormick, 1999, tested this experimentally and confirmed that the mRNA half-life was very long for ten out of twelve mRNAs. However, two mRNAs, one of which was *GLYCINE-RICH PROTEIN 2* (*GRP2*; a MADS-interacting protein described in Chapter 3) were relatively short-lived in pollen, and in the case of *GRP2* the half life was considerably shorter than in somatic cells (Ylstra and McCormick, 1999).

Female gametophyte

The female gametophyte develops within the ovule. Although variations exist, female gametophyte development is typified by a haploid megaspore that undergoes three rounds of mitosis, without cellularization, to produce an eight-nucleate structure. In *Arabidopsis*, cellularization results in a seven-celled gametophyte containing three antipodal cells at the chalazal pole, one egg cell and two synergid cells at the micropylar pole, and a central cell. The largest of these cells is the central cell, which inherits two nuclei referred to as the polar nuclei (for review see Boavida *et al.,* 2005a). Many genes essential to female gametophytic development have been reported including the MADS-box gene, *AGAMOUS-LIKE 80* (*AGL80*), which has no effect on male gametogenesis (Portereiko *et al.,* 2006). A comprehensive review of gametogenesis and gametophyte lethal mutants can be found in a review written by Wilson *et al.,* 2004.

1.1.1.4 Double fertilization

Male and female gametophytes and tissues interact to produce a viable embryo. Double fertilization is a process whereby a pair of sperm cells are delivered by the pollen tube (male gametophyte), which elongates towards an embryo sac (female gametophyte) enclosing an egg and a central cell (for a recent review see Boavida *et al.,* 2005b).

Fertilization independent

Many plants, including citrus and dandelion, are known to produce seeds asexually through a process called apomixis. In apomixis diploid maternal cells develop into embryos without going through meiosis and fertilization, resulting in seeds that are genetic clones of the mother. In *Arabidopsis*, a plant that does not normally exhibit

apomixis, mutations in the genes *FERTILIZATION INDEPENDENT* (*FIE)*, *FERTILIZATION INDEPENDENT SEED 2* (*FIS2*) or *MEDEA (MEA)* disrupt the normal dependence of seed development on fertilization (reviewed by Ma, 1999). Apomixis allows a genotype that is adapted to a particular environment to spread quickly through seeds without losses due to heterogeneity. In agriculture, apomixis provides a means to propagate an elite crop variety clonally via seeds, which are more easily stored and transported than plants.

Imprinting

Some autosomal genes are expressed only from their maternally or paternally inherited copy. These genes are called imprinted genes and play important roles in growth and development. In plants, imprinted-gene expression seems to be confined to the endosperm (reviewed by Arnaud and Feil, 2006). Because the endosperm does not transmit genetic or epigenetic information to the next generation the epigenetic status needs not be subject to a developmental cycle of erasure and reestablishment as observed in mammalian imprinting. *DEMETER* (*DME*), a DNA glycosylase, is believed to be one of the main players in endosperm-specific imprinting that arises through specific demethylation in the female gametophyte (Choi *et al*., 2004). *DME*, whose expression is predominantly in the central cell, induces maternal expression of *MEA*, which is maintained in the endosperm after fertilization (Choi *et al*., 2004). *MEA* is an essential gene that confers maternal control over seed development and acts specifically to suppress the maternally supplied MADS-box gene *PHERES 1* (*PHE1*), but not the paternal equivalent (Köhler *et al*., 2003, 2005). *DME* also induces the maternal activation of another *Arabidopsis* imprinted gene, the flowering time gene *FWA*, by antagonizing the action of DNA methylation (Kinoshita *et al*., 2004).

1.1.2 Zygotic and somatic embryogenesis

The human diet includes plant material and the meat and produce of animals raised on plant material. We wear plant material (cotton), we build shelters from plant material (wood), and our health often depends on compounds produced by plants. Without plants there would be no humans. The seed houses the next generation, the embryo. The seed nourishes the embryo, and protects it during dispersal and until environmental conditions are optimal for the survival of a newly emerged seedling. For agriculturally

important crops the seed is not only sold for human, pet and livestock consumption, but also serves a “store house” for the next generation. The farmer planting his or her crop is concerned with seed quantity, quality, and consistency. There is also much interest in understanding the biochemistry behind lipid metabolism in oilseeds, with a focus on improving the human diet and producing industrial oils or alternative fuel sources. It has been estimated that 70% of the human diet comes directly from seeds (Bewley and Black, 1994). Therefore, basic research that strives to understand seed development and embryogenesis offers the potential for improving agriculture and ultimately sustaining a growing population.

1.1.2.1 Zygotic Embryogenesis

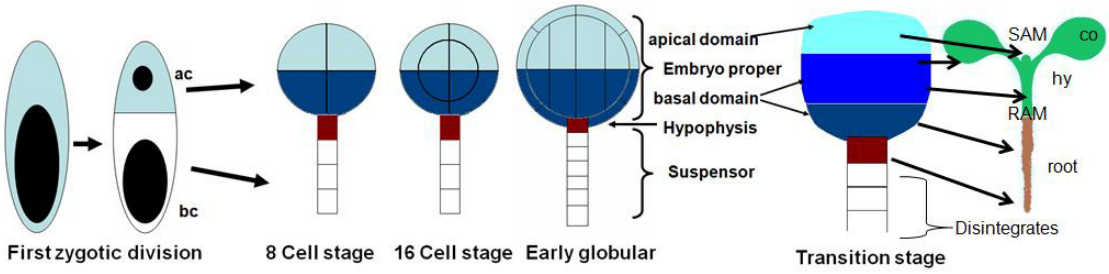
Embryogenesis, which transforms the fertilized egg cell into a multi-cellular organism, is a crucial period in the development of eukaryotes. The process of zygotic embryogenesis in higher plants is initiated by double fertilization, where by the fusion of the egg and one sperm cell leads to the formation of a zygote and the fusion of the central nuclei with the other sperm nucleus leads to a triploid endosperm (Figure 1.1). The embryo develops within an embryo sac, which is surrounded by maternal diploid tissue of the ovule. Seed development can be divided into two major phases: Embryo development (morphogenesis) and seed maturation. In eudicots the morphogenesis phase ends at the heart stage when all embryo structures have been formed. During the seed maturation phase the embryo grows and fills the seed sac, and accumulates food reserves and acquires dormancy and desiccation tolerance (Goldberg *et al*., 1994).

In *Arabidopsis*, zygotic embryogenesis begins with an asymmetric cell division that gives rise to a polar embryo having a larger basal cell and a smaller apical cell. The embryo proper develops from the apical cell, and the basal cell develops into the suspensor, which is attached to the ovule and serves as a conduit for nutrient transfer to the developing embryo (Figure 1.2). In *Arabidopsis*, the egg cell and zygote initially display a polar organization, with a large vacuole at the basal end. The zygote subsequently elongates and then divides asymmetrically to form daughter cells (apical and basal) of different sizes. The apical daughter cell undergoes two rounds of longitudinal and one round of transverse divisions to give rise to an eight-cell embryo proper. Descendants of the basal daughter cells divide transversely to form the suspensor and the hypophysis.

Unlike animal cells, plant cells do not move and positional information instead of lineage is the primary determinant of cell fate. During higher plant embryogenesis a simple body plan is established, which requires organization and coordination. The cells of the embryo need to become specified and must differentiate into cell types in an integrated manner. Recent reviews on pattern formation during plant embryogenesis include, Souter and Lindsay, 2000, Jurgens, 2001, Laux *et al.,* 2004, Willemsen and Scheres, 2004, Jenik *et al.,* 2007. In Arabidopsis four regions with different developmental fates can be recognized at the eight-cell stage:

1. The apical embryo domain that gives rise to the shoot meristem and most of the cotyledons.
2. The central embryo domain that form the hypocotyl and root as well as contributing to the cotyledons and the root meristem.
3. The hypophysis that will give rise to the distal parts of the root meristem, the quiescent center, and the stem cells of the central root cap.
4. The extra embryonic suspensor, which eventually pushes the embryo into the endosperm tissue and provides a connection to the mother tissue.

Figure 1.2 Pattern formation during *Arabidopsis* embryogenesis



Developmental fates are defined by the first asymmetric cell division, and pattern formation describes how lineages of cells, determined during early embryogenesis, are destined to produce the progenitor cells defining specific tissues and organ in the seedling. This figure was constructed based on reviews written by Souter and Lindsay, 2000, Jurgens, 2001, Laux *et al.,* 2004, Willemsen and Scheres, 2004, and Jenik *et al.,* 2007. ac, apical cell; bc, basal cell; SAM, shoot apical meristem; RAM, root apical meristem; hy, hypocotyl;, co, cotyledons

Concurrent with the transition from globular to heart stage in eudicots, two groups of cells divide periclinally causing bulges that emerge as cotyledon lobes. The apical-most suspensor cell, named the hypophysis, becomes incorporated into the embryo proper and the rest of the suspensor disintegrates, providing a source of nutrients for the embryo. In the subsequent torpedo and bent-cotyledon stages of eudicots such as *Arabidopsis*, the embryo completes its growth by elongating and enlarging.

Quiescence and Germination

During the later stages of embryogenesis, endogenous abscisic acid (ABA) levels in embryo increase, followed by the accumulation of seed storage proteins, thus enabling the seed to acquire desiccation tolerance (reviewed by Ikeda *et al.,* 2006). Plant development is interrupted during the seed maturation processes which results in a quiescent seed. Once desiccation has been achieved, the seed remains quiescent until conditions are right for completion of germination. Environmental conditions trigger germination and embryo arrest is reversed. Germination is initially characterized by cell elongation, which allows the embryo to break through the surrounding envelopes (Mansfield and Briarty, 1996, Bewley, 1997).

GA and ABA act antagonistically to regulate the germination versus maturation programs. ABA promotes maturation while GA promotes germination. *Arabidopsis ABA-insensitive* 3 (*abi3*) and maize *viviparous 1* (*vp1*) are ABA-insensitive mutants that undergo viviparous germination, and acquire no desiccation tolerance and accumulate little seed storage proteins (Ooms *et al*., 1993, Hollung, 1997, Parcy *et al*., 1994). The low expression levels of some *LATE EMBRYOGENESIS ABUNDANT* (*LEA*) genes in these mutants suggest that *ABI3*/*VP1* may be an important factor in controlling the expression of the *LEA* genes (Parcy *et al*. 1994, Ooms *et al*., 1993, Hollung, 1997, Baumbusch, 2004). The expression levels of the *LEA* genes are very low in somatic embryos (Ikeda-Iwai *et al.,* 2002), which bypass desiccation and dormancy. The *fusca 3* (*fus3), leafy cotyledons 1 and 2 (lec1* and *lec2)* mutants also exhibit premature germination (Raz *et al.,* 2001). Seeds have mechanisms to ensure germination occurs only under favorable environmental conditions for seedling growth (Bewley and Jack, 1994). All seeds must imbibe before they will complete germination. For some this is the only requirement, yet others have photoperiod, temperature, or mechanical requirements

before they will complete germination. Recent reviews include Koornneef *et al.,* 2002, Finch-Savage and Leubner-Metzger, 2006.

1.1.2.2 Somatic Embryogenesis

Unlike cells of other eukaryotes, almost all plant cells have the capacity to dedifferentiate under defined conditions. Somatic embryogenesis is a process whereby differentiated somatic cells dedifferentiate, and then form embryos that are morphologically similar to zygotic embryos, and are capable of developing into seedlings. Of the many benefits *Arabidopsis* offers plant scientists, accessibility to the embryo is not one of them. Not only is the embryo surrounded by maternal tissue, but the small seed size makes harvesting sufficient amounts of tissue for many experimental procedures a challenge. *Arabidopsis* somatic embryogenesis systems have circumvented this limitation and enabled researchers to ask biological questions pertaining to embryo development. For a comprehensive review of these systems Ideka *et al.,* 2006, is recommended.

In addition to a providing a tool for understanding embryo development, somatic embryos may be used as a convenient way to propagate large numbers of genetically identical individuals. The discovery and exploitation of embryogenic tissue cultures has led to the development of efficient procedures for plant regeneration in almost all of the agriculturally important grasses, and recovery of mature plants from protoplasts in crops such as maize, rice and sugarcane (Vasil, 1988). However, only a limited number of cells will form somatic embryos at any one time and this fraction is highly variable among plant species and cultivars. Agrobacterium-mediated transformation of cotton is highly efficient in inserting transgenes into cotton cells (Sunilkumar and Rathore, 2001), and the major obstacle lies in the ability to recover transgenic plants (Wilkins *et al*., 2000). Most cotton cultivars are recalcitrant to plant regeneration through somatic embryogenesis, and the ability to initiate somatic embryos appears to be genotype-dependent (Trolinder and Xhixian, 1989, Sakhanokho *et al.,* 2004). A similar story exists for *Agrobacterium*-mediated transgene introduction into soybean, and successful transformation has been limited to a few cultivars able to be regenerated via somatic embryogenesis (Ko *et al.,* 2004). The list goes on, but the take home message is that understanding, and ultimately improving somatic embryogenesis in agriculturally

important crop plants will expedite the recovery and propagation of transgenic lines and permit clonal propagation of desirable genotypes.

1.1.2.3 Induction of somatic embryogenesis

In a short letter published in Science a couple of years back Vogel, 2005, asked the question, what makes a plant cell become embryogenic? The answer was, and still is, that we do not really know. However, a number of genes when ectopically expressed promote the induction of somatic embryogenesis, namely *SERK1* (Schmidt *et al*., 1997, Hecht *et al*., 2001, Hu *et al.,* 2005), *BABY BOOM* (*BBM*; Boutilier *et al.,* 2002), *WUSCHEL* (*WUS*; Zuo *et al*., 2002), *LEC1* (Lotan *et al*., 1998) *LEC2* (Stone *et al*., 2001), and *AGL15* (Harding *et al.,* 2003). *EMBRYONIC FACTOR 1* (*FAC1*) is one of the first genes known to be expressed in the diploid plant and high level can be observed in putative embryogenic cells (Xu *et al.,* 2005). *SERK1* expression also seems to be concurrent with cells that develop into somatic embryos (Schmidt, *et al.,* 1997). Mutants have also been described that exhibit enhanced somatic embryogenesis. The *primordial timing* (*pt*) and *clavata* (*clv*) mutants and the *vpi/abscisic acid insensitive 3-like* (*val1/val2*) double mutant show enhanced somatic embryogenesis (Mordhorst *et al*., 1998, Suzuki *et al.,* 2007). *LEC1* expression is restricted to embryogenesis (Lotan *et al.,* 1998) and is repressed in vegetative tissue postgermination in part by *PICKLE* (*PKL*), a putative chromatin-remodeling factor (Ogas *et al*., 1999). The *pickle* (*pkl*) mutant exhibits an incomplete penetrance phenotype typified by enhanced somatic embryogenesis that is increased by inhibitors of GA biosynthesis (Ogas *et al*., 1999).

It has long been known that hormones play an important role in plant morphogenesis. Skoog and Miller's famous experiments showed that the ratio of auxin to cytokinin in the growth medium determined whether roots or shoots developed from cultured cell clusters (Skoog and Miller, 1957).

Auxin

Exogenous auxin is a requirement for most somatic embryogenesis systems (Ideka *et al.,* 2006). Auxin is usually a requirement for induction of somatic embryogenesis, but a reduction of auxin is necessary for the development of somatic embryos (Ideka *et al.,* 2006). Exactly how auxin works during somatic embryogenesis is not understood, and it

is worth remembering that most studies use the synthetic analog 2,4-dichlorophenoxyacetic acid (2,4-D) rather than the biological prominent auxin indole-3-acetic acid (IAA). Auxin induces the expression of *AGL15* (Zhu and Perry, 2005), which could account for the fact that *35S:AGL15* cultures no longer require exogenous auxin for the induction somatic embryogenesis (Harding *et al.,* 2003). In fact ectopic expression of several other genes including *BBM*, *LEC1*, *LEC2*, and *WUS*, promote somatic embryogenesis in the absence of external hormonal inducers (Boutilier *et al*., 2002; Lotan *et al*., 1998; Stone *et al*., 2001; Zuo *et al*., 2002)

Gibberellic Acid (GA)

GA levels also seem to be important to embryogenesis. In the embryo GA biosynthesis appears to be regulated by LEC2 and FUS3. The level of bioactive GAs is increased in immature seeds of *lec2* and *fus3* mutants relative to wild-type level and the GA-biosynthesis gene *AtGA3ox2*, which encodes the key enzyme that catalyzes the conversion of inactive to bioactive GAs, is ectopically activated in these mutants (Curaba *et al.,* 2004). AGL15 directly binds and induces the expression of *AtGA2ox6*, which encodes for an enzyme that converts bioactive GA into inactive GA (Wang *et al.,* 2004). Ectopic expression of *AGL15* induces somatic embryogenesis (Harding *et al.,* 2003) but somatic embryo induction is reduced in the *ga2ox6* background (Wang *et al.,* 2004).

Abscisic Acid (ABA)

In the late stages of zygotic embryogenesis, endogenous ABA levels increase, and accumulation of seed storage proteins ensues, followed by desiccation. In contrast somatic embryos do not normally undergo desiccation and development is not interrupted by a period of quiescence. The maternally produced ABA in the seed can inhibit viviparous germination in mutants that lack the growth arrest capacity after the embryo phase (Raz *et al.,* 2001). Unlike zygotic embryos, somatic embryos are not surrounded by maternal tissues. However ABA-treated somatic embryos do acquire some desiccation tolerance (Shiota *et al*., 1998). Endogenous ABA is important for the induction of secondary embryogenesis on carrot somatic embryos (Ogata *et al.,* 2005) and supplying ABA enhances the production of somatic embryos in coconut (Fernando and Gamage, 2000). ABA is traditionally described as the “stress hormone” and is the major player in mediating the adaptation of the plant to stress. Somatic embryos can be induced by stresses such as osmotic, heavy metal ion, drought, and cold (Ikeda-Iwai *et*

al., 2002, Umehara, unpublished data cited by Ikeda *et al.,* 2006). It is interesting to note that some of the genes regulated by AGL15 appear to be involved in stress responses (S. Perry and C. Zhu, unpublished data).

Brassinosteroids

AGL15 has been recovered along with another promoter of somatic embryogenesis, SERK1, from a complex also comprised of two members of the brassinosteroid signaling pathway, the main brassinosteroid receptor BRASSINOSTEROID-INSENSITIVE 1 (BRI1) and its co-receptor SERK3 (Karlova *et al.,* 2006). Brassinosteroids regulate a wide range of developmental and physiological processes, including cell elongation, cell division, stem elongation, vascular differentiation, senescence, and photomorphogenesis.

1.1.3 Genes essential to development

In embryo defective mutants a gene, essential for the completion of embryogenesis, has been disrupted. Consequently the embryo aborts and dead seeds are observed. Some dead seeds will be due to environmental stresses, however, siliques producing approximately one quarter dead seeds (or in the case of a gametophytic lethal mutation, half dead seeds) are indicative of a recessive gene mutation. *Embryo-defective* (*EMB*) mutants differ in their terminal phenotypes, extent of abnormal development, allele strength, nature of the underlying mutation, size and color of aborted seeds and embryos, efficiency of transmission through male and female gametes, capacity to produce mutant seedlings, and level of phenotypic analysis (Tzafrir *et al*., 2004). In many cases basic cellular function is protected by functional redundancy. However there exists a set of non-redundant genes that are essential for life and a drastic phenotype is observed in absence of each. In Arabidopsis there are postulated to be 500 to 750 genes encoding non-redundant functions in this minimal set of genes essential to life (McElver *et al.,* 2001, Tzafrir *et al.,* 2003). Information on more than 400 mutants with embryo-defective mutants has been deposited in the SeedGenes database (Tzafrir *et al*., 2003). One of these is the *LARIAT DEBRANCHING 1* (*DBR1*) enzyme (Wang *et al.,* 2004).

Although some of these genes are specific to embryo development, such as *LEC1* and *LEC2* (Meinke *et al.,* 1994), many are house keeping genes that also have important functions beyond embryo development, but because they are essential the plant cannot progress though the earliest stages of development without them. Tzafrir *et al*., 2004 compared 250 genes with embryo-defect mutant phenotypes and 550 genes with other mutant phenotypes and found that genes required for basal functions associated with DNA, RNA, and protein synthesis are more prevalent among the former (30% versus 9%), while transcription factors and components of signal transduction pathways are more common among the later (39% versus 15%). When compared to 550 genes with other knockout phenotypes, *EMB* genes have fewer paralogs, and are more likely to have counterparts among essential genes of yeast (*Saccharomyces cerevisiae*) and worm (*Caenorhabditis elegans*; Tzafrir *et al*., 2004). Of the 69 Arabidopsis genes with a significant match in both yeast and worm, 60% are *EMB* genes and 35% are predicted to perform a basal cellular function (Tzafrir *et al*., 2004).

1.2 Mechanisms of gene regulation

1.2.1 Chromatin structure and remodeling

Chromatin is historically categorized as one of two distinct domains, heterochromatin and euchromatin. Heterochromatin is defined as chromosome regions that remain densely stained and highly condensed throughout the cell cycle, whereas euchromatin is decondensed during interphase. Heterochromatin is often associated with telomeres and pericentric regions of the chromosomes that are rich in repetitive sequences and low in gene density. Euchromatin by contrast is gene rich and characterized by irregularly spaced nucleosomes arrays. A high-frequency of nuclease hypersensitive sites with euchromatic regions indicates accessibility and is characteristic of active transcribed genes (Grewal and Elgin, 2002)

1.2.1.1 Chromatin modifiers

The term “chromatin remodeling” is used to describe a range of biochemical processes that lead to an altered or reconfigured chromatin structure and changes its accessibility to a variety of factors involved in replication, transcription, recombination, and DNA repair (for a recent reviews see Meyer, 2001, Hsieh and Fischer, 2005). The building block of chromatin is the nucleosome, which is comprised of 146 base pairs of DNA wrapped around the histone octamer (a H2A/H2B tetramer and two H3/H4 dimers). Chromatin modifiers can be grouped into three classes: histone chaperones, ATP-dependent chromatin-remodeling enzymes, and histone modification enzymes (Reyes, 2006). One example of histone chaperones is the HIR/HIRA proteins, which interact with histones and function in nucleosome assembly and gene silencing (Prochasson *et al.,* 2005). In plants a HIRA homolog, together with *ROUGH SHEATH 2* (*RS2*) and *ASYMMETRIC LEAVES 1* (*AS1*) mediate epigenetic silencing, possibly by modulating chromatin structure (Phelps-Durr *et al.,* 2005). The MYB domain proteins RS2 and AS1 can form complexes through interaction with the DNA binding LOB-domain factor ASYMMETRIC LEAVES 2 (AS2), a predicted RNA binding protein (RIK, for RS2-Interacting KH protein), and a homologue of the chromatin-remodeling protein HIRA (Phelps-Durr *et al.,* 2005). ATP-dependent chromatin-remodeling enzymes alter

interactions between the DNA and the histone octamer, which result in destabilization of the nucleosome structure and allow general and specific transcription factors to access the DNA (Reyes, 2006). The amino-terminal tails of histones (mostly H3 and H4) can be chemically modified by a large number of histone modification enzymes that add or remove small chemical groups. Chromatin modification is achieved by phosphorylation, acetylation, methylation, ADP-ribosylation and ubiquitination of histone tails. These modifications influence the interaction of histone tails with DNA and adjacent nucleosomes, and, as histone tails have multiple target sites for acetylation, phosphorylation and methylation, the options for specific combination of these modifications are vast (for a recent reviews see Meyer, 2001, Hsieh and Fischer, 2005). These covalent modifications have been proposed to function as a 'histone code', which provides information about the transcriptional state of a gene and that can be inherited as epigenetic marks during cellular differentiation (Jenuwein and Allis, 2001, Reyes *et al*., 2002, Loidl, 2004). The histone modifications were first examined in yeast and mammals but appear to be largely conserved in plant chromatin (Grasser, 2005). Histone modifications, as well as DNA methylation are important for imprinting. Imprinted genes play important roles in growth and development. In plants imprinting is an essential epigenetic process that controls the size of seeds (Arnaud and Feil, 2006). Epigenetic inheritance is the underlying mechanism enabling plants to "remember" cold winter and flower accordingly in the spring (Henderson and Dean, 2004). Upon vernalization, the *FLC* locus is epigenetically silenced by increased methylation of lysines K9 and K27 of histone H3 (Bastow *et al.,* 2004)

1.2.1.2 Histone acetyl transferases and deacetylases

The interplay between histone acetyl transferases (HATs) and histone deacetylases (HDACs) results in a dynamic equilibrium between acetylation and deacetylation at promoters and regulatory regions that affect chromatin structure and transcription (reviewed by Reyes *et al*., 2002).

Histone acetyl transferases (HATs)

Transcriptionally active genes are enriched in hyperacetylated histones and in histones methylated at certain positions (Grasser, 2005). Acetylation occurs at lysine residues on the amino-terminal tails of the histones and neutralizes the positive charge of the histone

tails, thus decreasing their affinity for DNA and altering nucleosome conformation. Histone acetyl transferases (HATs) include the TAF130/250 subunit of the TFIID complex, and p300/CBP, which is are associated with the RNA polymerase II holoenzyme (reviewed by Struhl, 1998). GCN5, the first histone acetylase to be identified, although not essential for cell growth, is important for the expression of a subset of genes. In yeast GCN5 is found in at least two distinct multi-protein complexes, ADA and SAGA, neither of which is tightly associated with TFIID or the RNA polymerase II holoenzyme (reviewed by Struhl, 1998). Histone acetylase activity is also an intrinsic function of ACTR and SRC-1, two transcriptional co-activators that associate with a variety of nuclear receptors in a hormone-dependent manner (reviewed by Struhl, 1998).

Histone deacetylases (HDACs)

Histone deacetylases (HDACs) are enzymes that catalyze the removal of acetyl groups from the N-terminal tails of histones. However, most recombinant expressed enzymes are found to be inactive *in vitro* indicating other cofactors are required for HDAC activity (reviewed by de Ruijter *et al.,* 2003). In the *Arabidopsis* genome, 16 potentially functional HDACs have been identified, and these can be classified into three families, the RPD3/HDAC1-like histone deacetylases, the members of the SIR2-like family, and the plant-specific HD2-like HDACs (Pandey *et al*., 2002). The HDACs form complexes with other factors *in* vivo. For example, in addition to HDACs, SIN3 can sequester other enzymatic functions, including nucleosome remodeling, DNA and histone methylation (Silverstein and Ekwall, 2005). The SIN3/HDAC1 complex lacks any DNA-binding activity, so must be targeted to gene promoters by interacting with DNA-binding proteins (Silverstein and Ekwall, 2005).

1.2.2 Transcriptional regulation

Large multi-subunit, DNA-dependent RNA polymerases are used by organisms in all kingdoms to produce mRNAs. Despite their structural complexity, these multi-subunit enzymes require one or more auxiliary factors to facilitate and regulate transcription of protein encoding genes. The general transcription factors help recognize promoter DNA, position the RNA polymerases at the transcription start site and support DNA strand separation during transcription initiation. Transcription factors and promoter DNA sequences serve as targets for negative and positive cofactors that modulate mRNA

production via reversible interactions that transiently affect the function of the transcription machinery, thus conferring even greater complexity and scope for refined regulation. In eubacteria five-subunit RNA polymerase ($\alpha_2\beta\beta'\omega$) acts with one of several σ factors to form a functional holoenzyme and regulate gene expression under various environmental conditions. In prokaryotes the RNA polymerase recognize the –35 and –10 DNA regions upstream of the transcription start site. In eukaryotes, transcription initiation is decidedly more complicated. RNA polymerases I, II, and III are responsible for rRNA, mRNA and tRNA synthesis respectively. In mRNA transcription five class II general transcription factors (TFIID, TFIIB, TFIIF, TFIIE and TFIIH) assemble on the core promoter with RNA polymerase II to form a functional pre-initiation complex (reviewed by Burley and Kamada, 2002).

1.2.2.1 Transcription activation domains

Transcriptional activators contain a sequence-specific DNA binding domain and an activation domain that mediates a variety of protein-protein interactions that lead to gene activation (Minter *et al.,* 2004, Titz and Thomas *et al.,* 2006). Glutamine, asparagine and acidic residues are often found in activation domains (Titz and Thomas *et al.,* 2006). Types of activation domains include acidic, glutamine-rich, proline-rich regions, as well as a few unspecific properties like dispersed hydrophobic patches (Titz and Thomas *et al.,* 2006). Even small chemical compounds with a certain pattern of hydrophobic and hydrophilic residues are sufficient for transcriptional activation (Minter *et al.,* 2004). Proteins containing a glutamine-rich activation domain include the *Drosophila* proteins GAGA (Espinas *et al*., 2000) and BICOID (Zhu and Hanes, 2000). Proline-rich activations domains can be found in the murine proteins HOXD-4 and AP-2 (Rambaldi *et al.,* 1994). Example of proteins containing an acidic activation domain are the mammalian protein YY1 (Yang *et al.,* 1996), the plant EARLY RESPONSIVE TO DEHYDRATION 15 (ERD15; Wang and Grumet, 2004), the herpes simplex virus type 1 VP16 protein (Dalrymple *et al.,* 1985), and the yeast protein that is utilized in yeast two-hybrid assays, GAL4 (Traven *et al.,* 2006). GAL4 activates transcription by recruiting co-activators and the general transcription machinery to promoter regions through its activation domain (Traven *et al.,* 2006).

1.2.2.2 Repressor proteins

Repression may occur via simple exclusion of activators from their target sequences promoter regions (competitive binding), masking of regulation domains through dimerization with other factors (Liu *et al.,* 1999), or by actively altering the acetylation state of the chromatin (de Ruijter *et al.,* 2003). Hypoacetylation results in a decrease in the space between the nucleosome and the DNA that is wrapped around it. The widely accepted picture is that tighter wrapping of the DNA diminishes accessibility for transcription factors, leading to transcriptional repression (de Ruijter *et al.,* 2003).

LxLxL/EAR repression motif

Comparison of amino acid sequences of class II ETHYLENE RESPONSE FACTOR (ERF) repressors reveal the conservation of the sequence motif L /F DLN L /F (x)P, designated the ERF-ASSOCIATED AMPHIPHILIC REPRESSION (EAR) motif and mutational analysis demonstrated that this motif is essential for repression (Ohta *et al.,* 2001). This EAR motif was also identified in a number of zinc-finger proteins from wheat, Arabidopsis, and petunia plants, and these proteins functioned as repressors, with their repression domains mapping to regions that contained an EAR motif (Ohta *et al.,* 2001). A similar motif, designated LxLxL, can be found in AUX/IAA protein is also responsible for the *in vivo* repressive capacity of these protein (Tiwari *et al.,* 2004). A similar motif also exits in the MADS-domain transcription factor AGL15 and may account for the proteins ability to act as a repressor *in* vivo (Hill *et al.,* 2007; Chapter 2). The nature of this short motif suggests that it does not possess an inherent enzymatic function that could direct account for its repressive activity. Work described in this study explains how such a motif might confer repressive function via recruitment of the SIN3/HDAC1 complex (Hill *et al.,* 2007; Chapter 2).

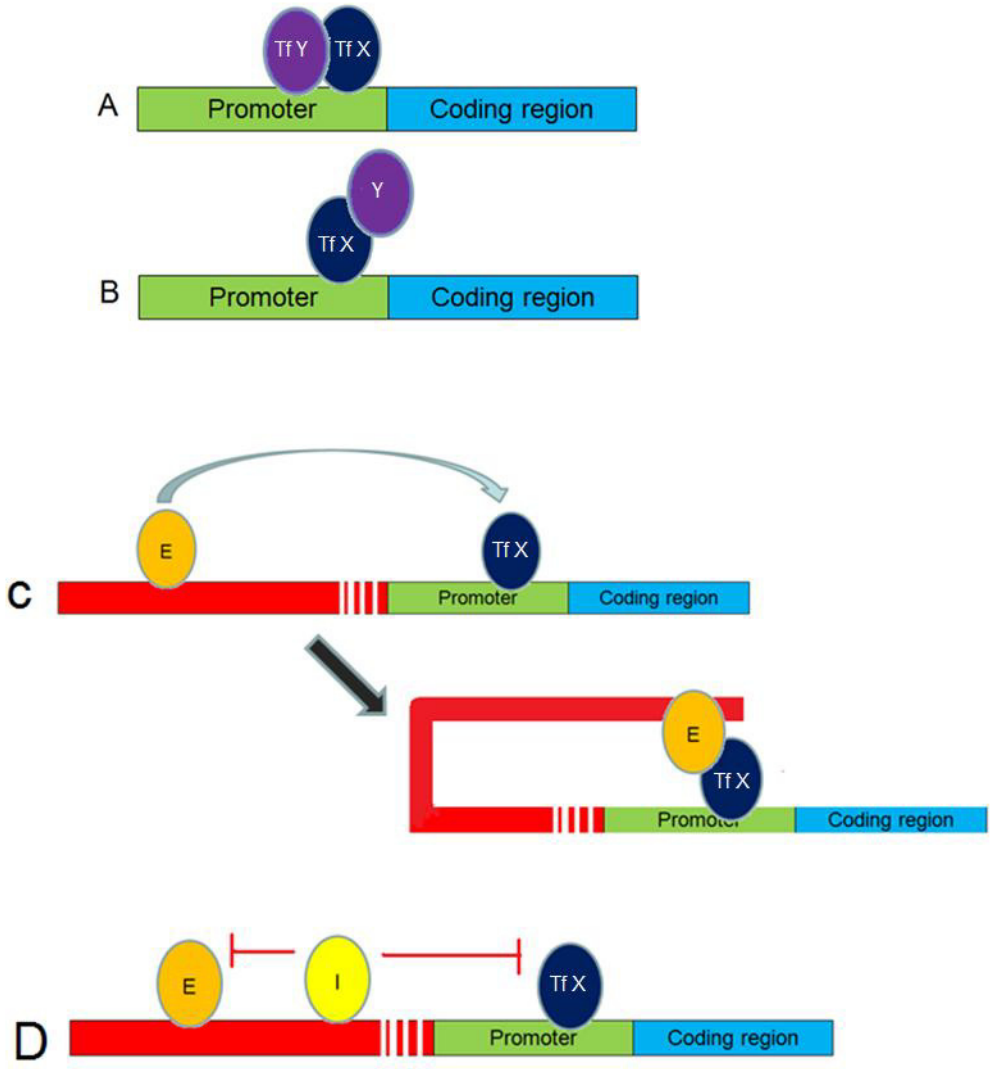
There are numerous examples in the literature of protein interactions determining whether a protein acts as a transcriptional activator or repressor. Known transcriptional activators have been shown to exert a repressive function through recruitment of the SIN3/HDAC1. The mouse homolog of the transcriptional activator Su(fu) interacts and functionally cooperates with SAP18 to repress transcription by recruiting the SAP18-mSin3 complex to promoters containing the Gli-binding element (Cheng and Bishop, 2002). The Drosophila transcription factor BICOID is converted from an activator into a

repressor by recruitment of a co-repressor to BICOID-dependent promoters (Zhu and Hanes, 2000). The mammalian zinc-finger transcription factor, YY1, contains an acidic activation domain and a glycine-rich region that is responsible for mediating the interaction of YY1 with the RPD3/HDAC1 complex and the consequential repressive activity (Yang *et al.,* 1996). Co-activators that interact with SIN3/HDAC1 members, giving the complex an activator function, have also been reported. In the literature HDACs are generally associated with the repression of gene expression, however, in plants HDA19 has been shown to interact with BnKCP, and exerts a modest transactivation of reporter genes (Gao *et al.,* 2003). HDA19 has also been shown to interact with BnSCL1, an ortholog of the *Arabidopsis* SCARECROW-LIKE 15 (SCL15) protein and likewise exert transcriptional activation of reporter genes (Gao *et al.,* 2004).

1.2.2.3 Enhancers and Insulators

Enhancer bound proteins bind to promoter bound transcription factors and enhance initiation of transcription, and they can accomplish this independently of distance and location with respect to the gene being regulated (Mongelard and Corces, 2001). Insulators interfere with interactions between enhancers and promoters and inhibit enhancer-activated transcription. In yeast, insulators delimit the boundaries of silenced chromatin at telomeres and the mating-type loci HML and HMR (Mongelard and Corces, 2001).

Figure 1.3 Simplified model depicting modes of transcriptional regulation



TF X, transcription factor X: Y, regulatory factor Y; E, enhancer; I, insulator.

A = Transcription factors X and Y bind the promoter in close proximity and interact to regulate transcription. B = Transcription factor X binds the promoter and recruits a co-activator or co-repressor, Y. C = Transcription factor X bind the promoter and interacts with an enhancer (E) bound to distant (several kilo bases) DNA elements. D = An Insulator (I) prevents interaction between a transcription factor and enhancer.

1.2.2.4 Combinatorial control

Combinatorial regulation describes how the activity of a combined set of transcription factors is greater than the additive effects of individual transcription factors. Cooperative recruitment of co-activators to form enhanceosomes depends on stereo specific alignment of DNA-binding sites such that adjacently bound factors are aligned with one another on the same face of the helix. Changing the spacing between binding sites by a half-integral multiple of the DNA helical repeat length often has much more deleterious effects than does changing the spacing by an integral multiple of the helical repeat length (Liu and Little, 1998). The concept of combinatorial control was explored by Kato *et al.,* 2004, who systematically identified combinations of transcription factors and motifs involved in multiple binding mechanisms by integrating chromatin immunoprecipitation (ChIP) data with microarray expression data and with combinatorial transcription factor (TF)-motif analysis. From this study three general features of combinatorial control pertaining to the regulation of the yeast cell cycle are apparent: waiting-activating systems, joint-phase combinations, and joint-process combinations (Kato *et al.,* 2004). A waiting-activating system is an apparatus that is already in place but awaits a signal before it activates transcription. Joint-phase combinations describe how some gene promoters are bound by one regulator that works primarily in the previous cell-cycle phase, and another that works primarily in the next cell-cycle phase. Joint-process combinations are combinations of transcription factors that allow genes to respond to two (or more) different transcriptional programs.

Repression can be described as long or short range (reviewed in Courey and Jia, 2001). In long range repression, the promoter is rendered resistant to the influence of all enhancers, even those located thousands of base pairs from the repressor binding site. Because the entire chromosomal locus is inactivated by this type of repression it is often referred to as gene silencing. Physical analyses suggest that the silenced loci are organized into a compact conformation that may be inaccessible to the transcriptional machinery or that may reduce the processivity of transcription by hindering RNA-polymerase II and associated factors (Courey and Jia, 2001). In contrast, short-range repressors function in a less broad, but more flexible manner. Rather than interfering with all transcription at a given locus, they block the function of nearby DNA-bound activators while not interfering with more distantly bound activators. The phenomenon of

long-range repression was first characterized in studies of the yeast silent mating type loci HMR and HML (Loo and Rine 1994). A large number of proteins are required for optimal silencing of the *HMR* and *HML* loci, which form a complex that appears to blocks transcription by remodeling a large domain of chromatin into a repressed, hypoacetylated, conformation (Loo and Rine 1994). A well studied example of long range repression is GROUCHO-mediated repression in animals (reviewed in Courey and Jia, 2001). GROUCHO protein does not bind directly to DNA, but is recruited by protein–protein interactions with a variety of DNA-bound repressors. Studies of GROUCHO-mediated repression suggest that large nucleoprotein complexes, analogous to the enhanceosome, may also mediate repression (Courey and Jia, 2001). Evidence supporting this concept of "repressosomes" comes from studies of DORSAL, a GROUCHO-dependent repressor that can also function as an activator of transcription (Dubnicoff *et al*. 1997). The ability of DORSAL to discriminate between targets that it should activate and targets that it should repress depends on the context of the DORSAL binding sites in any given target gene (Dubnicoff *et al*. 1997). Recruitment of GROUCHO to the template by protein-protein interactions is required for the conversion of DORSAL from an activator to a repressor (Dubnicoff *et al*. 1997).

Both long- and short-range corepressors may function through histone deacetylation. However, long-range corepressors might have the ability to spread along the template recruiting histone deacetylases and/or other chromatin modifying activities to a large domain, in a cooperative fashion, whereas short-range repressors may lack the capacity to spread. Alternatively, the differences between long- and short-range corepressors could relate to the inherent properties of different histone deacetylases (Courey and Jia, 2001). Indeed, the long-range co-repressor, GROUCHO binds only class I histone deacetylases, whereas the short-range repressor, CTBP appears to bind both class I and class II histone deacetylases (Bertos *et al*. 2001).

1.2.4 Post transcriptional regulation

Gene-expression studies in eukaryotes mostly measure steady-state mRNA levels as an indicator of the synthetic rates of transcript production. However, this approach fails to differentiate between mRNA stabilities and transcriptional activation. This is not trivial

considering that mRNA stability is an important player in post-transcriptional control and variations in transcript stability can be tissue specific (Ylstra and McCormick, 1999).

Transcription by RNA-polymerase II proceeds through multiple stages: promoter recruitment, initiation, elongation and termination. Much interest has been paid to the recruitment and initiation, but the elongation stage was once considered a mere extension of the initiated transcript. However, it is now apparent that it is a dynamic and highly regulated step in the production of mRNA (reviewed by Grasser, 2005). Transcript elongation plays a central role in coordinating transcription and various co-transcriptional RNA processing steps such as 5' capping, splicing and polyadenylation (Proudfoot *et al.,* 2002). DNA template is packaged into chromatin in the cell nucleus, and the elongating polymerase has to cope with the repressive effects of chromatin on efficient mRNA synthesis. A group of elongation factors has been identified that can assist RNAPII transcribing through chromatin (Grasser, 2005). ATP-dependent chromatin remodeling machines, SWI/SNF, CHD1 and ISW1, can alter the state of the nucleosomes in the path of the polymerase to allow productive elongation (Grasser, 2005).

A picture is emerging relating the importance of transcription factors in RNA-processing mechanism (reviewed by Kornblihtt *et al.,* 2004). Promoter structure appears to be important for alternative splicing because changing the promoter perturbs normal RNA processing, suggesting factors that regulate alternative splicing could be acting through promoter bound factors (reviewed by Proudfoot *et al.,* 2002, Kornblihtt *et al.,* 2004). An old observation is that expression of recombinant cDNAs transfected in mammalian cells is far less efficient compared to corresponding intron-containing constructs (reviewed by Kornblihtt *et al.,* 2004). Higher levels of *AGL15* transcription are observed in transgenic plants carrying a form of AGL15 containing the first three introns compared to those transformed with the cDNA version (Fernandez *et al.,* 2000). This indicates that factors controlling intron removal are important for normal levels of transcription. The presence of an intron, or simply a 5'-splice site immediately downstream from a promoter greatly enhances transcription, in mammalian and yeast genes (Furger *et al*. 2002). A member of the SIN3/HDAC1 complex, which is involved in transcriptional regulation, has also been identified in the multi-protein exon junction complex (Tange *et al.,* 2006).

RNA-binding proteins organize nascent RNA transcripts into groups, and mediate their progression along the chain of splicing, nuclear export, stability and translation (reviewed by Keene, 2007). This apparent coordination, especially at the level of mRNA stability and translation, formed the basis of the post-transcriptional 'RNA-operon' theory (Keene, 2007). The RNA-operon theory describes how *trans*-acting factors regulate multiple mRNAs in a combinatorial fashion along a coordinated pathway of RNA processing. This allows cells to respond with quickly to environmental cues.

1.3 MADS domain transcription factors

1.3.1 The MADS family

MADS-domain proteins comprise a large family of regulatory factors that have been identified in all major eukaryotic kingdoms and are involved in a diverse array of biological functions (for review see Becker and Theissen, 2003, Messenguy and Dubois, 2003, Kaufmann *et al.,* 2005). MADS is an acronym derived from the four founding member of this family; MCM1 from yeast, AGAMOUS and DEFICIENS from plants, and SERUM RESPONSE FACTOR from humans.

The only common denominator among all MADS-box genes is the ~180 base pair region, which encodes the DNA-binding MADS domain. Based on the MADS-domain sequence, two distinct types of MADS box genes are found in animal, fungi, and plant and are classified as SRF-like (type I) and MEF2-like (type II). This suggests that at least one gene-duplication event occurred before the divergence of plant and animals. MADS-domain proteins bind DNA as homo- or heterodimers and recognize AT-rich consensus sequences known as CArG boxes, which contain a highly conserved 10 bp core. SRF-like (type I) proteins bind the sequence $CC(A/T)_6GG$ and MEF2-like (type II) proteins the $CTA(A/T)_4TAG$ sequence (Messenguy and Dubois, 2003). Type I and type II proteins can be further classified into subfamilies on the basis of shared sequence similarity between regions outside of the highly conserved MADS domain. Type I proteins are comprised of SAM (fungi)/SRF (animals) and SRF-like (plants). Type II proteins are MEF2 (fungi and animals) and MIKC type in plants (reviewed by Messenguy and Dubois, 2003).

Animals and fungi contain only a few MADS-box genes, whereas plant genomes boast considerably more MADS-box genes. Indeed 107 MADS-box genes can be found in the humble *Arabidopsis* genome (Parenicová *et al.,* 2003). Plant type I MADS-box genes can be further grouped into Mα, Mβ, and Mγ, and type II MADS-box genes into the classical $MIKC^C$ and MIKC*(Mς) groups (Becker & Theissen, 2003, Parenicová *et al.,* 2003). In Arabidopsis there are 39 $MIKC^C$, 6 MIKC*(Mς), 25 Mα, 20 Mβ, 16 Mγ genes, and only one (AGL33) that could not be assigned a group (Parenicová *et al.,* 2003). The $MIKC^C$ members are the most studied and in eudicots 13 different paralogous MIKC-type

MADS-box gene subfamilies have been defined based on phylogeny: AG-, AGL2(SEP)-, AGL6- AGL12-, AGL15-, AGL17-, DEF-, FLC-, GGM13- (B-sister), GLO-,SQUA-, STMADS11-, and TM3-like genes (Becker and Theissen, 2003). Eight out of the thirteen clades have been identified in angiosperms and gymnosperms, and a further three can be found in both monocots and eudicots (Becker and Theissen, 2003). Monocots and eudicots diverged 160 -200 million years ago. Only for two clades, AGL15- (comprised of *AGL15* and *AGL18*) and FLC-like genes, have members not yet found outside of eudicots (Becker and Theissen, 2003).

A typical plant type II MADS-domain protein consists of four domains: MIKC. The MADS (M) domain itself is comprised of 55-60 highly conserved amino acids, and it is this domain that associates with the DNA (for review see Reichmann and Meyerowitz, 1997). The intervening (I) domain is less conserved, but forms part of the minimal DNA-binding domain (Reichmann *et al.,* 1996). The K-domain, whose name comes from its inferred structural similarity to the coil-coil motif of Keratin, is implicated in mediating protein-protein interactions. The K-domain of AGAMOUS (AG) has been shown to interact with four AGAMOUS-LIKE (AGL) proteins: AGL2, AGL4, AGL6, and AGL9 (Fan *et al.,* 1997). The I- and K-domains define the functional specificities of APETALA3 (AP3) and PISTILLATA (PI), whereas the MADS and I domain define those of APETALA1 (AP1) and AG (Reichmann *et al.*, 1996). The K-domain is predicted to form three hydrophobic α-helices, and it is the first two helices that appear to be critical for the strength of APETALA3/PISTILLATA (AP3/PI) dimerization (Yang *et al.,* 2003a). In *Antirrhium* DEFICIENS (DEF), GLOBOSA (GLO) and SQUAMOSA (SQUA) can form ternary complexes via the C-termini, and the DNA binding affinity of these complexes differs from that of the individual dimers (Egea-Cortines *et al*., 1999). In some cases the inclusion of the C-domain, although alone unable to mediate dimerization, enhances interactions involving the K-domain (Fan *et al.,* 1997). The carboxyl (C) domain is the most divergent, and in some cases has been shown to be a transactivation domain (Lim *et al*., 2000, Moon *et al*., 1999, Cho *et al.,* 1999, Ng and Yanofsky, 2001, Pelaz *et al.,* 2001, Honma and Goto, 2000, Honma and Goto, 2001).

MADS domain transcription factors may function as both transcriptional activators and repressors, depending on interaction partners present. For instance, ectopic expression of *SEPALLATA 3* (*SEP3*) has been reported to induce *AG* expression outside of the

floral context (Castillejo *et al.*, 2005). However AP1 and SEP3 are also able to interact with the transcription co-repressors LEUNIG (LUG) and SEUSS (SEU) (Sridhar *et al.,* 2006), which prevent ectopic *AG* transcription (Franks *et al.*, 2002; Liu and Meyerowitz, 1995). Neither *LUG* nor *SEU* are predicted to encode a recognizable DNA-binding motif, but SEU has been shown to associate *in vivo* with an *AG cis*-regulatory region containing a putative CArG motif, perhaps through binding to DNA-bound SEP3 (Sridhar *et al.,* 2006).

1.3.2 The role of MADS-box genes in plant development

Although the basic mechanisms of developmental pattern formation evolved independently in plants and animals, there are many similarities in the overall logic. For example in *Drosophila* segmental identity is established by the spatially specific transcriptional activation of an overlapping series of master regulatory genes, the *HOX* homeobox genes. In contrast to animals, homeotic genes in plants do not code for homeodomain-containing proteins, and most often encode for MADS domain proteins. The MADS box genes are not homologous to the homeobox genes. MADS box and homeobox family members are found in plants and in animals, and each family traces its origin to before the last common ancestor (Meyerowitz, 2002).

In plants, MADS-domain proteins are the central players in many developmental processes, including control of flowering-time, homeotic regulation of floral organogenesis, fruit development, and seed pigmentation (reviewed by Parenicová *et al*., 2003). With as many plant MADS-box genes as there are it is not that surprising that so much genetic redundancy is observed and that higher order mutant are often required before a phenotype is observed. One of the key driving forces behind evolution is gene duplication. One copy serves its original function, freeing the second to take on new roles. A high degree of partial or full redundancy is encountered within the MADS-box gene family, especially in recently duplicated clades (Rijpkema *et al.,* 2007).

1.3.2.1 Fruit development

The *SEPALLATA* (*SEP*) genes play an important role in ovule development. A reduction in SEP activity leads to the loss of normal ovule development (Favaro *et al.,* 2003).

MADS-box genes involved in ovule and fruit development include the classical type II MIKC^C group members *FRUITFULL* (*FUL*), *SEEDSTICK* (*STK*), *SHATTERPROOF 1* (*SHP1*) and *SHATTERPROOF 2* (*SHP2*). Ovule development is lost in *stk/shp1/shp2* triple mutants and ectopic expression of *STK* or *SHP* is sufficient to induce the transformation of sepals into carpeloid organs bearing ovules (Favaro *et al.,* 2003). The fruit is a complex structure unique to flowering plants. The fruit mediates the maturation and functions in seed dispersal. Following fertilization, fruits undergo a dramatic enlargement that is accompanied by differentiation of numerous distinct cell types. Seed dispersal in plants such as Arabidopsis occurs by a process called fruit dehiscence, or pod shatter. The terminal step of fruit development in Arabidopsis involves valve separation from the replum, allowing seed dispersal. The *fruitful* (*ful-1*) mutation abolishes elongation of the silique after fertilization (Gu *et al.,* 1998). *FUL* is also required for fruit valve differentiation, and is a negative regulator of *SHP* expression (Gu *et al.,* 1998, Ferrándiz, *et al.,* 2000). *SHP1* and *SHP2* redundantly control dehiscence zone differentiation and promote the lignification of adjacent cells (Liljegren *et al.,* 2000). AG interacts with SEP3 in yeast two-hybrid assays (de Folter *et al.,* 2005), and likely act together in the same complex to regulate the expression of *SHP2*, a MADS-box gene expressed in the fourth whorl of the flower. *AG* has been shown to be activated by ectopic expression of *SEP3* (Castillejo *et al*., 2005). *SHP2* has been described as a target gene of AG (Savidge *et al*., 1995), and in contrast to wildtype, *SHP2* is not activated in the fourth whorl of the flower in *sep1/2/3* triple mutant (Castillejo *et al*., 2005), perhaps due to loss of *AG*.

1.3.2.2 Gametogenesis

AGAMOUS (*AG*) is required for the specification of stamens and carpels and induces microsporogenesis, via activation of the *SPOROCYTELESS* (*SPL*; Ito *et al.,* 2004). SPL is a novel protein, related to the MADS-domain protein, which is essential for sporogenesis in both male and female organs in *Arabidopsis* plants (Yang *et al.,* 1999). The *spl* mutation causes the disruption of sporocyte formation, resulting in the absence of microspores and megaspores, without affecting other aspects of sporophytic development, with the exception of the anther walls and the nucellus (Yang *et al.,* 1999). Null alleles of the SRF-like (type I) MADS-box gene, *AGAMOUS-LIKE 80* (*AGL80*), have no effect on the male gametophyte, but in female gametophytes, the central cell's

nucleolus and vacuole fail to mature properly and endosperm development is not initiated after fertilization (Portereiko *et al.,* 2006). Despite the underrepresentation of transcription-associated transcripts in pollen transcriptome, type I and non-classical (MIKC*) MADS box genes emerge as a class with putative unique roles in pollen (Pina *et al.,* 2005).

1.3.2.3 Embryogenesis

Seed abortion in the *medea* mutant is largely mediated by deregulated expression of the SRF-like (type I) MADS-box gene *PHE1* (Köhler *et al*., 2003, 2005). Both MEA and FIE directly associate with the promoter region of *PHE1* and reduced levels of *PHE1* in the *medea* mutant seeds can suppress the seed abortion phenotype (Köhler *et al*., 2005). In the embryo defective mutant, *EMB 3008*, which arrests in the pre-globular stage of development, a SRF-like (type I) MADS-box gene (At5g39750) is disrupted (SeedGenes database; Tzafrir *et al*., 2003)

1.3.3. Combinatorial nature of MADS-domain proteins

The literature contains a plethora of data demonstrating interactions between plant MADS-domain proteins (Mizukami *et al*., 1996, Fan *et al.,* 1997, Egea-Cortines *et al.,* 1999, Moon *et al*., 1999, Lim *et al*., 2000, Honma and Goto, 2001, Immink *et al.,* 2002, Jang *et al.,* 2002, Favaro *et al.,* 2002, 2003, Causier *et al.,* 2003, Yang *et al.,* 2003a, b, Yang and Jack, 2004, Shchennikova *et al*., 2004, de Folter *et al.,* 2005, Cseke *et al.,* 2007) demonstrating a vast potential for modular based regulation.

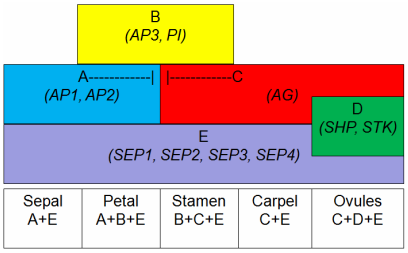
1.3.3.1 The ABC model of floral identity

The most studied example of the combinatorial nature of MADS-domain protein pertains to the floral organ identity. The quartet model is the revised “ABC” model of floral development, which makes predictions about the composition of the tetramers in the four whorls of the flower (for review see Jack, 2004). In the quartet model a combination of AP3/PI-SEP/AP1 is postulated to specify petals in whorl 2, a combination of AP3/PI-SEP/AG to specify stamens in whorl 3, and a combination of AG/AG-SEP/SEP to specify carpels in whorl 4 (Figure 1.4).

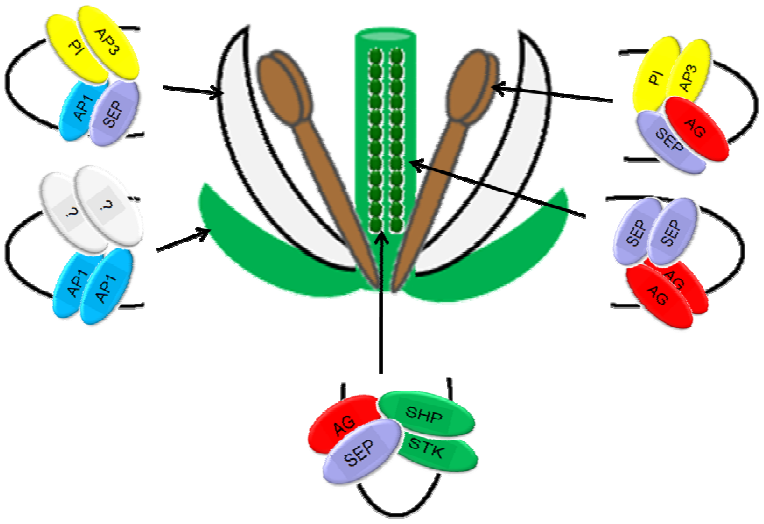
All components of the ABC model, with the exception of *AP2* are MADS-box genes. In support of the aforementioned model, quadruple transgenic plants carrying transgenes for *35S::PI*, *35S::AP3*, *35S::SEP3*, and *35S::AG* are transformed into staminoid organs (Honma & Goto, 2001). The leaves of *35S::AG*, or *35S::SEP3*, plants show subtle transformation, whereas those of plants doubly transgenic *35S::AG* and *35S::SEP3* displayed carpelloid features at a higher degree (Castillejo *et al*., 2005). Triple transgenic *35S::PI*, *35S::AP3*, and *35S::AP1* or *35S::PI*, *35S::AP3*, and *35S::SEP3*, are transformed into petaloid organs (Honma and Goto, 2001). *PI*, *AG* and *AP3* are required for petal, stamen, and carpel development, but not for sepal development. The four *Arabidopsis SEPALLATA (SEP)* genes are postulated to function redundantly to specify petals, stamens, and carpels as well as floral determinacy. Triple mutant *sep1/2/3* plants develop flowers composed entirely of sepals but show normal expression of *AG*, *AP3* and *PI*, suggesting that AG, AP3, and PI require SEP proteins for the normal development of floral organs (Pelaz *et al*., 2000). Mutant alleles of *sep3* produce petals that develop some characteristics of sepals, resembling the aberrant petals that form in intermediate alleles of *ap1*, thus suggesting that the loss of SEP3 function reduces the ability of AP1 to carry out its petal-identity function (Pelaz *et al*., 2001).

Figure 1.4 The revised "ABC" model of floral organ identity

a.



b.



This figure was constructed based on similar models published by Jack, 2004, and Theissen and Saedler, 2001, references therein, and cited in the text (1.3.3.1).

a. Floral organ identity depends on the expression of groups of genes, A thru E. All but *AP2* are MADS-box genes.

b. MADS domain transcription factors form complexes in a modular fashion to define floral organ identity.

1.3.3.2 *SEPALLATA* bridges the gap

SEP3 has been reported as directly interacting in yeast two-hybrid assays with a number of other MADS (Fan *et al.,* 1997, Pelaz *et al.,* 2001, Honma and Goto, 2001, Favaro *et al.,* 2003, Yang and Jack, 2004, de Folter *et al.,* 2005, Hill *et al.,* 2007, manuscript under review), and is postulated to function as a bridge protein in multimeric MADS-domain protein complexes (Honma and Goto, 2001, Favaro *et al.,* 2003). SEP3 has also been shown to interact with the co-repressor proteins SEUSS (SEU) and LEUNIG (LUG), and possibly recruit them to the AG promoter (Sridhar *et al.,* 2006).

1.3.3.3 Interactions with other proteins

In addition to interactions between MADS-domain proteins, there is a rapidly growing body of knowledge where plant MADS are reported to interact with other factors, such as putative transcription factors (Causier *et al.,* 2003, Masiero *et al.,* 2002), or co-repressors (Sridhar *et al.*, 2006), RNA-binding proteins (Pelaz *et al.,* 2001), post-translational modifying factors (Fujita *et al.,* 2003 Gamboa *et al.*, 2001, Yalovsky *et al*., 2000), and others (Honma and Goto, 2001, Cseke *et al.,* 2007). Both the PI homodimer and the AP3 homodimer are cytoplasmic, but the AP3-PI heterodimer is localized to the nucleus (McGonigle *et al.,* 1996). AGL15 is initially present in the cytoplasm of cells localizes to the nucleus early in embryo development (Perry *et al*., 1996). The question arises as to what causes these transcription factors to localize to the nucleus. AGAMOUS-Like 24 (AGL24) interacts with the kinase domain of MERISTEMATIC RECEPTOR-LIKE KINASE (MRLK), and is AGL24 localized to the nucleus when the two are co-expressed (Fujita *et al.,* 2003). AG interacts with FLOR1, a novel leucine-rich repeat protein and possible membrane receptor kinase (Gamboa *et al.,* 2001).

1.3.4 AGAMOUS Like 15 (AGL15)

AGAMOUS-Like 15 (AGL15) (*At5g13790*) is a member of the MIKC subfamily, which consist of four domains, the MADS (M)-, I-, K-, and C- domains (see 1.3.1). AGL15 preferentially accumulates in a wide variety of tissues that are developing in an embryonic mode (Heck *et al.,* 1995; Rounsley *et al*., 1995; Perry *et al*., 1996, 1999) and constitutive expression promotes somatic embryogenesis and can lead to long term

maintenance of development in this mode without a requirement for exogenous hormones (Harding *et al.,* 2003). *AGL15*, a promoter of somatic embryogenesis, is induced in response to auxin (Zhu and Perry, 2005), which is normally a requirement for induction of somatic embryogenesis (reviewed by Ideka *et al.,* 2006). AGL15 induces the expression of *AtGA2ox6*, which encodes for an enzyme that converts bioactive GA into inactive GA, and ectopic expression of this gene also induces somatic embryogenesis (Wang *et al.,* 2002, 2004).

Research programs lead by Dr. Sharyn Perry primarily address the role of AGL15 in embryo development. However, AGL15 is not exclusively expressed during the embryonic phase of development but is expressed at lower levels after completion of germination in restricted sets of cells including the vegetative shoot apical meristem, leaf primordia, young flower buds, and in the base of expanding lateral organs (rosette and cauline leaves, and floral organs; Fernandez *et al.*, 2000). *AGL15* and *AGL18* play a redundant role in regulating of flowering time. Recently the *agl15/agl18* double mutant was reported to flower early under short day conditions (Adamczyk *et al.,* 2007). Over-expression of *AGL15* delays fruit maturation, silique dehiscence, and seed desiccation, and all these processes involve senescence or abscission (Fang *et al.*, 2002). Ectopic expression of AGL15 also maintains development in the embryonic mode in culture (Harding *et al.,* 2003). Therefore, AGL15 might operate to maintain an immature or juvenile state.

AGL15 possesses the ability to directly interact with other MADS-domain proteins, including itself, SOC1, SVP, AP1, AGL6, AG, STK, SHP1, SHP2, AGL16, AGL21, AGL24 (de Folter *et al.,* 2005), and SEP3 (this study), some of which have overlapping expression patterns (de Folter *et al.,* 2005, Lehti-Shiu *et al.,* 2005) and are present in embryonic tissue (Lehti-Shiu *et al.,* 2005) where AGL15 accumulation is greatest (Heck *et al*., 1995, Rounsley *et al*., 1995, Perry *et al.,* 1996, 1999). Although AGL15 has been reported as a protein co-purifying in the same complex as SERK1 (Karlova *et al.,* 2006), no direct interaction between AGL15 and a non-MADS domain protein has yet been reported. MADS-domain proteins bind as either homo- or heterodimers to an A/T-rich *cis*-element named the CArG motif (C-A/T-rich-G with a canonical sequence of $CC[A/T]_6GG$; for review see Reichmann and Meyerowitz, 1997), and AGL15 has been shown to preferentially bind a CArG sequence with a longer A/T-rich core ($C[A/T]_8G$) *in*

vitro (Tang and Perry 2003). Research in the lab of Dr. Sharyn Perry has identified a number of downstream targets of AGL15 (Wang *et al.,* 2002, Tang and Perry, 2003, Wang *et al.,* 2004, Zhu and Perry, 2005, Hill *et al.,* 2007, unpublished data), and while some of these target genes are induced in response to AGL15, others are repressed. A number of direct target genes have been analyzed that exhibit strong association with AGL15 *in vivo*, yet *in vitro*, AGL15 binds only weakly. Taken together these data suggested that AGL15 may form heterodimers or ternary complexes with other proteins, thus modulating the specificity and function of AGL15 *in planta*.

1.4 Research perspectives and introduction to succeeding chapters

This study was undertaken in order better understanding how AGL15 functions, in conjunction with other factors, to regulate developmental programs. The specific aim of the project was to identify proteins capable of interacting with AGL15 and that modulate the function and/or specificity of AGL15 in a biologically relevant context. The yeast two-hybrid system has been employed to identify novel AGL15-interacting proteins and Chapters 2 through 5 describe the identification and subsequent characterization of AGL15-interacting proteins. Many tools are available to researchers studying the model plant *Arabidopsis*, and together with the short life cycle of this plant, allow projects to be undertaken that would not at this point in time be feasible in other species. While basic plant science is intellectually satisfying, the driving force and financial backing behind most research endeavors is to ultimately be able to apply what is learned to agriculturally important crops. Chapter 6 describes the initial steps taken towards identifying and characterizing a soybean *AGL15* ortholog, and discusses the potential this offers in regard to improving soybean regeneration. The identification and characterization of a gene essential for embryogenesis is described in Chapter 7. Finally, Chapter 8 surmises and reflects on the main findings presented in this study and recommends future research directions.

2.1 Introduction

MADS-domain proteins may function as both transcriptional activators and repressors. For instance, ectopic expression of SEP3 has been reported to induce *AGAMOUS* (*AG*) expression outside of the floral context (Castillejo *et al.*, 2005). However APETALA 1 (AP1) and SEPALLATA 3 (SEP3) are also able to interact with the transcription co-repressors LEUNIG (LUG) and SEUSS (SEU; Sridhar *et al.,* 2006), which prevent ectopic *AG* transcription (Franks *et al.*, 2002; Liu and Meyerowitz, 1995). Neither *LUG* nor *SEU* are predicted to encode a recognizable DNA-binding motif, but SEU has been shown to associate *in vivo* with an *AG cis*-regulatory region containing a putative CArG motif, perhaps through binding to DNA-bound SEP3 (Sridhar *et al.,* 2006). Although AGL15 has been reported as a protein co-purifying in the same complex as SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE 1 (SERK1; Karlova *et al.,* 2006), no direct interaction between AGL15 and a non-MADS domain protein has yet been reported. AGL15 possesses the ability to directly interact with other MADS-domain proteins, including itself, SOC1, SVP, AP1, AGL6, AG, STK, SHP1, SHP2, AGL16, AGL21, AGL24 (de Folter *et al.,* 2005), and SEP3 (this study), some of which have overlapping expression patterns (de Folter *et al.,* 2005, Lehti-Shiu *et al.,* 2005) and are present in embryonic tissue (Lehti-Shiu *et al.,* 2005) where AGL15 accumulation is greatest. Previous research has identified a number of downstream targets of AGL15 (Wang *et al.,* 2002, Tang and Perry, 2003, Wang *et al.,* 2004, Zhu and Perry, 2005, S. Perry, unpublished data), and while some of these target genes are induced in response to AGL15, others are repressed. A number of direct target genes have been analyzed that exhibit strong association with AGL15 *in vivo*, yet *in vitro*, AGL15 binds only weakly. Taken together these data suggest that AGL15 may form heterodimers or ternary complexes with other proteins, thus modulating the specificity and function of AGL15 *in planta*. The yeast two-hybrid system has been used to address this question and identify AGL15-interacting proteins (Chapter 4).

Here it is reported that AGL15 interacts with members of the SWI-INDEPENDENT 3/HISTONE DEACETYLASE (SIN3/HDAC) complex, and that a conserved LxLxL motif present in the C-terminal domain of AGL15 is required for its association with one member of this complex named SIN3 ASSOCIATED POLYPEPTIDE OF 18 KD (SAP18), in yeast two-hybrid studies. It is also shown that AGL15 functions as a

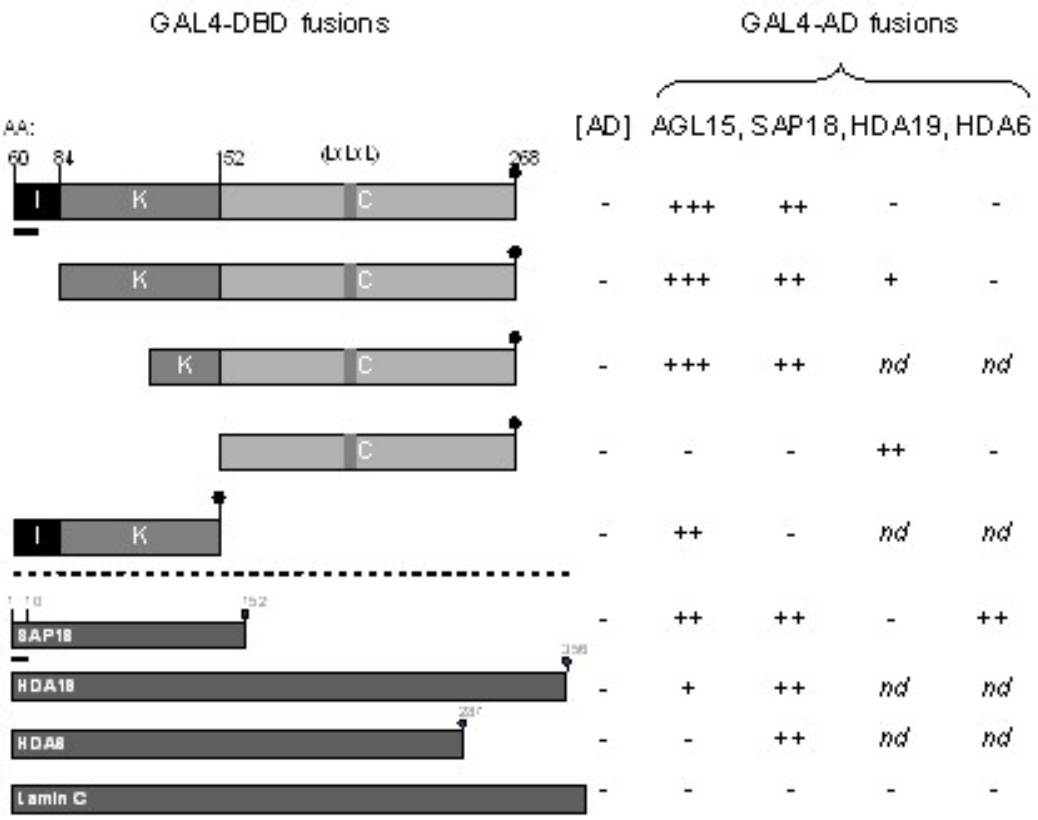
transcriptional repressor *in vivo*, and that the region where the LxLxL motif resides is essential to the repressive function. The interaction between AGL15 and members of the SIN3/HDAC1 complex suggests a mechanism that could explain its function as a transcriptional repressor *in planta*.

2.2 Results

2.1.1 AGL15, but not other MADS-domain proteins, associates with members of the AtSIN3 histone deacetylase complex

SIN3 ASSOCIATED POLYPEPTIDE OF 18 KD (SAP18) was recovered as an AGL15-interacting protein from a yeast two-hybrid screen (see Chapter 4) of a cDNA library derived from an *Arabidopsis* embryonic culture tissue (ECT; Harding *et al*., 2003). The recovered clone contained full length *SAP18* (*At2g45640*) cDNA, in frame with the GAL4 activation domain (AD), and was unable to activate reporter genes (*HIS3*, *ADE2*, and *MEL1*) in the presence of the GAL4-DNA binding domain (DBD) alone or the GAL4-DBD fused to lamin C (Figure 2.1 and not shown). A series of yeast two-hybrid assays, performed using truncated forms of AGL15 as bait, revealed that the latter half of the K domain together with the C-domain (AA 118 – 268) was required for the interaction between AGL15 and SAP18 (Figure 2.1). While the C-domain alone was not sufficient to mediate an interaction between AGL15 and SAP18, it was necessary (Figure 2.1). Like AGL15, SAP18 was also able to form homodimers in the yeast system (Figure 2.1). SAP18 directly interacted with putative members of the SIN3 histone deacetylase complex, HDA19 (At4g38130; Figure 2.1, Song and Galbraith, 2006) and HDA6 (At5g63110; Figure 2.1). AGL15 was also able to directly interact with HDA19 via the C-domain alone (Figure 2.1). Unlike some AGL15-interacting proteins, in which the interaction domain had been mapped to the relatively well conserved K-domain (Chapter 4), SAP18 did not interact with any of the other MADS domain proteins tested. The MADS domain proteins FLC, SVP, AGL18, SOC1, PI, SEP3 all failed to show any interactions with SAP18 in the yeast system (Chapter 4, Figure 4.4).

Figure 2.1 Yeast two-hybrid assays demonstrate homodimerization of AGL15 and SAP18, direct interactions between AGL15, SAP18 and HDA19 and between SAP18 and HDA6



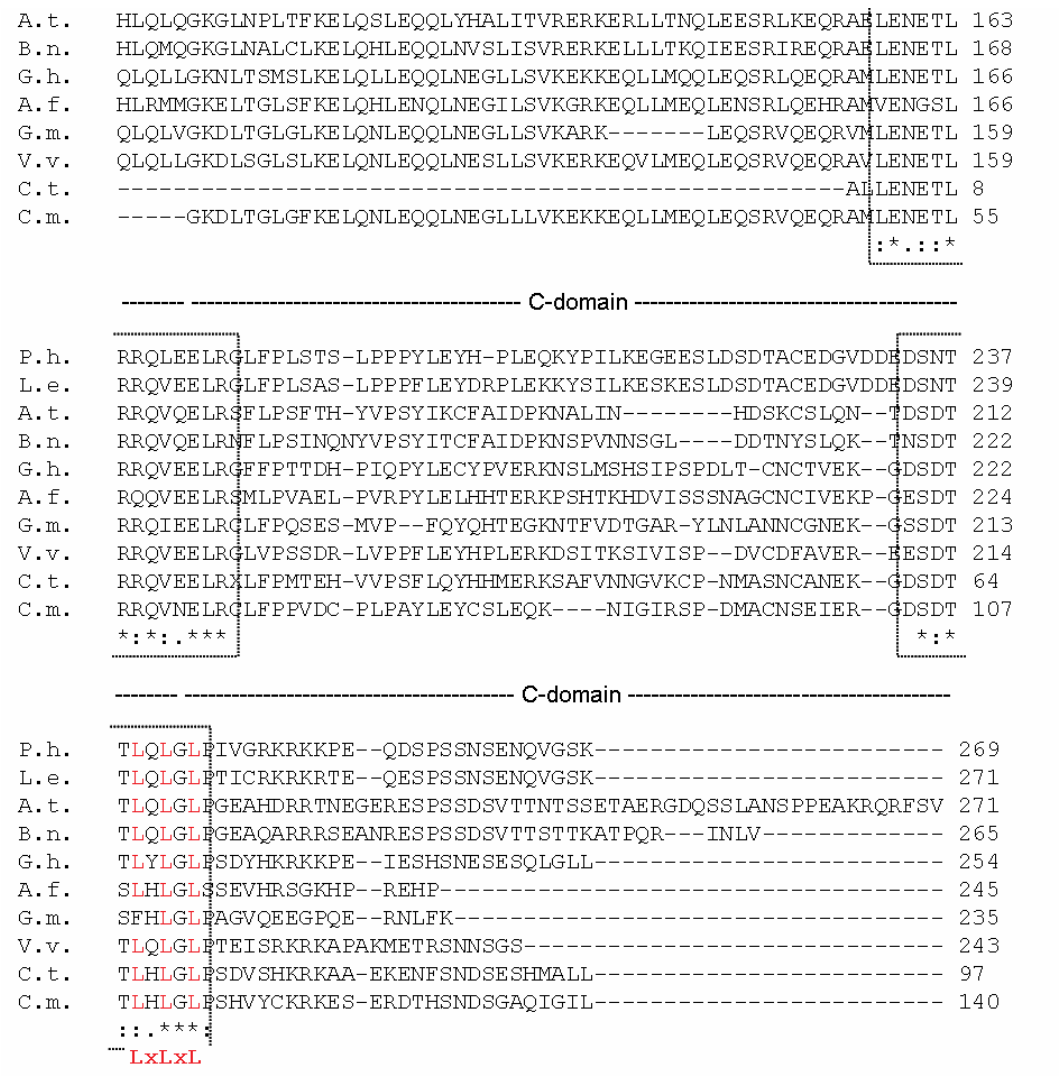
The form of AGL15 cloned into the prey construct for the data shown here contained the KC domains. All constructs above the dotted line represent different forms of AGL15; below the dotted line are non-AGL15 bait constructs.

Black bar, 10 aa; nd, not determined; +++, strong activation of reporter genes; ++, moderate activation of reporter genes; +, weak activation of reporter genes; +/-, very weak/transient activation of reporters genes; -, no activation of reporter genes (for a visual representation of how growth was scored see Figure 4.8).

2.2.2 The C-terminal domain of AGL15 contains two conserved motifs, one of which is a putative LxLxL repression motif, and resides within a region required for the repressive function of AGL15 *in planta*

While all plant MADS-domain proteins share a highly conserved DNA-binding MADS domain, the C-terminal domains are highly divergent. Thus conserved regions, especially within the C-terminal domain of AGL15, may shed light on its potential function. Putative AGL15 orthologs were obtained from EST databases and their translated amino acid sequences aligned using Clustal W software (Chenna *et al*., 2003). Two putative domains are apparent within the C-terminal domain of AGL15; LENETLRRQxxE and S[D/N]T[T/S]LxLGLP (Figure 2.2). The latter is strikingly similar to the EAR motif ([L/F]D L N [L/F](X) P) and the AUX/IAA LxLxL repression domains (Ohta *et al.,* 2001, Tiwari *et al*., 2004) reported in plant proteins. Thus transient expression assays were initiated to see if AGL15 could also function as a transcriptional repressor *in vivo*.

Figure 2.2 Multiple alignments of putative AGL15 orthologs reveal a conserved motif in the C-terminal domain, which shows similarity to an LxLxL repression domain

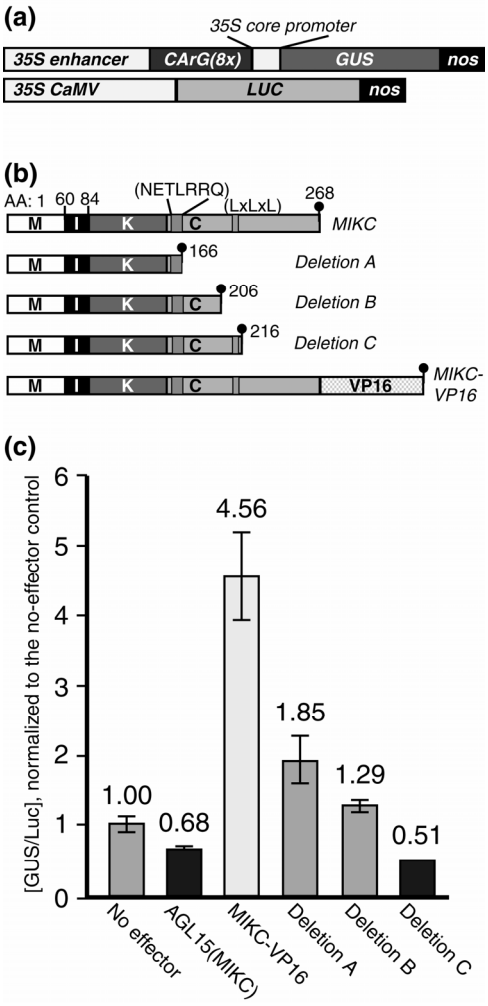


P.h., *Petunia x hybrida*; L.e., *Lycopersicon esculentum*; A.t., *Arabidopsis thaliana*; B.n., *Brassica napus*; G.h., *Gossypium hirsutum*; A.f., *Aquilegia formosa x Aquilegia pubescens*; G.m., *Glycine max*; V.v., *Vitis vinifera*; C.t., *Cyamopsis tetragonoloba*; C.m., *Cucumis melon*

Clustal W software (Chenna *et al*., 2003) was used to align AGL15 with other MADS-domain proteins.

In the transient expression system, two reporter constructs (Figure 2.3a) were co-infiltrated with different effector constructs (Figure 2.3b). One reporter (*35S:LUC*) served as an internal control. The other reporter also had the *35S* regulatory regions but the *35S* enhancer was separated from the *35S* minimal promoter by eight tandem repeats of a high affinity AGL15 binding site (Tang and Perry, 2003). A GUS/LUC ratio was calculated and the no effector control used to normalize all other results. As expected, AGL15 with a C-terminal fusion to VP16, a strong transcriptional activation domain (Dalrymple *et al*., 1985), showed significant activation of the GUS reporter indicating that the fusion protein could bind and activate expression of GUS (Figure 2.3c). Likewise, a form of AGL15 where the C-terminal domain was replaced with that of APETALA 1 (AP1), previously shown to function as a transactivation domain (Cho *et al.,* 1999, Ng and Yanofsky, 2001, Pelaz *et al.,* 2001, Honma and Goto, 2001), also had a significantly higher GUS/LUC ratio than the no effector control (data not shown). Figure 2.3c represents one of several independent experiments, and while absolute numbers varied, a trend is apparent and consistent, namely that full-length AGL15 and deletion C (that both contained the LxLxL motif) repressed expression of the GUS reporter, while deletion A and deletion B (that did not contain the LxLxL) did not. An interesting observation is that deletion A and B, which both contained the first half of the C-domain and the conserved LENETLNRRQxxE motif, appeared to be behaving as transcriptional activators in some experiments (Figure 2.3c). However this observation did not always hold true and in several experiments the GUS/LUC ratio was not significantly different from the no-effector control. A form of AGL15 lacking the entire C-domain (*i.e.*, MIK domains) never activated the reporter genes (data not shown) and only forms of AGL15 with the region of the C-domain containing the LxLxL motif showed reduced GUS activity indicating repression of *GUS* expression.

Figure 2.3 AGL15 acts as a transcriptional repressor *in planta* and this repressive activity requires the region of the C-domain containing the LxLxL motif

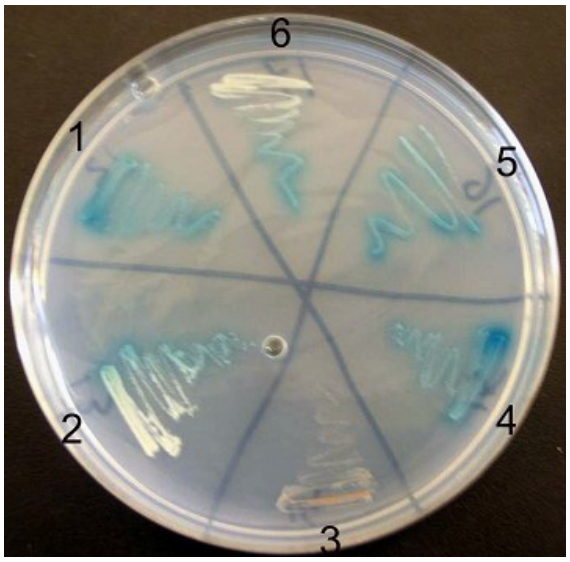


a. Two reporter constructs:- The *35S::LUC* construct, driven by the *35S* promoter, constitutively expresses *LUCIFERASE* and acts as an internal control. The second reporter consists of a modified *35S* promoter where the enhancer (-832 to -52) and the *35S* minimal promoter (-51) have been separated by eight copies of a high affinity AGL15 binding site (CArG, Tang and Perry, 2003) driving the expression of *GUS*.
b. Effector constructs: Truncations and fusion proteins of AGL15 under the control of the *35S* promoter
c. The effect various forms of AGL15 have on the transcription of a reporter gene. The Y-axis represents the ratio of GUS to luciferase activity, normalized to the no-effector control, and the X-axis shows the various effectors. The bars indicate the standard error of three data points. Different shades of grey indicate significant and consistent differences in GUS/LUC ratio resulting from different effectors. The experiment shown is representative of four biological replicates.

2.2.3 The LxLxL is required for the interaction of AGL15 with SAP18, but not with other partners

Because the LxLxL motif has been reported to function as a repression domain in other proteins (Ohta *et al.,* 2001, Tiwari *et al*., 2004) including AGL15 (Figure 2.3), and because the HDAC1/SIN3 complex generally functions to repress gene expression (for review see de Ruigter *et al.,* 2003), the question was asked whether the LxLxL motif was necessary to mediate the interaction between AGL15 and SAP18. It was found that mutating the LxLxL motif by replacing the leucine residues with alanine abolished the yeast two-hybrid interactions between AGL15 and SAP18, but not between AGL15 and other interacting proteins (Figure 2.4). The SEP3-AGL15 interaction was dependent on the latter half of the K-domain and the C-domain, which includes the LxLxL motif (data not shown), but mutation of the leucine residues in the aforementioned motif had no effect on the SEP3-AGL15 interaction (Figure 2.4).

Figure 2.4 AGL15 requires the LxLxL motif to interact with SAP18 in yeast



1. DBD-AGL15 (KC)*** + AD-SEP3 (MIKC)
2. DBD-AGL15 (KC) + AD-SEP3 (MIKC)
3. DBD-AGL15 (KC)*** + AD-SAP18
4. DBD-AGL15 (KC) + AD-SAP18
5. DBD-AGL15 (KC)*** + AD-AGL15 (IKC)
6. DBD-AGL15 (KC) + AD-AGL15 (IKC)

*** (Mutated LxLxL to AxAxA)

AH109 cells were co-transformed with a bait and prey construct and selected for on SD/-LW plates, and then re-streaked on SD/-LWHA, X-α-gal. The pink growth is dying cells resulting from heavily streaking colonies from SD/-LW plates. DBD refers to the DNA-binding domain of GAL4, whereas AD refers to the activation domain of GAL4.

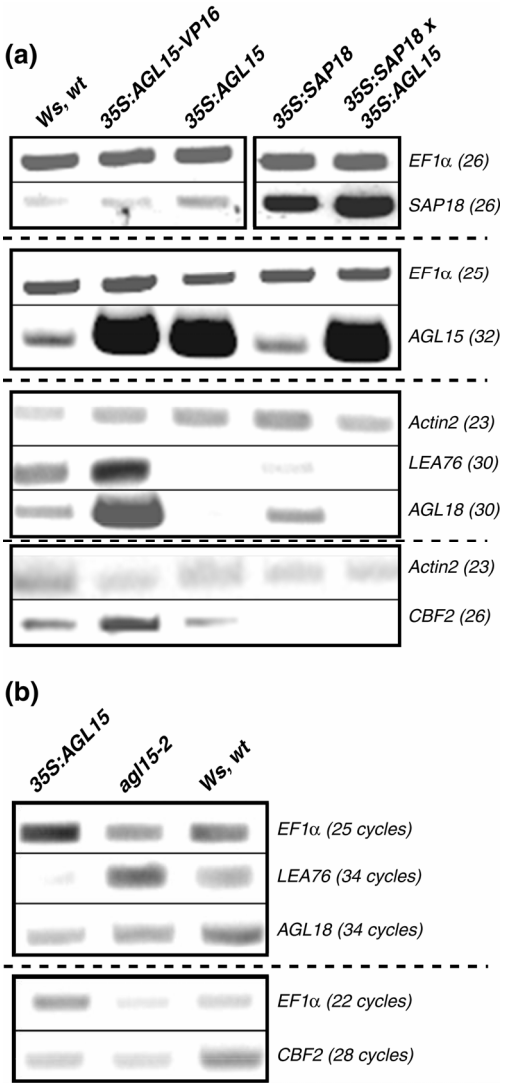
Mutating the LxLxL motif to AxAxA abolished the interaction of AGL15 with SAP18 (compare sector 3, DBD-AGL15 (KC)*** to sector 4, DBD-AGL15 (KC)). However, the interactions of AGL15 with SEP3 (compare sectors 1 and 2) and with itself (compare sectors 5 and 6) were not compromised by this mutation.

2.2.4 AGL15 decreases transcript accumulation of *LEA76*, *CBF2*, and *AGL18* indicating repression of expression

LEA76 (At1g52690) and *CBF2* (At4g25470) were first identified as putative downstream targets of AGL15 from a preliminary microarray experiment, comparing mRNA transcript amounts from Ws seeds and seed homozygous for a null allele of *AGL15* (*agl15-2*) (*CBF2* 2.99 fold change comparing *agl15-2* to Ws, P-value 0.02; *LEA76* 1.5 fold change, P-value 0.0286) (C. Zhu and S. Perry, unpublished data). Subsequent semi-quantitative and quantitative RT-PCR experiments, performed using independently harvested seed samples, confirmed the microarray data. *LEA76* and *CBF2* transcript levels were also higher in the seeds homozygous for another null allele of *AGL15* (*agl15-4*), relative to the Columbia wild type control (data not shown). *LEA76* mRNA was greatly reduced in seeds and 5-6 day siliques (data not shown), and seedlings (Figure 2.5) constitutively expressing *AGL15*. *LEA76* transcript was present at increased amounts in *agl15-2* seedlings compared to Ws (Figure 2.5b), and present at decreased amounts in *AGL15* constitutively expressing seedling compared to Ws (Figure 2.5a, b). What is more, *LEA76* was up-regulated in seedlings over-expressing a form of *AGL15* fused to a strong transcriptional activation domain (AGL15-VP16) (Figure 2.5a).

In contrast to accumulation of *LEA76* transcript in response to decreased amounts of AGL15 in seedlings, no obvious difference in *CBF2* transcript in *agl15-2* and Ws seedlings was observed, which could be due to genetic redundancy with other MADS-box genes expressed in the seedling but not present in the seed. However *CBF2* was repressed, although not as drastically as in seeds, in seedlings constitutively expressing *AGL15* relative to wild type plants and induced in seedlings accumulating the AGL15-VP16 fusion protein (Figure 2.5a, b). We have performed crosses between plants carrying a *GUS* reporter gene under the control of the *CBF2* promoter (Fowler *et al.,* 2005, generously donated by Dr. Michael Thomashow, Michigan State University), and an *AGL15* over-expresser line or an *AGL15-VP16* over-expresser line. The latter showed stronger GUS staining throughout the entire seedling, while the former appeared to have less GUS activity relative to the uncrossed *pCBF2:GUS* control (2.6).

Figure 2.5 RT-PCR on 6 day old seedlings shows transcriptional changes in downstream target genes *LEA76, AGL18 and CBF2* in response to SAP18 and AGL15



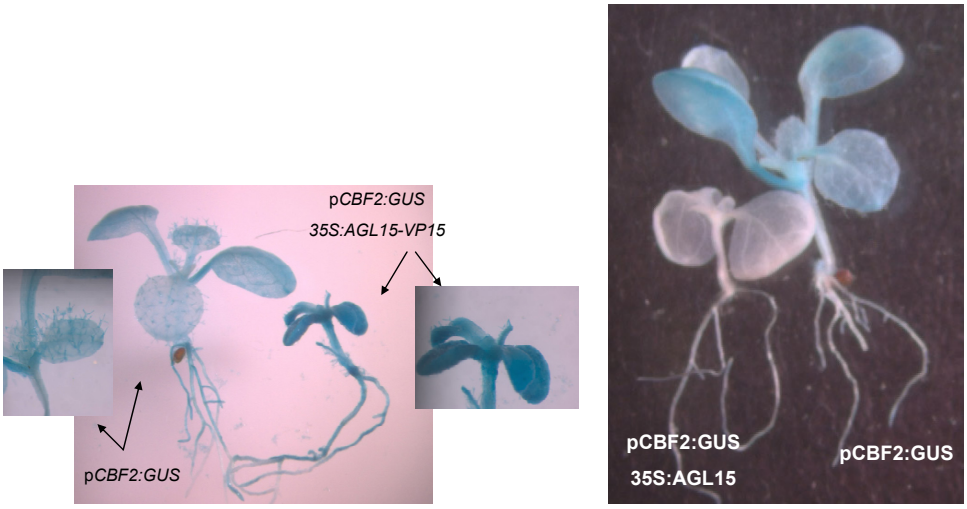
a.*AGL15* and *SAP18* transcript accumulated in the relevant over-expressing lines. *LEA76* transcript abundance was repressed by ectopic expression of *AGL15* and/or *SAP18* and induced by *AGL15-VP16*. *AGL18* was repressed by ectopic expression of *AGL15* but not by *SAP18,* and was induced by *AGL15-VP16*. *CBF2* was repressed by ectopic expression of *AGL15* and/or *SAP18* and induced by *AGL15-VP16*. Cycle numbers are shown in parentheses.

b. *LEA76* transcript levels were increased in *agl15-2* seedlings relative to Ws and decreased in AGL15 over-expressing lines relative to Ws. There was no apparent increase in *AGL18* or *CBF2* transcript in *agl15-2* seedlings, but *AGL18* and *CBF2* transcripts were decreased in *AGL15* over-expressing plants.

Results shown are representative of at least three biological replicates.

Figure 2.6 Expression of GUS in seedlings carrying the p*CBF2:GUS* reporter increases in response to *35S:AGL15-VP16* decreases in response to *35S:AGL15*

a. **b.**



Plants carrying the p*CBF2:GUS* reporter construct were crossed to plants ectopically expressing either *AGL15* or an activated form of *AGL15*, *AGL15-VP16*.

a. Expression of GUS in seedlings carrying the p*CBF2:GUS* reporter increases in response to AGL15-VP16
b. Expression of GUS in seedlings carrying the p*CBF2:GUS* reporter decreases in response to AGL15

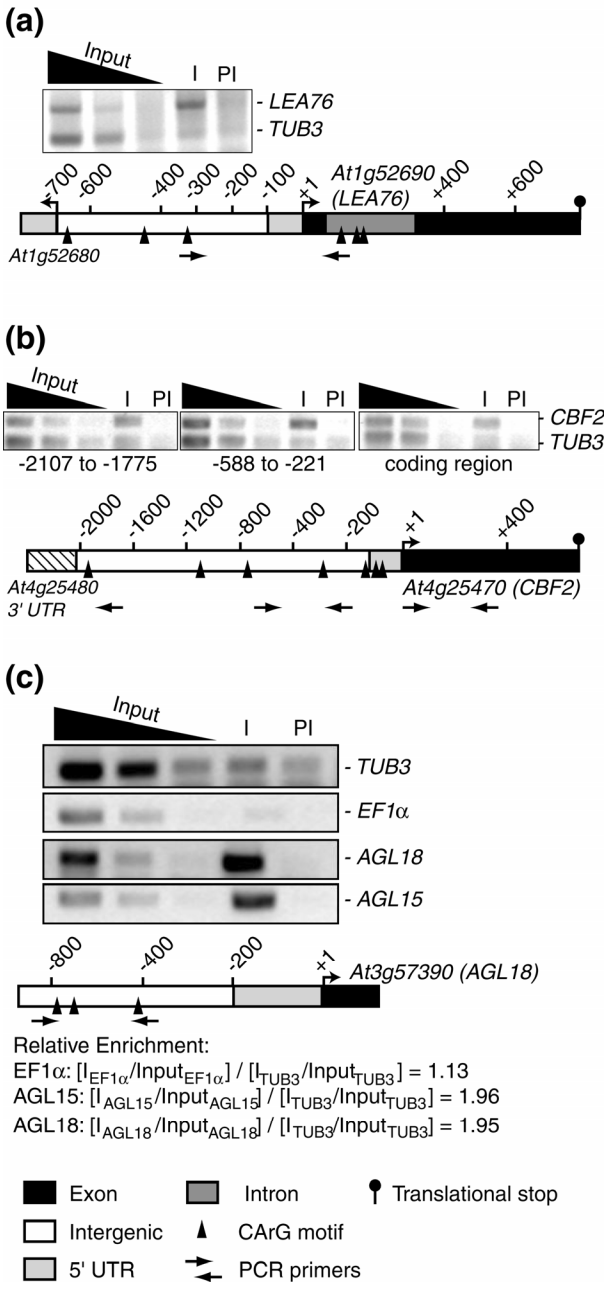
AGL15 has previously been shown to directly bind and repress *AGL15* transcript accumulation (Zhu and Perry, 2005). *AGL18* is the closest relative of AGL15 and is expressed in an overlapping pattern (Adamczyk *et al.,* 2007). Here it is shown that although no change in *AGL18* transcript was observed in plants homozygous for null alleles of *AGL15* in seeds, siliques (data not shown), or seedlings, constitutively expressing *AGL15* via the *35S* promoter resulted in a reduction in *AGL18* transcript accumulation (Figure 2.5a, b). What is more, levels of *AGL18* transcript were greatly increased in seedlings constitutively expressing *AGL15-VP16* (Figure 2.5a).

2.2.5 AGL15 binds to the promoter regions of *LEA76, CBF2*, and *AGL18 in planta*

Changes in transcript amounts of AGL15 downstream target genes could be an indirect consequence of AGL15 accumulation. However, association of AGL15 with the chromatin of the regulatory regions of its target gene would suggest a more direct regulation mechanism. Chromatin immunoprecipitation (ChIP) and enrichment test experiments were performed to test whether *LEA76, CBF2*, and *AGL18* regulatory regions may be bound *in vivo* by AGL15 and therefore may represent direct targets of AGL15. In ChIP, *in vivo* formed AGL15-DNA fragment complexes are immunoprecipitated using AGL15-specific antiserum. Association of AGL15 with suspected target DNA fragments was tested using an enrichment test where oligonucleotide primers that amplify select DNA fragments that are believed to be bound by AGL15 and oligonucleotide primers that amplify a control DNA fragment not expected to be bound by AGL15 (*e.g.,TUB3* or *EF1α*) are used in multiplex PCR. In the input (total) DNA, target PCR product and control PCR product should be present at approximately equal amounts (assuming equal efficiency of PCR). After selection of DNA by ChIP, target fragments should be enriched over controls and the PCR products should reflect this.

As shown in Figure 2.7, AGL15 directly associated with the promoter regions of *LEA76* in ECT (data not shown) and imbibed seeds (Figure 2.7a), and *CBF2* in ECT (Figure 6b). AGL15 has previously been shown to directly bind its own promoter and repress expression (Zhu and Perry, 2005). Here we show that AGL15 also associated with the *AGL18* promoter in ECT (Figure 2.7c). Figure 2.7c also demonstrates that in AGL15-immunoprecipitated chromatin, the *AGL18* regulatory region was as enriched as that of *AGL15*.

Figure 2.7 AGL15 directly binds to downstream target genes *LEA76, CBF2, and AGL18*



a. AGL15 binds *LEA* promoter in imbibed seeds
b. AGL15 binds *CBF2* in ECT
c. AGL15 binds *AGL18* promoter as strongly as *AGL15* promoter in ECT

Oligonucleotide primer pairs to amplify regulatory regions of suspected direct targets of AGL15 and to amplify control regions not expected to be bound by AGL15 (*TUB3* and *EF1α*) were used in multiplex PCR on total DNA diluted 25-, 125-, and 625-fold and on DNA recovered by immune precipitation (I) or pre-immune precipitation (PI) in ChIP.

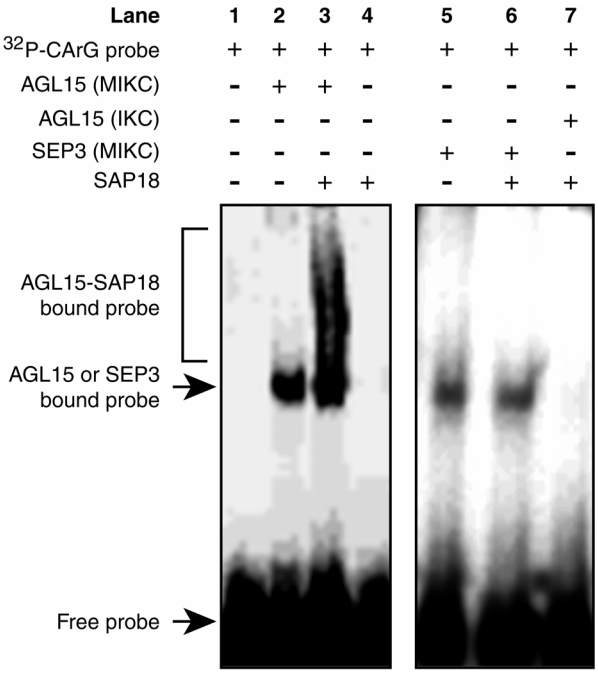
2.2.6 SAP18 represses *LEA76* and *CBF2*, but has no effect on transcript accumulation of *AGL15* or *AGL18*

Although SAP18 is not believed to bind DNA directly (for review see Silverstein and Ekwall, 2005), over-expression of *SAP18* alone, in a wild type background, was sufficient to reduce *LEA76* and *CBF2* transcript amounts in seedlings (Figure 2.5a). Conversely accumulation of SAP18 alone did not appear be involved in the decreased *AGL15* or *AGL18* transcript abundance (Figure 2.5a).

2.2.7 SAP18 does not bind to an AGL15-binding site *in vitro*, but does cause a retardation of CArG-bound AGL15 but not SEP3

Because SAP18 has been shown to interact with DNA-bound ERF3 *in vitro* but not with the ERF3 binding site (Song and Galbraith, 2006), we tested whether SAP18 was able to associate with DNA-bound AGL15. We found that SAP18 was indeed able to retard the migration of DNA-bound AGL15 through a polyacrylamide gel (Figure 2.8, compare lane 2 to lane 3, and to lane 7 in which the DNA-binding MADS domain of AGL18 is not present). SAP18 was not able to directly interact with a strong AGL15 binding site (Figure 2.8, lane 4), and neither did it cause a shift in retardation of CArG-bound SEP3 (Figure 2.8, compare lane 5 to lane 6).

Figure 2.8 SAP18 does not bind to an AGL15-binding site *in vitro*, but does cause a retardation of CArG-bound AGL15, but not SEP3



Electrophoretic mobility shift assay to assess the interaction of AGL15, SEP3, and SAP18 with a radiolabeled probe containing a binding site for MADS-domain proteins of the form $C[A/T]_8G$ that shows particularly strong interaction with AGL15. :Lane 1 contains the radiolabeled probe alone while lane 2 contains AGL15 (6 µg) and the ^{32}P-CArG probe. Lane 3 contains 3 µg each of AGL15 and SAP18 while lane 4 contains 6 µg SAP18. Lanes 5 through 7 represent a separate experiment testing binding of the ^{32}P-CArG probe by the MADS-domain protein SEP3 (3 µg, lane 5), compared to SEP3 and SAP18 together (3 µg of each protein, lane 6). Lane 7 contains a form of AGL15 lacking the DNA-binding MADS-domain (5 µg, IKC) with SAP18 (3 µg).

2.3 Discussion

2.3.1 AGL15 associates with members the AtSIN3 histone deacetylase complex

SAP18 (SIN3 ASSOCIATED PROTEIN OF 18 KD) was recovered as an AGL15-interacting protein from a yeast two-hybrid screen (Chapter 4). SAP18 is a member of the core SWI-INDEPENDENT/HISTONE DEACETYLASE (SIN3/HDAC) complex, which is conserved from yeast to man and consists of eight proteins: SIN3, HDAC1, HDAC2, RbAp46, RbAp48, SAP30, SAP18, and SDS3. However, yeast has only one HDAC, and lacks the SAP proteins (reviewed by Silverstein and Ekwall, 2005). In mammalian systems SAP18 is known to exist in a complex along with mSin3 and HDAC1 (Zhang *et al*., 1997), and in Arabidopsis with the HDAC1 ortholog, HDA19 (At4g38130; Song *et al.,* 2005). Arabidopsis has two *HDAC1* orthologs, *HDA19* and *HDA6* (At5g63110), and the Arabidopsis ortholog of SAP18 has recently been shown to interact directly with HDA19 (Song and Galbraith, 2006). Here it is shown that SAP18 also interacts with HDA6 and that AGL15 can directly interact with HDA19 via its C-terminal domain (Figure 2.1). It is also demonstrated here that SAP18 is able to homodimerize in yeast (Figure 2.1), and while the C-domain alone is not sufficient to mediate an interaction between AGL15 and SAP18, it is necessary (Figure 2.1).

2.3.2 The terminal domain of AGL15 contains two conserved motifs, one of which is a putative LxLxL repression motif

Comparison of amino acid sequences of class II ERF repressors revealed the conservation of the sequence motif L/FDLNL/F(x)P, designated the ERF-ASSOCIATED AMPHIPHILIC REPRESSION (EAR) motif and mutational analysis demonstrated that this motif is essential for repression (Ohta *et al.,* 2001). This EAR motif was also identified in a number of zinc-finger proteins from wheat, Arabidopsis, and petunia plants, and these zinc finger proteins functioned as repressors, with their repression domains mapping to regions that contained an EAR motif (Ohta *et al.,* 2001). Additionally, Aux/IAA proteins contain a LxLxL motif that functions in repression (Tiwari *et al.,* 2004). Here it is demonstrated that the repressive capacity of AGL15 is dependent on the presence of the C-terminal domain containing the LxLxL motif (Figure 3c). What is more, the LxLxL was required for the interaction of AGL15 with SAP18 (Figure 2.4).

Interestingly AtSAP18 has also been shown to interact with ERF3 and ERF4 (Song and Galbraith, 2006), which contain the motifs IDLDLNLAP and LDLDLNLPP respectively.

SAP18 does not interact with any other MADS domain proteins tested here (Chapter 4, Figure 4.1), which is unsurprising given that the interaction occurs at least in part via the divergent C-terminal domain and not the highly conserved MADS or moderately conserved I and K domains. What is surprising is that SAP18 did not interact, at least in the yeast system, with the closest relative of AGL15, AGL18, which also contains a similar LxLxL motif (DTSLQLGLSS) and exhibits functional redundancy *in planta* with AGL15 (Adamczyk *et al.,* 2007). However, no AGL18-interacting partners have been demonstrated using the yeast system in our lab or have been reported in the literature, even though such experiments have been performed (de Folter *et al.,* 2005, Verelst *et al.,* 2007). The literature contains a number of examples where a full-length MADS-domain protein, when fused to the GAL4-DBD is unable to interact with partners with which truncated forms do interact (Yang *et al.,* 2003a, Yang *et al.,* 2003 b, Yang and Jack, 2004, Fujita *et al.,* 2003). Additionally, there are examples where an interaction shown to occur between full-length MADS *in planta* does not occur in the yeast system (Immink *et al.,* 2002). Thus it is likely that a conformation effect brought about by the artificial nature of the GAL4-DBD fusion masks some interactions. This might be the case with AGL18 also, although the truncated form of AGL18, lacking the MADS domain also failed to interact with any of the AGL15-interacting proteins (see Chapter 4).

2.3.3 The interaction of AGL15 with members of the SIN3/HDAC1 complex suggests a mechanism that could explain its function as a transcriptional repressor *in planta*

HDAC enzymes remove acetyl groups from histones and hypoacetylation results in a decrease in the space between the nucleosome and the DNA that is wrapped around it. Tighter wrapping of the DNA diminishes accessibility for transcription factors, leading to transcriptional repression (for review, see de Ruijter *et al.,* 2003). SIN3-ASSOCIATED PROTEINS (SAP18 and SAP30) have been hypothesized to stabilize the SIN3–HDAC interaction (discussed by Silverstein and Ekwall, 2005). SAP18, like other core members of the SIN3/HDAC1 complex, is not believed to directly associate with the chromatin (reviewed by Silverstein and Ekwall, 2005) but does interact with sequence-

specific DNA binding proteins (For example, Zhu and Hanes 2000; Espinas *et al.* 2000; Song and Galbraith 2006; this study), implicating it as a bridge protein connecting the core SIN3/HDAC1 complex targeted genes. The solution structure of SAP18 reveals an ubiquitin-like fold with several large loop insertions relative to other family members (McCallum *et al.,* 2006). This fold supports the functional role of SAP18 as a protein-protein adapter module and provides insight for how SAP18 may bridge the Sin3-HDAC complex to transcription factors.

When directed to regulatory regions of genes, SAP18 or other components of the HDAC complex, lead to repression of gene expression. In transient expression assays dSAP18 inhibits the ability of Bicoid to activate reporter genes, and overexpression of dSAP18 inhibits Bicoid-dependent transcription in *Drosophila* cells (Zhu *et al*., 2001). Mammalian SAP18, when tethered to a promoter, is able to repress transcription *in vivo* (Zhang *et al*., 1997), and a GAL4 DNA binding domain-mammalian RPD3 homolog fusion protein strongly represses transcription from a promoter containing GAL4 binding sites (Yang *et al.,* 1996). In plants a GAL4-AtRPD3A (HDA19) fusion, when directed to GAL4 binding sites, is able to partially inhibit transcription of reporter genes (Wu *et al*., 2000). The transcriptional repression activity of AtERF7, a GCC-box binding protein, is enhanced by AtHDA19 and AtSin3 (Song *et al.,* 2005). The DNA-binding factor, ERF3, together with either SAP18 or HDA19 is able to repress transcription of reporter genes over ERF3 alone, and a modest additive affect was observed when both SAP18 and HDA19 were present along with ERF3 (Song and Galbraith, 2006)

Here it is demonstrated that SAP18 is able to associate with DNA-bound AGL15 *in vitro*, but is not able to directly interact with a strong AGL15 binding site (Figure 2.8). Likewise SAP18 also interacts with DNA-bound ERF3 *in vitro* but not with the ERF3 binding site (Song and Galbraith, 2006).

2.3.4 AGL15 induces expression of some direct downstream target genes and represses the expression others.

Downstream target genes of AGL15 have previously been identified whose expression is induced in response to AGL15 binding a CArG motif within the promoter region (Wang *et al.,* 2002, 2004). In the transient expression assay, when the effector lacked any portion

of the C-terminal domain, the GUS/LUC reporter ratio was no different than the no-effector control, but forms of AGL15 that included part of the C-terminal domain often, but not always, showed activation of the *GUS* reporter gene. These effectors (deletions A and B) included a conserved region (LENETLNRRQxxE), but not the LxLxL motif. The LENETLNRRQxxE (Figure 2.2) contains a number of glutamine, asparagines and acidic residues and may under certain configurations function as an activation domain (Titz and Thomas *et al.,* 2006). Alternately activation of direct AGL15 downstream targets (Wang *et al.,* 2002, 2004, S. Perry, unpublished data) may be mediated via its association with transcriptional activators. Indeed, yeast two-hybrid studies have demonstrated that AGL15 can interact with AP1 (de Folter *et al.,* 2005) and SEP3 (Figure 2.4), both of which have reported transactivation functions; AP1 (Cho *et al.,* 1999, Ng and Yanofsky, 2001, Pelaz *et al.,* 2001, Honma and Goto, 2001) and SEP3 (Honma and Goto, 2001). It is interesting to note that the conserved LENETLNRRQxxE motif resides within a region of AGL15 (AA 118-180) that is necessary and sufficient to allow interaction with SEP3 in yeast (see Chapter 5).

There are numerous examples in the literature of protein interactions determining whether a protein acts as a transcriptional activator or repressor. Known transcriptional activators have been shown to exert a repressive function through recruitment of SIN3/HDAC1. The mouse homolog of the transcriptional activator SU(FU) interacts and functionally cooperates with SAP18 to repress transcription by recruiting the SAP18-mSIN3 complex to promoters containing the Gli-binding element (Cheng and Bishop, 2002). The Drosophila transcription factor BICOID is converted from an activator into a repressor by recruitment of a co-repressor to BICOID-dependent promoters (Zhu and Hanes, 2000). The mammalian zinc-finger transcription factor, YY1, contains an acidic activation domain and a glycine-rich region that is responsible for mediating the interaction of YY1 with the RPD3/HDAC1 complex and the consequential repressive activity (Yang *et al.,* 1996). Co-activators that interact with SIN3/HDAC1 members, giving the complex an activator function, have also been reported. In the literature HDACs are generally associated with the repression of gene expression, however, in plants HDA19 has been shown to interact with bnKCP, and exerts a modest transactivation of reporter genes (Gao *et al.,* 2003). HDA19 has also been shown to interact with BnSCL1, an ortholog of the *Arabidopsis* SCARECROW-LIKE 15 (SCL15) protein and likewise exert transcriptional activation of reporter genes (Gao *et al.,* 2004).

2.3.5 AGL15 binds directly to and represses *LEA76*, *CBF2*, and *AGL18 in planta*

It is becoming increasingly apparent that auto-regulatory loops are a common phenomenon in the regulation of MADS-box genes in plants (Honma and Goto, 2000, Gómez-Mena *et al.*, 2005, Zhu and Perry, 2005). We previously reported that AGL15 represses *AGL15* transcription (Zhu and Perry, 2005). Here we show that AGL15 can also directly associate with regulatory regions of *AGL18* and repress accumulation of *AGL18* transcript (Figure 2.5a, b). What is more, levels of *AGL18* transcript are greatly increased in seedlings constitutively expressing *AGL15-VP16* (Figure 2.5a). No noticeable increase in *AGL18* transcription has been observed in plants homozygous for null alleles of *AGL15* (Figure 2.5b) or *vice versa* (unpublished data), but given that AGL15 is able to repress its own transcription (Zhu and Perry, 2005), and the close similarity and redundancy between AGL15 and AGL18 (Adamczyk *et al.,* 2007), it is possible that any increase in expression is masked by a subsequent auto-repression or higher levels of redundancy involving other MADS domain proteins.

AGL15 binds the promoter of *LEA76* and represses its transcription. There is a clear increase in *LEA76* transcript accumulation in plants homozygous for null alleles of *agl15* in all tissues tested, and a repression observed in response to increased AGL15 amounts. In contrast to what was observed in seeds, no noticeable difference in *CBF2* transcript between *agl15-2* and Ws seedlings was observed, which could be due to genetic redundancy with other MADS expressed in the seedling not present in the seed. There are numerous examples of genetic redundancy, and complex regulatory interactions among plant MADS-box genes (for review see Gutierrez-Cortines and Davies 2000). However *CBF2* is repressed (although not as drastically as in seeds) in seedlings constitutively expressing AGL15 relative to wild type plants and is induced in seedlings accumulating the AGL15-VP16 fusion protein (Figure 2.5a). *LEA76, CBF2* (Figure 2.5a), and other putative AGL15 downstream target genes (unpublished data), are not as highly induced by ectopic expression of *AGL15-VP16* as expected, especially when compared to levels of *AGL18* transcript accumulation brought about by AGL15-VP16 (Figure 2.5a). It is has been demonstrated that AGL15 and AGL18 perform at least partially redundant functions (Adamczyk *et al.,* 2007, unpublished observations), and given that AGL15-VP16 induces accumulation of *AGL18* transcript (Figure 2.5a),

AGL18 might be subsequently repressing *CBF2* and *LEA76*, thus masking to some degree AGL15-VP16 mediated transcription, although an increase in response to AGL15-VP16 is still apparent.

Ectopic expression of *SAP18* alone repressed *LEA* and *CBF2* (Figure 2.5a) possibly through interaction with *AGL15* that is expressed in non-transgenic seedlings or through interaction with other DNA-binding proteins. However, *AGL15* and *AGL18* levels were unaffected by ectopic expression of *SAP18* alone, possibly indicating a more specific interaction at these loci. We have demonstrated that AGL15 binds to *LEA76* (*At1g52690*) in ECT (data not shown) and in imbibed seeds (Figure 2.7a) and that *LEA76* is repressed by AGL15 and/or SAP18 in a variety of tissues, including seedlings (Figure 2.5). Published results support a role for SIN3/HDAC1 mediated repression in regulation of LEA76: trichostatin A, a specific inhibitor of histone deacetylase, causes an increase in *LEA76* during seed germination (Tai *et al*., 2005)

2.3.6 Possible biological role of the AGL15-SAP18 interaction

AGL15 accumulates in the nuclei of cells in the embryo beginning very early in development (by the eight-cell stage for Arabidopsis) and remains at relatively high levels throughout morphogenesis and into maturation stage (Perry *et al.*, 1996). The level of AGL15 decreases during desiccation. *AGL15* is expressed after germination in Arabidopsis, in the vegetative shoot apical meristem (SAM), but not at the transition to reproductive development, although it is once again expressed at the bases of lateral organs such as leaves, cauline leaves, and floral organs (Fernandez *et al.*, 2000). The level of expression in these tissues is generally at least 10-fold lower than found in the embryo (Heck *et al.*, 1995; Fernandez *et al.*, 2000). One possible role for AGL15 is in regulating developmental transitions during the plant life cycle. The expression data supports this hypothesis. Ectopic expression of *AGL15* via the *35S* promoter causes delays at many points in the life cycle, including late flowering, a longer period for seeds to mature, and delayed senescence and abscission (Fernandez *et al.,* 2000). Additionally, *35:AGL15* transgenic lines show increased capability for cultured zygotic embryos and for apical meristem tissue in a liquid culture system to form somatic embryos (Harding *et al.,* 2003). One possible explanation is that these tissues are delayed in exiting embryonic programs. Recently, a double mutant *agl15/agl18* was

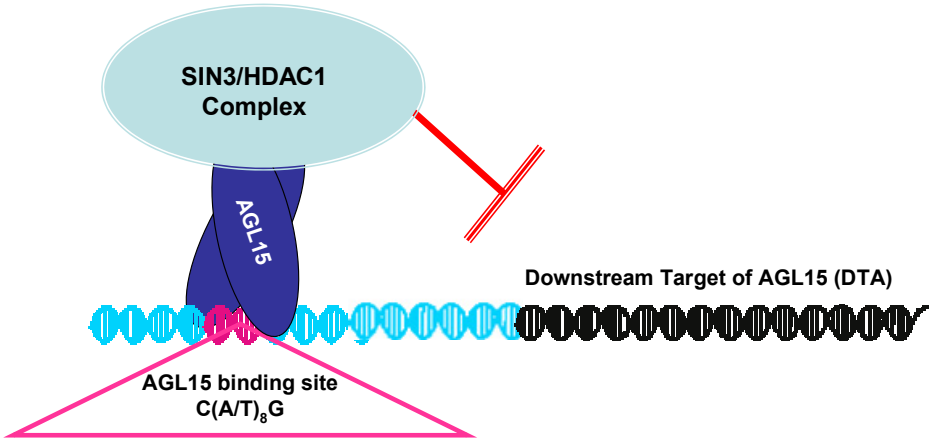
documented to flower early under short days (Adamczyk *et al.,* 2007). Epigenetic control is one means by which developmental transitions are regulated (for review, see Berger and Gaudin, 2003 Henderson and Dean, 2004, Boss *et al.,* 2004). Therefore, interaction of AGL15 with chromatin modifying proteins is very intriguing.

Phenotypes reported as being associated with HDA19 over-expression include decreased acetylation of histones, short siliques, abnormal cotyledons and leaves, late flowering, and decreased fertility (Zhou *et al*., 2005). Some of these phenotypes, notably late flowering, have also been documented in plants that constitutively express *AGL15* (Fernandez *et al.,* 2000). However, both antisense *HDA19* transgenic and *athd-t1* homozygous lines have also been reported to cause a delay in flowering, in some cases observable only when plants are grown under long day conditions (Tian and Chen, 2001, Tian *et al.,* 2003, Wu *et al*., 2000). In agreement with the latter observation, treatment of *Arabidopsis* plants with the histone deacetylase inhibitor trichostatin A also induces a late flowering affect (unpublished results, cited in Wu *et al*., 2000). Consistent with these results is the observation that a loss-of-function allele of *fld*, a component of an HDAC complex causes hyperacetylation of histones associated with *FLC*, increased transcript abundance of *FLC*, and a late flowering phenotype (He *et al*., 2003). However, loss-of-function of the histone acetyltransferase *AtHAC1* also results in increased transcript abundance of *FLC* and a delay in flowering, perhaps due to regulation of components upstream of FLC (Deng *et al*., 2007).

Despite the fact that SAP18 is a single copy gene, highly conserved even between plants and animals, plants homozygous for null alleles of *sap18* are viable and do not exhibit any obvious phenotype until the plant is stressed (Song and Galbraith, 2006). Similarly *Drosophila*, homozygous for null *sap18* alleles, are able to complete their life cycle (Singh *et el.,* 2005). AGL15 is able to interact directly with at least one other member of the SIN3/HDAC1 complex (HDA19), suggesting that SAP18 is not essential in recruiting a functional HDAC1 complex to AGL15 bound promoters. Plants over expressing SAP18 also do not have any obvious phenotype under unstressed conditions (K.Hill, unpublished observations). *SAP18* is broadly expressed, but transcript amounts increase in response to NaCl, drought, and cold stress (Song and Galbraith, 2006). Therefore regulation of targets involved in stress response such at *CBF2* are also interesting. A number of sequence-specific DNA binding proteins are likely serving as

platforms, targeting the SIN3/HDAC1 complex to these various target genes. Thus it is proposed that AGL15 functions as one of these platforms thus recruiting the SIN3/HDAC1 complex to a subset of target genes (Figure 2.9).

Figure 2.9 AGL15 likely acts as a platform recruiting the SIN3/HDAC1 complex to the promoter regions of AGL15 downstream target genes



A model depicting how AGL15 might function as a platform recruiting the SIN3/HDAC1 complex to promoters of AGL15 regulated genes.

2.4 Materials and methods

2.4.1 Yeast two-hybrid assays

The Clontech™ Matchmaker Library Construction and Screening Kit (Clontech, Mountain View, CA) was used, in accordance with manufacturers instructions, to screen a cDNA expression library, derived from Arabidopsis embryonic culture tissue (ECT; Harding *et al*., 2003) for putative protein-protein interactions involving the MADS-domain transcription factor, AGL15 (At5g13790). For a detailed description of this procedure see Chapter 4.

For directed tests AH109 cells were co-transformed with specific bait and prey constructs and plated onto SD media (2% dextrose, 0.67% nitrogen base, 1% agar) lacking leucine (-L), tryptophan (-W) (SD/-LW). After several days transformed AH109 yeast colonies were transferred onto to plates lacking histidine (-H), and alanine (-A) and supplemented with 0.2mg/ml X-α-gal (SD/-LWHA, 0.2mg/ml X-α-gal) to assay for transformants that are able to activate the reporter genes (*HIS3*, *ADE2*, and *MEL1*). For a detailed description see Chapter 4.

Bait constructs used in this study:

DBD-AGL15 (IKC), DBD-AGL15 (KC), DBD-AGL15 (½KC), DBD-AGL15(C), DBD-AGL15 (IK), DBD-SAP18, DBD-HDA6, and DBD-HDA19

The oligonucleotides and restriction sites used generate the pGBKT7 (bait) constructs can be found in Appendix A

Prey constructs used in this study:

AD-AGL15 (IKC), AD-AGL15 (KC), AD-HDA6, AD-HDA19, and AD-SEP3

The oligonucleotides and restriction sites used generate the pGADT7 (prey) constructs used in this study can be found in Appendix B

Mutagenesis

For DBD-AGL15 (KC***), AA 84-268, the leucine residues in the LxLxL motif were changed to alanine using the QuikChange II Site Directed Mutagenesis Kit (Stratagene, La Jolla, CA) as directed using the following oligonucleotides:

5'GATTCAGACACAACTGCGCAAGCAGGGGCGCCGGGAGAGGCACATG'5'C
ATGTGCCTCTCCCGGCGCCCCTGCTTGCGCAGTTGTGTCTGAATC'3

All constructs were verified by DNA sequencing carried out by UK-AGTC. Protein accumulation in yeast cells was determined via Western blotting, using AGL15 antibody (Heck *et al.,* 1995), c-myc monoclonal antibody or HA polyclonal antibody (Clontech, Mountain View, CA).

2.4.2 Multiple Alignments of Putative Orthologs

Putative orthologs of AGL15 were obtained by blasting the predicted amino acid sequence of AGL15, minus the highly conserved MADS domain, against the NCBI database (http://www.ncbi.nlm.nih.gov/BLAST/). The following entries were translated in amino acid sequences and re-blasted against the Arabidopsis database to verify that they were most similar to Arabidopsis AGL15 and then aligned using ClustalW software http://www.ebi.ac.uk/clustalw/ (Chenna *et al*., 2003). Soybean (Accession AY370659), Brassica (Accession U22681), Tomato (Accession BT014220), Petunia (Accession AY370526), Cotton (Accession AY631395), Muskmelon (Accession DV635005), Aquilegia (Accession DR941287), Wine grape (Accession EC993423), Cluster bean (Accession EG989688).

2.4.3 Generation of Transgenic Plants and growth conditions

The plant expression vector, *pBIMC-35S:cmyc-SAP18*, was generated from pDBD-SAP18 by PCR with oligonucleotides that amplify the c-myc tag and cDNA already cloned into the multiple cloning site of pGBKT7 and adding restriction enzyme cleavage sites, *Xba*I and *Xho*I (underlined):

Forward 5'GCTCTAGAATGGAGGAGCAGAAGCTG'3
Reverse 5'CGAGCTCCCTCAAGACCCGTTTAG'3

The *c-myc-SAP18* cDNA was cloned into the binary expression vector, pBIMC (a gift from Dr. D. Falcone, University of Massachusetts) downstream of the *35S* CaMV promoter. *pBIMC-35S:gAGL15 (MIKC)-VP16* encodes a form of AGL15 containing the first three introns and an additional C-terminal viral VP16 (AA 413-490) transcription activation domain (Dalrymple *et al*., 1985) cloned into pBIMC downstream of the *35S* CaMV promoter.

The constructs were checked by sequencing and then transformed into the *Agrobacterium tumefaciens* strain GV3101, and used to transform Ws flower buds using the floral dip method (Clough and Bent, 1998). Transgenic T1 seeds selected on GM plates containing 50ug/ml Kanamycin. *SAP18* transcript levels of putative c-myc-SAP18 over-expressing lines were analyzed by RT-PCR, and protein accumulation verified by Western analysis using an anti-c-myc monoclonal antibody (Sigma-Aldrich, St. Louis, MO).

Plants constitutively expressing *c-myc-SAP18* were cross pollinated to a previously described *35S:AGL15* line (Fernandez *et al.,* 2000, Harding *et al.,* 2003) to produce plants overexpressing both transgenes.

Plant growth conditions

Seeds of *A. thaliana* ecotype Wassilewskija (Ws) or Columbia (Col) were sterilized by 3-4 brief washes with 70% ethanol containing 0.1% Triton X-100, followed by two rinses with 95% ethanol, then allowed to dry in a sterile laminar flow hood before sprinkling on GM (germination media) plates containing Murashige-Skoog (MS) salts, vitamins, 1% (w/v) sucrose, 0.05% (w/v) MES, and 0.7% (w/v) agar, pH 5.8. Seeds were chilled for several days at 4°C before being grown at 20/18°C under a 16-h light/8-h dark regime. Seedlings were transplanted at 8–10 days to ProMix BX (Premier Brands, Inc., Riviere-du-Loup, Quebec, Canada) and grown in a Conviron growth chamber with fluorescent and incandescent lights set to approximately 110 μmol m^{-2} sec^{-1}.

2.4.4 Chromatin Immunoprecipitation

Tissue Fixation

Plant tissue (5-10g) was fixed in MC buffer (10 mM potassium phosphate, pH 7.0, 50 mM NaCl and 0.1 M sucrose) containing 1% formaldehyde and incubated on ice under vacuum. After 1 hr the crosslinking was stopped by the addition of cold glycine to a final concentration of 0.125 M and incubated for a further 10-30 minutes before being thoroughly washed in MC buffer and flash frozen.

Nuclei isolation

Nuclei were isolated from fixed tissue as described by Bowler *et al.,* 2004. The fixed tissue (0.5 -12g) was ground to powder with a mortar and pestle in liquid nitrogen. Approximately 5ml of EB1 (0.4 M Sucrose 10 mM Tris–HCl, pH 8.0 5 mM β-ME 0.1 mM PMSF) buffer was added per 1g of tissue. The resulting tissue slurry was filtered through Miracloth and centrifuged at 5000 x g for 20 min at 4°C in a Sorvall RC-5B centrifuge (Du Pont Instruments). The pellet was suspended in 1ml of EB2 (0.25 M Sucrose 10 mM Tris–HCl, pH 8.0 10 mM $MgCl_2$ 1% Triton X-100 5 mM β-ME 0.1 mM PMSF), transferred to a 1.5ml Ependorf tube and centrifuged at 12000 x g for 10 min at 4°C. The pellet was suspended in 400 ul of EB3 (1.7 M Sucrose 10 mM Tris–HCl, pH 8.0, 0.15% Triton X-100 2 mM MgCl2 5 mM β-ME 0.1 mM PMSF) and overlayed on top of another 400 ul of EB3 and centrifuged at 16000 x g for 30 min at 4°C.

Chromatin solubilization

The nuclear pellet was suspend in 1 ml of sonication buffer (10 mM potassium phosphate, pH 7, 0.1 M NaCl, 0.5% sarkosyl, 10 mM EDTA) with PMSF (200 mM stock in isopropanol) added to a final concentration 1mM just prior to sonication. The nuclei were sonicated for 10 to 15 sec. x 4 pulses with a probe sonicator (Fisher, Model 300 sonic dismembrator), with 2 minutes of incubation on ice in between each pulse. The insoluble material was removed by centrifugation at 12000 x g for 5 min at 4°C, and the supernatant transferred to a fresh 1.5ml ependorf tube.

Immunoprecipitation

chromatin immunoprecipitation was performed, with minor modifications, as described in Wang *et al.,* 2004, using a polyclonal antibody raised against AGL15 (Heck *et al.*, 1995; Perry *et al.*, 1996, 1999, Wang *et al.,* 2000) or preimmune serum as a control. 50ul of the supernatant was kept aside to recover the input (non-immunoprecipitated) DNA used in the PCR enrichment tests. 150ul of cold glycine elution buffer (0.1 M glycine, 0.5 M NaCl, 0.05% Tween-20, pH 2.8) and 100ul of 1 M Tris, pH 9, were added and the sample treated in the same way as the immunoprecipitated samples (see below). The remaining supernatant was divided into equal aliquots for immunoprecipitation with anti-AGL15 specific sera or preimmune sera as a control. An equal volume of immunoprecipitation buffer (50 mM HEPES, pH 7.5, 150 mM KCl, 5 mM $MgCl_2$, 10 µM ZnSO4, 1% Triton X-100, 0.05% SDS) was added to each tube along with 30ul of protein A-Sepharose beads (Invitrogen, Carlsbad, CA) and incubated for two hours at 4°C with rotation. The beads were pelleted by spinning at top speed for 1-2 min. The supernatant was removed and saved as "post-bind" to check the depletion of the protein from the supernatant. The beads were washed with immunoprecipitation buffer (1 ml each tube) for 10 minutes at room temperature with rotation and pelleted by centrifugation at top speed for 1 min. The wash and centrifugation was repeated for 3-5 times.

Elution and DNA recovery

100 µl of cold glycine elution buffer (0.1 M glycine, 0.5 M NaCl, 0.05% Tween-20, pH 2.8) was added to the beads. The sample was mixed by vortexing and pelleted by centrifugation at top speed for 1 min at room temperature. The supernatant was removed and added to a tube with 50 µl of 1 M Tris, pH 9 to neutralize. The elution and neutralization were repeated twice more to give a 450 µl total volume of the eluted sample. The eluted sample was centrifuged at top speed for 2 minutes at room temperature. The top 300 µl was moved to a new tube and the remainder (~150 µl) was saved to verify recovery of the protein by Western analyses. To the 300 µl elution sample and the input sample (non-immunoprecipitated), 1 µl RNase A (stock, 1 mg/ml) was added and then incubated at 37°C for 15-30 min. After RNase A treatment, proteinase K was added to final concentration of 0.5 mg/ml and incubated overnight at

37 °C. The following day a second aliquot of proteinase K was added and the mix was incubated at 65°C for at least 6 hours to reverse the formaldehyde crosslinks. The sample was then cooled to room temperature and chilled on ice. DNA was extracted by phenol: chloroform extraction (phenol: chloroform: isoamyl alcohol at 25:24:1). The DNA was then recovered by ethanol precipitation and used for enrichment PCR tests. The input DNA was suspended in 40ul of ddH_2O and the I and PI in 20 ul. 5-10ul of the input DNA was electrophoresed on an agarose gel to check the size of the sonicated DNA, which should ideally have been sheared into fragments ranging between 200-1000 bp,

Protein analysis

Samples reserved during ChIP were used for protein analysis. Proteins were separated on 12.5 % (w/v) polyacrylamide denaturing gels using vertical gel apparatus (Hoefer Scientific Instruments, San Francisco, CA) and then blotted onto Immobilon™ PVDF Transfer Membranes (Millipore, Billerica, MA) using a Genie blotter (Idea Scientific Co., Minneapolis, MN). Blots were blocked with 5% non-fat dry milk, washed in TBST (0.1% Tween 20, 100mM Tris-Cl, 150mM NaCl, pH 7) and probed with 1:1,000 diluted AGL15 antiserum. A 1:5,000 diluted secondary antibody (HRP-conjugated goat-anti-rabbit) and the Lumi-Glo System (Kirkegaard and Perry Laboratories, Gaithersburg, MD) were subsequently applied. Blots were exposed to Kodak XAR5 X-ray films (Eastman Kodak, Rochester, NY) and the films developed in a Konica film processor (SRX-101, Konica Corp., Tokyo, Japan). Blots were exposed to X-ray film (Kodak XAR5) for 1-5 min. Sample analyzed were typically aliquots of the soluble, nuclei-presonication, sonicated nuclei, eluted, and post-bind fractions, which were boiled in 1x sample buffer for 5 minutes.

Enrichment tests

Multiplex PCR tests were performed on a series of dilutions of total (input) DNA, DNA recovered from immunoprecipitation (I) with an anti-AGL15 antibody (Heck *et al.,* 1996), and the preimmune serum (PI). 30-35 cycles of PCR, with an annealing temperature of 52°C were performed using KlenTaq1 (Ab Peptides, St. Louis, MO).

Oligonucleotides specific for house-keeping genes, not believed to be bound by AGL15 were used as internal controls:

EF1α (At1g07920)
Forward 5'ACGCTCTACTTGCTTTCACC'3
Reverse 5'GCACCGTTCCAATACCACC'3
TUB2/3 (At5g62690 and AT5g62700)
Forward 5'GTCCTACTTTGTGGAGTGGA'3
Reverse 5'CTGTGTACCAATGCAAGAA'3

Oligonucleotides designed to amplify promoter regions potentially bound by AGL15 *in vivo* are as follows:

pLEA76 (At1g52690)
Forward 5'GTCTAACATCTTCCGTAGCTCCGTT'3
Reverse 5'TTGCCTCTGGTTTCACCAGCTTTG'3
pCBF2-1 (At4g25470)
Forward 5'TGCAAGTATTTTTAGAGCAGTAAC'3
Reverse 5'CAATAAAATATCTTCACAACGAAC'3
pCBF2-2 (At4g25470)
Forward 5'GCAATGCACGATATGTGAATGGAGA'3
5'ACGCGGAGTTTCTGTCTCTGTGAA'3
pCBF2-3 (At4g25470)
Forward 5'GAATTAGCAGAAAGGCAGAA'3
Reverse 5'GACGTGTCCTTATGGAGCTA'3
pAGL18 (At3g57390)
Forward 5'GCCACGTTTGGCCATTCTA3'
Reverse 5'ATTTCGTGTATCGCCTCCCT'3
pAGL15 (At5g13790)
Forward 5'GGAAGAAAAGGGAAAGTAGGACC'3
Reverse 5'GAGAGAAGAAGGTAGAAGGAAGA'3

PCR products were analyzed by agarose gel electrophoresis and gel images were captured using a ChemiImager (Alpha Innotech Corporation, San Leandro, CA).

2.4.5 Semi-quantitative RT-PCR

Either TRIZOL Reagent (Invitrogen, Carlsbad, CA) or the RNeasy® plant mini kit (Qiagen, Valencia, CA) was used to isolate total RNA from ~50 mg of whole 6-8 day old seedlings, grown on GM media. 1.0 µg of total RNA was first treated with DNase I (Invitrogen) and then used for first strand cDNA synthesis. Reverse transcription was performed using A-MLV Reverse Transcriptase System (Promega, Madison, WI) according to the manufacturer's instructions. 1-2 µl aliquot of each first strand cDNA reaction was amplified by specific primer pairs in a reaction containing 1x PCR buffer, dNTPs at 200 µM each, 0.2 µM of each primer and 1.0 unit of KlenTaq1 (Ab Peptides, St. Louis, MO) in a final volume of 20 µl. Amplification reactions were performed in a PCT-100 (MJ Research Inc., Watertown, MA), under conditions that varied only in the number of cycles of denaturation at 95°C (30 sec), annealing at 55°C (30 sec), and extension at 72°C (30 sec).

Control oligonucleotides specific for "house-keeping" genes were used as controls:

EF1α (At1g07920):
Forward 5'ACGCTCTACTTGCTTTCACC'3
Reverse 5'GCACCGTTCCAATACCACC'3
Actin2 (At3g18780)
Forward 5'GAGACCTTTAACTCTCCCGCTATG'3
Reverse 5'GAGGTAATCAGTAAGGTCACGTCC3'

Oligonucleotides specific to the respective target genes are as follows:

AGL15 (At5g13790)
Forward 5'TCCAAGAGGCGTTCTGGGTTACTT'3
Reverse 5'CTGCTCAAGGCTTTGCAGCTCTTT'3
SAP18 (At2g45640)
Forward 5'AAGACAAGGTGGTGGGAGACCATT'3
Reverse 5'CTCAAACGGAAGTTCGGAAAGCGT'3

LEA76 (At1g52690)
Forward 5'TAGGGCTTCGCACTGATGAAGGAA'3
Reverse 5'GGCATAACCTCACGAACGCAACAA'3
CBF2 (At4g25470)
Forward 5'AAACTCCGGTAAGTGGGTGTGTGA'3
Reverse 5'AAGACCATGAGCATCCGTCGTCAT'3
AGL18 (At3g57390)
Forward 5'ACACTACTGCGTCCACTGAGCATA'3
Reverse 5'AGAAGCCACTTGACTCCCAGAGTT'3

All RT-PCR experiments are representative of at least three biological replicates.

2.4.6 Transient Expression Assays

Petunia leaves were infiltrated with two reporter constructs (Figure 3a): The *35S:LUC* construct, driven by the *35S* promoter, constitutively expresses *LUCIFERASE* and acts as an internal control. The second reporter consists of a modified *35S* promoter where the enhancer (-832 to -52) and the *35S* minimal promoter (-51) have been separated by eight copies of a high affinity AGL15 binding site (CArG, Tang and Perry, 2003) driving the expression of *GUS*. The various effector constructs (Figure 3b) are truncations of AGL15 or fusion-proteins comprised of AGL15 and the VP16 activation domain driven by the *35S* promoter. *pBIMC-35S:gAGL15* (Fernandez *et al.,* 2000, Harding *et al.,* 2003) contains full-length *AGL15* cloned downstream of the 35S promoter and includes the first three introns; pBIMC-35S:gAGL15 deletion A encodes the first 166 amino acids, deletion B the first 206 amino acids, and deletion C the first 216 amino acids. All deletions are derivatives of *pBIMC-35S:gAGL15* and contain an engineered stop codon as indicated. *pBIMC-35S:gAGL15-VP16* encodes a form of *AGL15* containing the first three introns and an additional C-terminal viral VP16 transcription activation domain (Dalrymple *et al*., 1985) cloned into pBIMC downstream of the *35S* CaMV promoter.

Agrobacteria (GV3101) were transformed, using the freeze thaw method, with either the empty *pBIMC* vector or one of the effector constructs (Figure 2.3b), or with the *35S:CArGx8:GUS*, and *35S:LUC* constructs (Figure 2.3a). Overnight agrobacteria cultures were re-suspended to an OD600 of 0.5 in 10% sucrose containing 200 μM 3'5'-

dimethoxy,4'-hydroxyactophenone and then mixed 20:10:1, respectively. This mixture was used to infiltrate the abaxial side of young petunia leaves. After 3-5 days three samples were taken from each leaf, crushed in liquid nitrogen, suspended in 300 µL CCLR (Promega) and the used in subsequent MUG and luciferase assays. GUS (β-galactosidase) activity was measured by MUG (1mM 4 methylumbelliforyl β D glucuronide in 1XCCLR) assay (Yang *et al*., 2000) using a DynaQuant 200 fluorometer (Hoefer Scientific Instruments, San Francisco, CA), and at least three time points. GUS to LUC ratios were normalized relative to the no effector control and standard errors calculated. All transient expression assays are representative of at least three biological replicates.

2.4.7 Electro-Mobility Shift Assay (EMSA)

The protein coding regions of full-length AGL15, SAP18, SEP3, and AGL15 minus the 60 amino acid MADS domain, were cloned into an expression vector, pET-15b (Novagen, San Diego, CA). The N-terminal HIS tag was removed from AGL15 and SEP3 containing constructs by cloning these inserts into the *Nco* I site, which lies upstream of the HIS tag, and *Bam*HI. The *E.coli* strain BL21$^{(DE3)}$ (Novagen, San Diego, CA) was used to express the proteins, and inclusion bodies were harvested and solublized in 8 M urea, 1X binding buffer (10mMTris–HCl, pH 7.5, 50mM NaCl, 1mM DTT, 1mM EDTA). The inclusion bodies were subsequently dialyzed to remove urea. AGL15 binding site (CTATATATAG) probes were generated and labeled by PCR with [α - 32 P]dCTP (Amersham, Pittsburg, PA) and CArG containing oligonucleotides:

Forward 5'AGATCTGGTTACTATATATAGTAAGG'3
Reverse 5'GGATCCCCTTACTATATATAGTAACC'3

Equal amounts of probe (cpm 10^4 -10^5) were incubated for 15 minutes at room temperature with 3-8 µg protein (as indicated) in 1X binding buffer with 0.1 mg ml^{-1} poly(dI-dC), 0.5 mg ml^{-1} BSA, 5% glycerol, and resolved on a 5% polyacrylamide gel (0.5×TBE, 5% glycerol). The gel was dried and exposed to a PhosphorImager screen (Molecular Dynamics, Sunnyvale, CA).

2.5 Summary

AGAMOUS-Like 15 (*AGL15*) encodes a MADS-domain transcription factor that is preferentially expressed in the plant embryo, and may function as a regulator in embryonic developmental programs. A number of direct downstream targets of AGL15 have been identified, and while some of these target genes are induced in response to AGL15, others are repressed. Additionally, direct target genes have been analyzed that exhibit strong association with AGL15 *in vivo*, yet *in vitro*, AGL15 binds only weakly. Taken together these data suggest that AGL15 may form hetero-dimers or ternary complexes with other proteins, thus modulating AGL15's specificity and function *in planta*. Here it is reported that AGL15 interacts with members of the SWI-independent 3/Histone Deacetylase (SIN3/HDAC) complex, and that AGL15 target genes are also responsive to an AGL15 interacting protein that is also a member of this complex, SIN3 Associated Polypeptide of 18 KD (SAP18). AGL15 can repress transcription *in vivo,* and a region essential to this repressive function contains a motif that is conserved among putative orthologs of AGL15. What is more, the aforementioned motif mediates the association of AGL15 with SAP18 in yeast two-hybrid assays, thus providing a possible mechanism for AGL15's role in regulating gene expression via recruitment of an HDAC complex.

3.1 Introduction

GLYCINE RICH PROTEIN 2 (GRP2; At4g38680) was recovered as an AGL15-interacting protein from two independent yeast two-hybrid screens (see Chapter 4). *GRP2* encodes a protein containing a conserved COLD SHOCK DOMAIN (CSD), and two CCCH zinc finger motifs, interspersed by glycine rich regions (Kingsley & Palis, 1994). The high sequence similarity of GRP2 to known nucleic acid binding proteins, including a wheat CSD-containing protein (Karlson *et al*., 2002) rendered it an interesting candidate for interaction with AGL15 in a biologically relevant context. AGL15 accumulates in the nuclei of cells in the embryo beginning very early in development and remains at relatively high levels throughout morphogenesis and into maturation stage (Perry *et al.*, 1996). *GRP2* promoter activity can also be observed in the cells of the globular through to torpedo stage embryos (Fusaro *et al.,* 2007). What is more, down-regulation of *GRP2* produces plants that flower early, have altered stamen number, and are defective in seed development (Fusaro *et al.,* 2007). These observations are consistent with GRP2 playing role in development, perhaps via its interaction with AGL15.

The observation that ectopic expression of *AGL15* appears to enhance the freezing stress survival rate of seedlings is also of particular interest because low temperature is one of the most important environmental factors affecting plant growth and crop yields. *GRP2* is one of four CSD-encoding genes present in the Arabidopsis genome, all of which likewise encode CCCH zinc finger motifs, interspersed by glycine rich regions. Thus it was especially intriguing to discover that AGL15 binds the promoter of two of these genes *in vivo*, and regulates the transcription of at least one of them. The regulation of CSD-containing proteins, many of which are known to function as RNA-chaperones in other organisms thus permitting translation at low temperatures (Jiang *et al.,* 1997, Graumann and Marahiel, 1998), may account of the enhanced cold survival of *35S:AGL15* seedlings.

3.2 Results

3.2.1 GRP2 interacts with AGL15 and other MADS domain proteins in yeast two-hybrid assays

GLYCINE RICH PROTEIN 2 (GRP2; At4g38680) was recovered as an AGL15-interacting protein from two independent yeast two-hybrid screens, one using full-length (MIKC) AGL15 as bait and another with AGL15-IKC as bait (see Chapter 4). In order to elucidate the regions of AGL15 (At5g13790) that mediate the interaction with GRP2 (At4g38680), yeast two-hybrid assays were performed using various truncations of AGL15 as bait (Figure 3.1a). The minimum region of AGL15 able to interact with GRP2 lies between amino acids 84-105 (Figure 3.1b), a region corresponding to the first predicted α-helix of the K domain (Chapter 4, Figure 4.5).

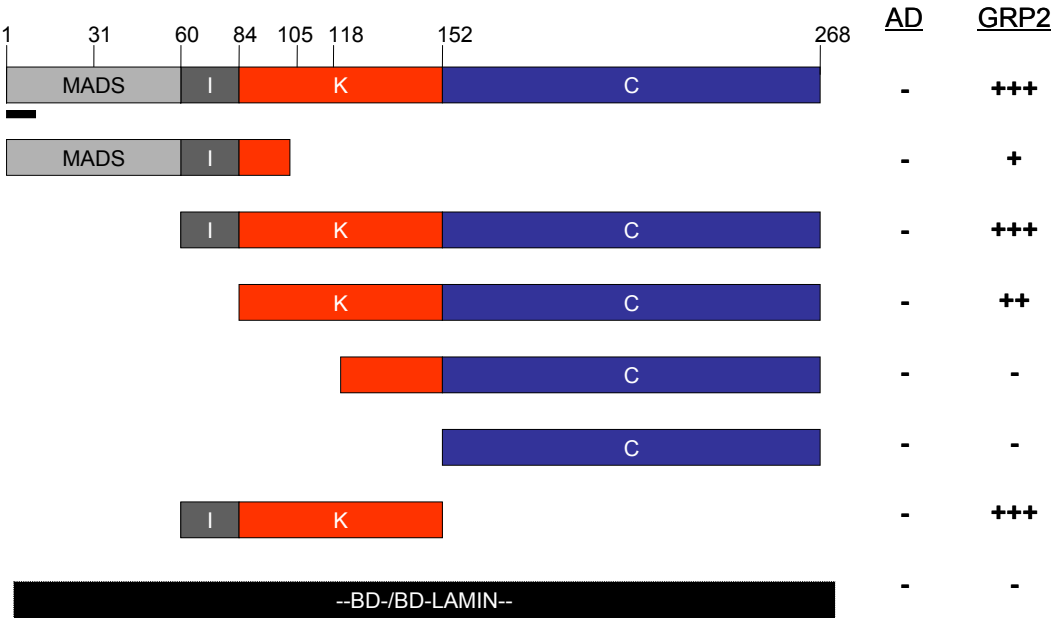
All *GRP2*-containing prey, recovered from yeast two-hybrid screens, encoded for full-length, or almost full-length GRP2 (Chapter 4). A truncated form of GRP2 (AA, 85-203), which lacks the CSD was subsequently found to be incapable of interacting with AGL15 (MIKC or IKC) in yeast two hybrid assays (Figure 3.2). This suggests that the CSD might be required for the interaction. However, yeast transformed with DBD-AGL15 and AD-CSD (GRP2 AA, 1-132), which encodes for the CSD and first glycine rich region, failed to grow on media selective for constructs (SD/-LW), indicating that either the co-transformation was unsuccessful or that transformed cells were not viable. The former seems unlikely because the experiment was repeated several times along side successfully co-transforming combinations.

Yeast two-hybrid assays were performed in order to determine if GRP2 interacts specifically with AGL15, or if it is also able to interact with other MADS-domain proteins. It was found that in addition to AGL15, GRP2 is able to interact with SVP, SOC1, SHP1, PI, and SEP3 in yeast two-hybrid assays (Figure 3.3). The exception was AGL18, for which no interaction with other proteins has been reported here or elsewhere, and FLC. GRP2 interacts with AGL15 via the structurally conserved first α-helix of the K-domain, thus it is unsurprising that GRP2 is also able to interact with other $MIKC^C$ MADS-domain proteins.

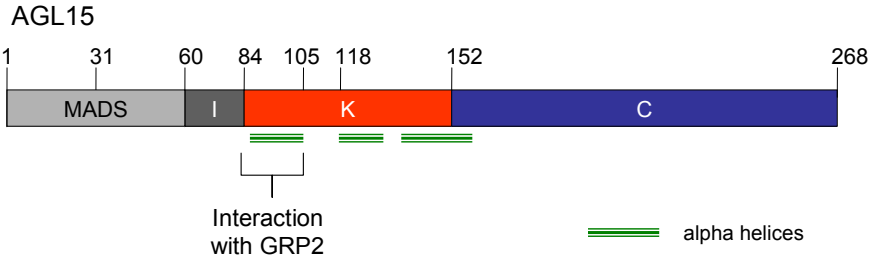
Figure 3.1 Elucidation of the regions of AGL15 that mediate protein-protein interactions with GRP2 in yeast two-hybrid assays

a.

AGL15 truncations in frame with GAL4 DBD:



b.

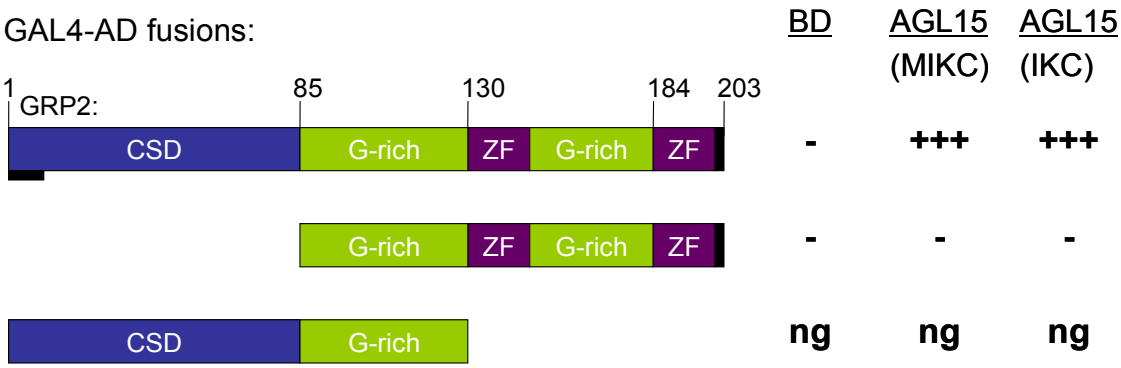


a. Yeast two-hybrid assays

Black bar, 10 aa; nd, not determined; +++, strong activation of reporter genes; ++, moderate activation of reporter genes; +, weak activation of reporter genes; +/-, very weak/transient activation of reporters genes; -, no activation of reporter genes (for a visual representation of how growth was scored see Figure 4.8). Results represent a minimum of three independent assays.

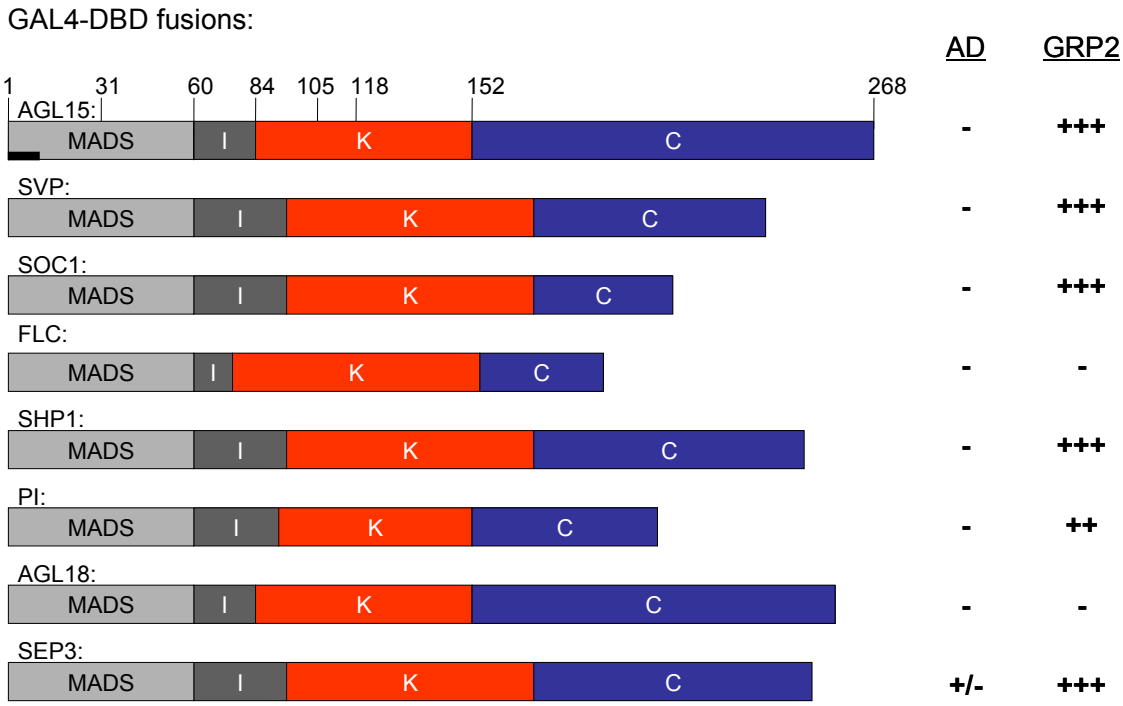
b. Schematic depicting the region of AGL15 that mediates an interaction with GRP2 in yeast two-hybrid assays

Figure 3.2 The CSD of GRP2 is required for interaction with AGL15 in yeast two-hybrid assays



Black bar, 10 aa; nd, not determined; +++, strong activation of reporter genes; ++, moderate activation of reporter genes; +, weak activation of reporter genes; +/-, very weak/transient activation of reporters genes; -, no activation of reporter genes (for a visual representation of how growth was scored see Figure 4.8). ng, no growth of transformed yeast cells on SD/-LW, indication either failed transformation or non-viability of transformed cells. Results represent a minimum of three independent assays.

Figure 3.3 AGL15 is able to interact with other MADS-domain proteins in yeast two-hybrid assays.



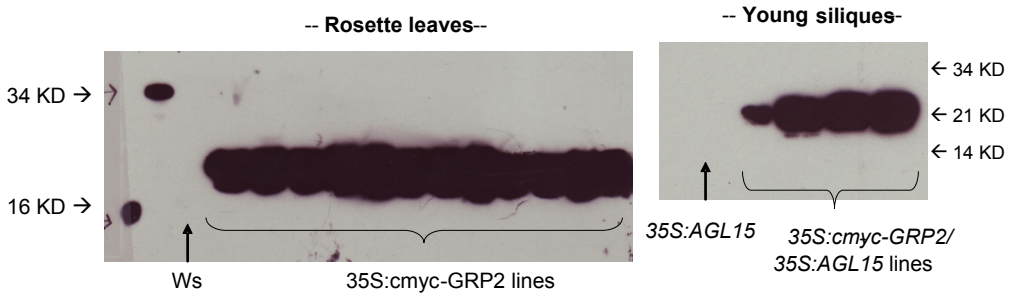
Black bar, 10 aa; nd, not determined; +++, strong activation of reporter genes; ++, moderate activation of reporter genes; +, weak activation of reporter genes; +/-, very weak/transient activation of reporters genes; -, no activation of reporter genes (for a visual representation of how growth was scored see Figure 4.8). Results represent a minimum of three independent assays.

3.2.2 GRP2 interacts with AGL15 *in vivo*

Transgenic plants, ectopically expressing an N-terminal c-myc tagged GRP2 fusion protein were generated, and lines accumulating high amounts of c-myc tagged protein crossed to *35S:AGL15* plants. C-myc protein accumulation can be readily detected in total extract from a range of tissues (seedlings, rosette leaves, and young siliques) isolated from *35S:c-myc-GRP2* transgenic plants (Figure 3.4). Protein markers of known molecular weights indicate that the migration of the c-myc-GRP2 protein through a polyacrylamide gel is consistent with that of a 22 KDa protein, the predicted size of myc-GRP2. Western analysis using a polyclonal antibody, raised against GRP2 (Fusaro *et al.,* 2007, generously donated by Dr. Gilberto Saccho-Martins, University of Rio de Janeiro, Brazil), also demonstrates much greater accumulation in transgenic *35S:myc-GRP2* seedlings compared to non-transgenic seedlings (data not shown). Myc-GRP2 protein can be detected in as little as 100 µg total extract. Despite high levels of myc protein accumulation, *GRP2* RNA transcript levels are only subtly elevated relative to non-transgenic controls (Figure 3.8a, and data not shown). GRP2 can be detected in nuclear extract and nuclear depleted samples (Fusaro *et al.,* 2007, K. Hill, data not shown).

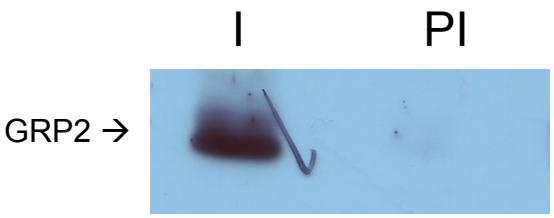
C-myc-GRP2 can be detected after immunoprecipitating nuclear extracts using AGL15 antiserum but not using the pre-immune control (Figure 3.5). Proteins from nuclear extract, isolated from embryonic tissue culture carrying both transgenes, *35S:AGL15* and *35S:c-myc-GRP2*, were immunoprecipitated with AGL15 antiserum and a pre-immune control serum. Eluted proteins were separated on a polyacrylamide gel and probed with GRP2 antiserum and anti-c-myc (Sigma-Aldrich, St. Louis, MO). C-myc-GRP2 protein can be detected with GRP2 antiserum (Figure 3.5) and anti-myc (data not shown) in the AGL15-immunoprecipitated fraction but not the pre-immune control. The experiment was repeated three times in total, and twice with an additional DNase treatment, which was included to eliminate the possibility of co-immunoprecipitation via interaction with adjacent chromatin regions. AGL15 could not be detected in the AGL15 immunoprecipitated fractions, and was barely detectable in the starting material. However, chromatin-immunoprecipitation experiments performed on the same tissue did show an obvious, albeit subtle, enrichment of AGL15 binding sites, suggesting that at least some protein was being precipitated.

Figure 3.4 C-myc-GRP2 protein accumulates in *35S:c-myc-GRP2* and *35S:c-myc-GRP2*/*35S:AGL15* lines



C-myc-GRP2 protein accumulates in total extract from leaves and siliques harvested from *35S:c-myc-GRP2* and *35S:c-myc-GRP2*/*35S:AGL15* lines. Protein extract from yeast expressing c-myc-tagged proteins of 34 and 16 KD were run as molecular weight markers (left hand lanes 1 and 2). Arrows on the right mark the positions of 34, 21, 14 KD markers (not shown).

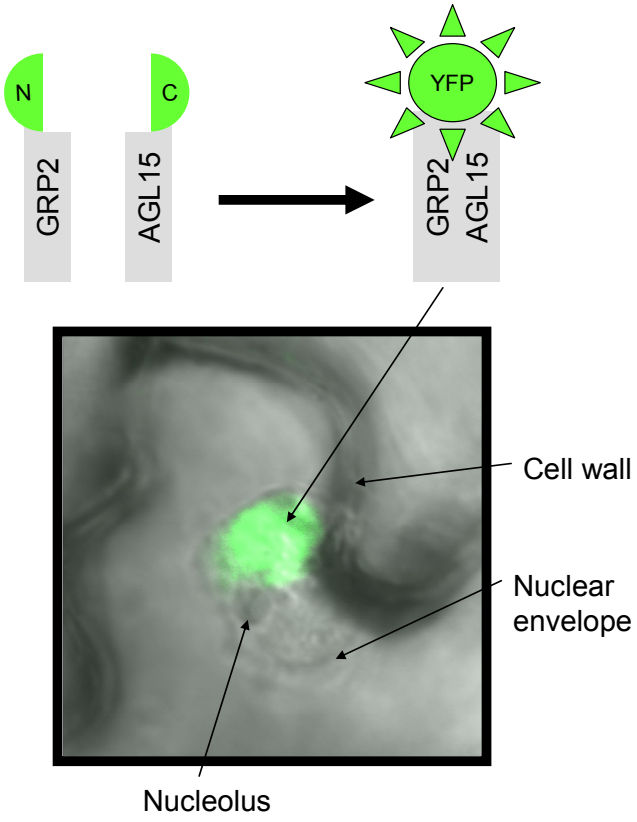
Figure 3.5 GRP2 co-immunoprecipitates with GRP2



Proteins from nuclear extract, derived from embryonic culture tissue carrying both the *35S:AGL15* and *35S:c-myc-GRP2* transgenes were immunoprecipitated with an AGL15 antibody (I) or with pre-immune serum as a negative control (PI). Immunopreciptated fractions were separated on a polyacrylamide gel and probed with a GRP2 antibody.

Another method to verify *an in vivo* interaction is bimolecular fluorescence complementation (BiFC). In this approach, fusion proteins containing either the N-terminal or C-terminal regions of yellow fluorescent protein (YFP) (Citovsky *et al*., 2006), in frame with either GRP2 or AGL15 respectably, were expressed in N. *benthamiana* leaves. Visualization via laser scanning confocal microscopy, using a FluoView^{TM} FV1000 Confocal Microscope (Olympus, Center Valley, PA) revealed that cYFP-GRP2 and nYFP-AGL15 interact in the nucleus of N. *benthamiana* leaves (Figure 3.6). However, although the cYFP-GRP2 and nYFP-AGL15 interaction was exclusively nuclear, signal throughout the cytoplasm could be detected when controls were performed using cYFP together with nYFP-AGL15 or nYFP together with cYFP-GRP2. Subsequent tests revealed that although no signal could be detected for either cYFP or nYFP alone, when the two halves were expressed together a strong signal was detected throughout the cell. This finding does cast some doubt over the claim that the YFP halves are being brought together in the nucleus via AGL15-GRP2 mediated interactions. However when cYFP or nYFP is co-expressed with nYFP-AGL15 or cYFP-GRP2, respectively, YFP signal can be detected throughout the cell, indicating that neither of the fusion proteins, when driven by the *35S* promoter is exclusively nuclear.

Figure 3.6 GRP2-nYFP and AGL15-cYFP interact in the nucleus

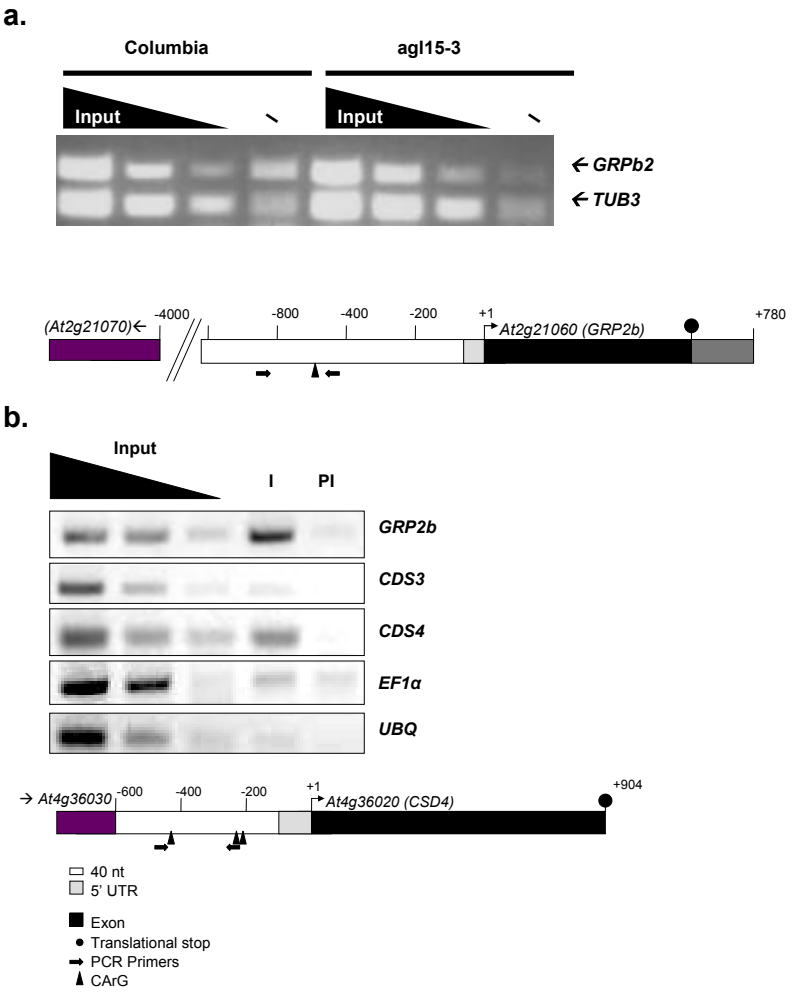


Agrobacteria were transformed with *35S:GRP2-nYFP* and *35S:AGL15-cYFP* constructs and co-infiltrated into *N. benthamiana* leaves. Laser scanning confocal microscopy, using a FluoView™ FV1000 Confocal Microscope (Olympus, Center Valley, PA), was employed to visualize fluorescence caused by the two the halves of YFP being brought together in close proximity.

3.2.3 AGL15 binds to the promoter regions of CSD genes *in vivo*.

GRP2 is one of four closely related genes present in the Arabidopsis genome, which all possess a CSD in conjunction with CCCH zinc finger motifs, interspersed by glycine rich regions (Karlson and Imai, 2003). AGL15 binds the promoter of *GRP2b* (At2g21060) *in vivo* (Figure 3.7a, b). A putative non-canonical $C(A/T)_8G$ AGL15 binding site resides in the *GRP2b* promoter at -654 to -646. Chromatin immunoprecipitation experiments were performed on B5 embryonic cultures (Ikeda *et al.,* 2002) derived from wild type Columbia seeds and seeds homozygous for null allele of AGL15, *agl15-3*. In Columbia, but not in the *agl15-3* negative control, a region of the *GRP2b* promoter (-510 to –980) was enriched in DNA populations that were immunoprecipitated with anti-AGL15, relative to a non AGL15-bound region (Figure 3.7a). AGL15 also binds to the promoter of *CSD4* (At4g36020) but not *CSD3* (At2g17870) in *35S:AGL15* embryonic tissue culture (Figure 3.7b). The relative enrichment of *GRP2b*, *CSD4*, and *CSD3*, compared to *EF1α*, to which AGL15 does not bind, is 2.13, 1.7, and 1.15 respectively, where 1 equals no enrichment. As a comparison, the relative enrichment of another negative control, *UBQ*, relative *to EF1α* is 1.1.

Figure 3.7 AGL15 binds to the promoter region of *GRP2b* and *CSD4* but not the *CSD3* promoter *in vivo*



a. Multiplex PCR performed on DNA immunoprecipitated with AGL15 from B5 cultures of Columbia and *agl15-3* genotypes. Oligonucleotide primer pairs to amplify regulatory regions of suspected direct targets of AGL15 and to amplify control regions not expected to be bound by AGL15 (*TUB3*) were used in multiplex PCR on total DNA diluted 125-, and 625-fold and on DNA recovered by immune precipitation (I) with AGL15 antiserum.
b. PCR performed on DNA immunoprecipitated with AGL15 and preimmune serum from *35S:AGL15* embryonic tissue culture. Oligonucleotide primer pairs to amplify regulatory regions of suspected direct targets of AGL15 and to amplify control regions not expected to be bound by AGL15 (*EF1α* and *UBQ*) were used in PCR on total DNA diluted 25-, 125-, and 625-fold and on DNA recovered by immune precipitation (I) or pre-immune precipitation (PI)

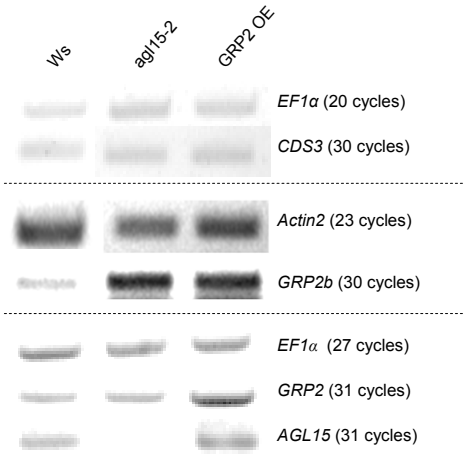
Experiments are representative of a minimum of three biological repeats.

3.2.4 *GRP2b* transcript accumulation is effected by both by GRP2 and AGL15

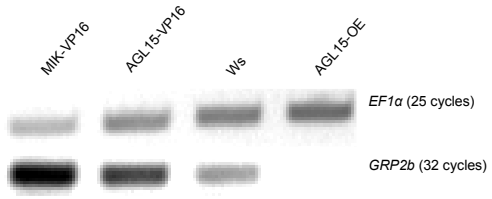
AGL15 transcript is not detected in *agl15-2* seedlings and *GRP2* transcript accumulates in seedlings ectopically expressing *GRP2*, although only moderately (Figure 3.8a). Despite the relatively low mRNA accumulation in *35S:c-myc-GRP2* seedlings, c-myc tagged protein accumulation can be readily detected in a range of tissues (Figure 3.4). Western analysis using a polyclonal antibody, raised against GRP2, also demonstrates much greater accumulation in *35S:c-myc-GRP2* lines compared to non-transgenic controls (Figure 3.4). Neither GRP2 nor *AGL15* transcript accumulation appears to be regulated by one another (Figure 3.8a, b). *GRP2b*, but not *CSD3*, transcript accumulates in *agl15-2* seedlings and in seedlings ectopically expressing GRP2 (Figure 3.8a). *GRP2b* transcript accumulation is reduced in seedlings ectopically expressing AGL15 (Figure 3.8b). *GRP2b* transcript also accumulates in seedlings ectopically expressing *AGL15-VP16*, and to an even greater extent in seedlings ectopically expressing a form of AGL15 where the C-terminal domain has been replaced by VP16 activation domain (*MIK-VP16*) (Figure 3.8b). Contrary to what was expected, *GRP2b* transcript accumulates in seedlings ectopically expressing GRP2 (Figure 3.8a).

Figure 3.8 RT-PCR on 8 day old seedlings

a.



b.



a. *GRP2b*, but not *CSD3*, transcript accumulates in *agl15-2* seedlings and in seedlings ectopically expressing GRP2. *AGL15* transcript is absent from *agl15-2* seedlings and *GRP2* transcript accumulates in seedlings ectopically expressing *GRP2*.

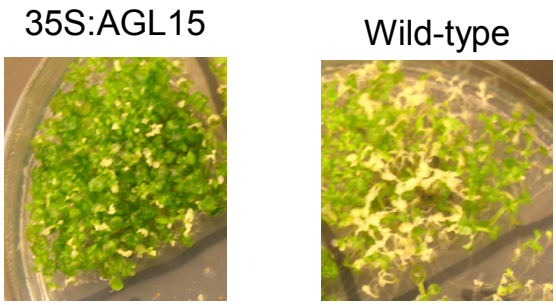
b. *GRP2b* transcript accumulates in seedlings ectopically expressing *AGL15-VP16*, and to an even greater extent in seedlings ectopically expressing a form of *AGL15* where the C-terminal domain has been replaced by VP16 activation domain (*MIK-VP16*). *GRP2b* transcript accumulation is reduced in seedlings ectopically expressing *AGL15*.

Experiments are representative of a minimum of three biological repeats

3.2.5 Plants ectopically expressing AGL15 have increased tolerance of freezing stress

Over expression of *Arabidopsis* RZ1a, a protein which like GRP2 contains zinc finger motifs interspersed by glycine-rich regions was found to confer freezing tolerance to plants and enhance the growth of *E. coli* exposed to cold temperature (Kim *et al.,* 2005). What is more, it has been demonstrated that it is the C-terminal half containing the zinc finger and glycine rich regions that is important for the growth-stimulating activity of *E. coli* under cold stress (Kim *et al.,* 2005). In similar experiments the induction of GRP2 had no effect on the cold survival of *E. coli* (K. Hill, unpublished data, Kim *et al.,* 2007). Likewise, over expression of *GRP2* had no effect on the survival rate of cold-shocked *Arabidopsis* plants (K. Hill, unpublished data, Kim *et al.,* 2007). Interestingly over-expression of *AGL15* did enhance the cold-shock survival rate of *Arabidopsis* plants in preliminary experiments (Figure 3.9). Seven day old seedlings, grown at 22ºC on GM media under long day conditions, were exposed to a temperature of -20°C for 40-120 minutes, and allowed to recover for another 7 days at 22ºC. Figure 3.9 is representative of several independent experiments demonstrating that the survival rate of *35S:AGL15* seedlings subjected to freezing temperatures is much greater than wild type seedlings grown concurrently. Because plants homozygous for the *35S:AGL15* transgene do not set seed, the seeds used in this study were collected from a *35S:AGL15* hemizygote plant. Therefore, one quarter of the seedlings in the *35S:AGL15* sector will not carry the transgene, and these could account for the seedlings that succumbed to the damage brought about by freezing temperatures.

Figure 3.9 Ectopic expression of AGL15 enhances the seedlings tolerance of freezing stress



Severn day old seedlings, grown at 22ºC under long day conditions, were exposed to a temperature of -20˚C for 40 minutes, and allowed to recover for another 7 days at 22ºC under long day conditions. *35S:AGL15* seeds were collected from hemizygous plants, therefore ~1/4 of the seedlings shown in *35S:AGL15* quadrant do not contain the transgene.

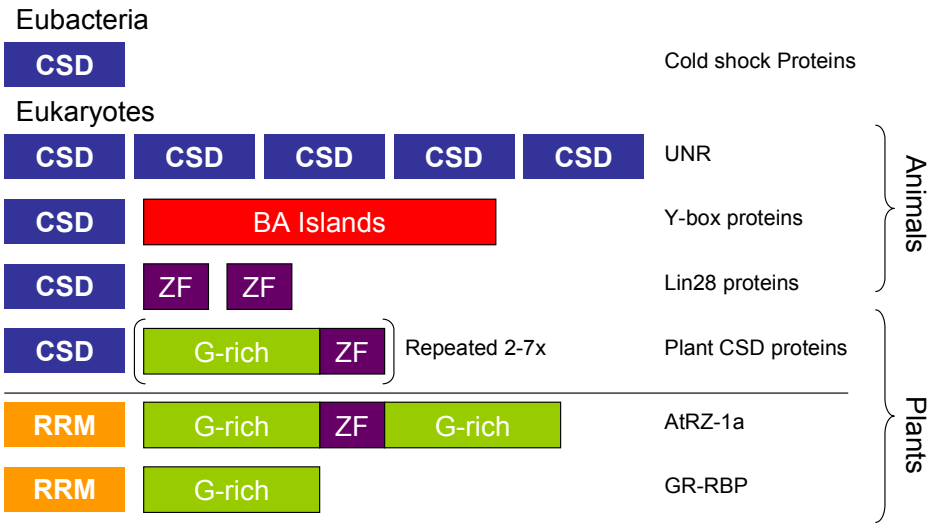
3.3 Discussion

3.3.1 Four genes coding for proteins containing a CSD in conjunction with zinc knuckle motifs, interspersed by glycine rich regions, are present in the *Arabidopsis* genome

The COLD-SHOCK DOMAIN (CSD) is the most conserved nucleic acid-binding sequence ever described, with greater than 40% identity and 60% similarity between bacteria and vertebrates (reviewed by Sommerville, 1999). The first cold-shock protein (CSP) described was CSPA from E.*coli* (Jones *et al.,* 1987), but CSD-containing proteins from a diverse array of organisms, including plants and animals have since been described (Figure 3.10a). Interestingly no CSD-containing protein is predicted to be encoded in the genomes of the yeasts *Saccharomyces cerevisiae* or *Schizosaccharomyces pombe* (Moss and Tang, 2003). Four CSD proteins are encoded in the *Arabidopsis* genome: GRP2 (At4g38680), GRP2b (At2g21060), CSD3 (At2g17870), and CSD4 (At4g36020). The first two have two zinc finger motifs whereas the later two have seven zinc finger motifs (Figure 3.10b). The retroviral-type CCHC zinc fingers, interspersed by glycine rich regions, are also found in conjunction with an RNA-recognition motif (RRM) in plants (Kim *et al.,* 2005, 2007a, 2007b).

The CSD consists of approximately 70-amino-acid residues that form a closed β-barrel structure with five β-strands with two β-sheets (reviewed by Yamanaka *et al,* 1998, Graumann and Marahiel, 1998). Located on the neighboring β2 and β3 are conserved RNA binding motifs, RNP1 (consensus K/NGY/FGFIE/T/NV/P/R) and RNP2 (consensus VFVHF), which are crucial for ssDNA and RNA binding (reviewed by Yamanaka *et al,* 1998, Graumann and Marahiel, 1998). These features are conserved in the four *Arabidopsis* CSD proteins (Figure 3.11). One notable difference between the plant GRP2 proteins is in the first β-strand (position 19 of AtGRP2), where the GRP2 proteins have either an aspartic acid or a serine and the rest of the cold shock domain proteins have an asparagine (Kingsley & Palis, 1994). Alignment of the four CSD proteins from Arabidopsis reveals that while 3 out of 4 have an aspartic acid or a serine residue at this position, one, CSD4, does have an asparagine (Figure 3.11).

Figure 3.10 CSD-containing proteins and plant proteins containing glycine-rich regions and zinc finger motifs

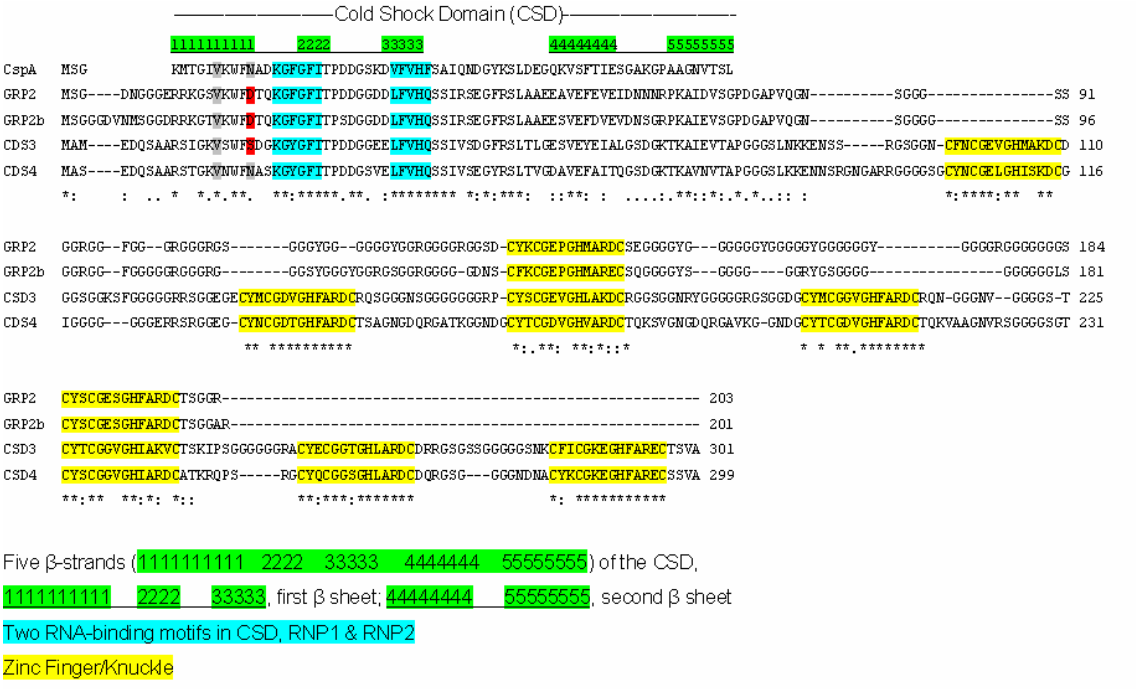


a. The CSD is found in proteins from bacteria, animals and plants and the G-rich, zinc finger motifs can be found in other plant RNA-binding proteins. Constructed using data reported in Fusaro *et al.,* 2007, Graumann and Marahiel, 1998, Karlson *et al*., 2002, Karlson and Imai, 2003, Kim *et al*., 2005, 2007a, b, Kingsley and Palis, 1994, Sommerville, 1999. CSD, cold shock domain; UNR; upstream of N-*ras*, a mammalian protein with 5 CSDs; BA, basic/aromatic islands; Zn, CCHC zinc fingers; G-rich, glycine-rich regions; RRM; RNA recognition motif; AtAZ-1a, an Arabidopsis protein containing zinc finger and glycine rich motifs in conjunction with and RNA recognition motif; GR-RNP, glycine-rich RNA binding protein.

Atg designation	Name used in this study	Other names	Schematic
At4g38680	GRP2[1,,3,4, 6]	CSD1[2] CSDP2[5]	CSD G-rich ZF X 2
At2g21060	GRP2b[1,,3]	CSD2[2]	CSD G-rich ZF X 2
At2g17870	CSD3[1,2]		CSD G-rich ZF X 7
At4g36020	CSD4[1,2]	CSDP4[5]	CSD G-rich ZF X 7
At4g13850	-	GRP2[5]	RRM G-rich

b. The *Arabidopsis* CSD-containing proteins have been assigned different names. *Arabidopsis* accession numbers and the corresponding names are given. 1, Thus study; 2, Karlson, D, personal correspondence; 3, Karlson and Imai, 2003; 4, Fusaro *et al.,* 2007; 5, Kim *et al.,* 2007a, b; 6, Kingsley & Palis, 1994.

Figure 3.11 Multiple alignment the four *Arabidopsis* CSD proteins and *E. coli* CSPA



Five β-strands (1.., 2.., 3…, 4.., and 5…) of the CSD are highlighted green. The two β sheets comprised of first three β strands and the fourth and fifth β strand are underlined. The RNA-binding motifs (RNP1 and RNP2) in the CSD are highlighted blue and the zinc fingers are highlighted yellow. The conserved valine in the first β-strand in highlighted grey, the conserved asparagine in the first β-strand of CspA is also highlighted grey, and the aspartic acid or serine in the corresponding position of the plant CSD proteins is highlighted red.

3.3.2 CSD-containing proteins can be found in a diverse array of organisms and are known to bind nucleic acids

Bacterial cold shock proteins (CSPs) bind to both single-stranded RNA and DNA and preferentially to ssDNA containing the sequence ATTGG as well as the complementary CCAAT sequence (reviewed by Yamanaka *et al,* 1998). The CSD of a Xenopus Y-box protein, FRGY2 preferentially binds the sequence AACAUCU in RNA (Bouvet *et al*., 1995). However, some CSD-containing animal proteins also bind specific elements in dsDNA. For example the CSD of the Y-box protein, YB-1 binds dsDNA and recognizes the Y-box (CTGATTGGCCAA), a sequence motif identical to that of an inverted CCAAT box (reviewed by Sommerville, 1999, Kohno *et al.,* 2003). The CSD of animal Y-box proteins are longer and more basic between the β3- and β4-strands than the bacterial CSD counterparts (Moss and Tang, 2003) and this appears to be important for dsDNA binding ability of the proteins. Replacing the six residue loop between the third and fourth β-strands of CSPA with the corresponding region of YB-1 results in a hybrid protein capable of binding dsDNA, but which retains the ability to bind ssDNA and RNA (Wang *et al.,* 2000). High salt concentrations abolish the dsDNA-binding capacity of the aforementioned hybrid protein, but RNA binding is unaffected (Wang *et al.,* 2000). The CSD of animal LIN-28 proteins, homologs of which can be found in a diverse range of animals, do not possess the extended region between the β3- and β4-strands and are closer to the bacterial CSD proteins than to that of the Y-box proteins (Moss and Tang, 2003). LIN-28 homologs all feature a cold shock domain (CSD) in conjunction with a pair of retroviral-type CCHC zinc knuckles (Moss and Tang, 2003).

Like the LIN-28 proteins, the four plant CSD genes also encode for CHHH zinc fingers. However, the CSDs of the plant proteins share greater sequence similarity with bacteria than they do with animal CSPs (Karlson *et al*., 2002). The CSD of the plant proteins have features of both Y-box and LIN-28 proteins. The plant CSD proteins, like the animal Lin-28 proteins, do not possess the extended basic regions between the third and fourth β-strands. However, the LIN-28 proteins possess a cysteine in the first beta strand of the CSD whereas animal Y-box proteins and the bacterial CSPs have a valine (Moss and Tang, 2003). Like the bacterial CSPs, the four *Arabidopsis* CSD-containing proteins also have a valine at this position (Figure 3.11) and this feature is appears to be conserved among all plant CSD proteins (Karslon and Imai, 2003). In fact plant CSD by itself, like the bacterial counterparts, appears not to bind dsDNA but does bind RNA and

ssDNA with high affinity (Karlson *et al*., 2002). However, the C-terminal zinc finger might render the whole protein capable of binding dsDNA, and perhaps even recognizing specific cis-elements. Indeed, a CSD-containing protein from wheat (WCSP1) does specifically bind dsDNA via a C-terminal glycine-rich and zinc-finger containing region, but not via the CSD (Karlson *et al*., 2002). AtRZ-1a, a protein which shares homology with the C-terminal zinc-finger/glycine rich regions of plant CSD-containing proteins is also able to bind dsDNA (Kim *et al.,* 2005).

A putative Y-box resides in GRP2b promoter (CTTAGTGGCCAA) – 966 to -952, which is in close proximity to the putative AGL15 binding site -654 to -646, and within a region enriched in AGL15-immunoprecipitated DNA. Several CCAAT and inverted CCAAT (ATTGG) motifs can also be found in the -1000 promoter and 3' UTR regions of GRP2b. CCAAT and ATTGG sequences can also be found within the CSD4 gene. Because protein-protein interactions between AGL15 and GRP2 have been observed one could hypothesize that AGL15 and GRP2 might associate together on the GRP2b and CSD4 promoter and work together to regulate the transcription of these genes. However, the presence of cis-motifs alone is not very informative. When allowing for one mismatch, over 90% of *Arabidopsis* genes contain a putative MADS-domain binding motif (CArG) within their regulatory regions (de Folter and Angenent, 2006). The CCAAT motif occurs even more frequently and is one of the most common elements in eukaryotic promoters, found in the forward or reverse orientation (Mantovani, 1998).

3.3.3 Transcriptional and translational regulation is mediated by CSD-containing proteins

Prior to fertilization AGL15 is cytoplasmic but becomes nuclear localized soon after the first embryogenic cell divisions (Perry *et al.,* 1996). GRP2 lacks any known nuclear localization or secretion signals and can be found in both nuclear and cytoplasmic factions (Fusaro *et al.,* 2007, K. Hill, unpublished data). In addition GFP-GRP2 can be detected in the cytoplasm and nucleus (Fusaro *et al.,* 2007). In nuclear extracts, GRP2 can be co-immunoprecipitated along with AGL15 (Figure 3.5) and preliminary data suggests that they directly interact exclusively in the nucleus (Figure 3.6). This is consistent with the hypothesis that they are working together to regulate transcription of downstream target genes.

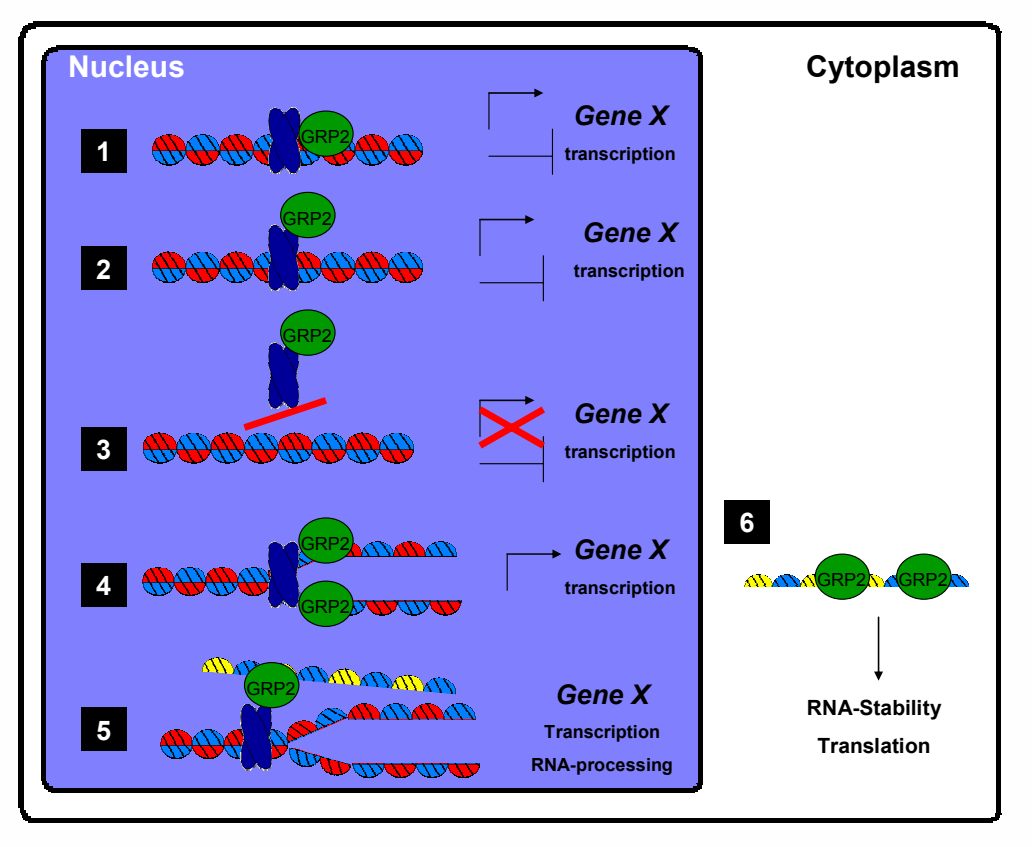
In *E.coli*, the cold shock protein CSPA acts as a transcriptional activator of at least two cold-shock genes, *hns* (La Teana *et al*., 1991) and *gyrA* (Jones *et al*., 1992). The mammalian CSD-containing protein, YB-1 also acts as transcriptional activator as well as a repressor. Three basic modes of operation have been hypothesized to explain how complexes containing YB-1 regulate gene expression (Kohno *et al.,* 2003):

1. YB-1 directly binds to Y-box alone or in association with other transcription factors.
2. YB-1 interacts with other transcription factors and functions as either co-activator or co-repressor.
3. YB-1 binds to the single-stranded region of the promoter either to enhance or inhibit the DNA binding of transcription factors.

A similar model could be used to describe how GRP2 might function together with AGL15 as a transcriptional activator or repressor in the nucleus. Figure 3.12 is a pictorial representation of six hypothesized and testable modes of action that might explain how GRP2 acts to affect transcription and/or translation of target genes.

In vitro dsDNA melting assays, the wheat CSD-containing protein, WCSP1, is able to melt dsDNA, an activity that was positively correlated to the ability to bind ssDNA (Nakaminami *et al.,* 2005). CSD4, but not GRP2, also demonstrates DNA melting activity (Kim *et al.,* 2007b). However, this does not preclude the possibility of GRP2 binding dsDNA or ssDNA. In fact, GRP2 has recently been shown to also be capable of binding nucleic acids (Fusaro *et al.,* 2007). A strong *in vitro* interaction with ssDNA and RNA, and a weak interaction with dsDNA has been reported for GRP2 (Fusaro *et al.,* 2007). The binding of GRP2 to dsDNA *in vivo* might be enhanced through combinatorial association with other proteins recruited to the promoter, such as AGL15.

Figure 3.12 Six models explaining how GRP2 might function together with, or independently of, AGL15 to regulate the transcription and/or translation of target genes



Blue dimers represent AGL15 (or other GRP2-interacting transcription factors); red and blue represents dsDNA and ssDNA; yellow and blue represents RNA.

1, GRP2 binds directly to DNA adjacent to AGL15 to affect transcription of a shared downstream target gene; 2, GRP2 binds to DNA-bound AGL15 to affect transcription of a downstream target gene; 3, GRP2 affects transcription of AGL15 downstream target genes by preventing AGL15 binding to and activating/repressing downstream target genes; 4, GRP2 affects transcription by binding ssDNA and facilitating recruitment of transcriptional machinery to the promoter; 5, GRP2 binds the nascent mRNA and is involved in transcript processing; 6, GRP2 plays a role in RNA-stability and/or translational regulation, independent of its interaction with AGL15 in the nucleus.

Ectopic expression of AGL15 results in reduced accumulation of *GRP2b* transcript and in *agl15-2* seedling an increased *GRP2b* transcript levels were observed (Figure 3.8a, b). Ectopic expression of *GRP2* also has an effect on *GRP2b* transcript accumulation, but its effect is contrary to what would be anticipated for a protein working together with AGL15 to regulate gene expression. *GRP2b* transcript levels decrease in response to increased AGL15 levels but are elevated in seedling ectopically expressing GRP2 (Figure 3.8a, b). However, GRP2 is able to interact with other MADS-domain proteins (Figure 3.3), including SEP3, which exhibits transactivation of reporter genes in yeast (Honma and Goto, 2001, supplementary data included in de Folter *et al.,* 2005, K. Hill, unpublished observation). The gene-expression studies performed herein measured steady-state mRNA levels as an indicator of the synthetic rates of transcript production. However, this approach fails to account for differences in mRNA stabilities or translational activation. The literature contains a number of examples whereby a CSD-containing proteins function to regulate translation. The bacterial CspA stimulates translation of cold-shock and cold-tolerant mRNAs at low temperature (Giuliodori *et al.,* 2004). The animal CSD-containing protein, FRGY2 also functions to control translation by masking RNA (Bouvet and Wolffe 1994, Ranjan *et al.,* 1993). It could be hypothesized that GRP2 plays a role in stabilizing mRNA in the cytoplasm independent of its interaction with AGL15 in the nucleus (Figure 3.12). It has been suggested that bacterial CSPs function as mRNA chaperones by destabilizing the over-stabilized secondary structures in mRNAs for efficient translation at low temperatures (Jiang *et al.,* 1997, Graumann and Marahiel, 1998). The addition of CSD4 but not GRP2 enhances the susceptibility of RNAs to RNase activity *in vitro* (Kim *et al.,* 2007b), suggesting that the latter is not acting to enhance translation by destabilizing secondary structures. However, this does not exclude the possibility that GRP2 binds and acts as a molecular chaperone, increasing the stability of mRNAs.

3.3.4 The regulation of CSD-genes by AGL15 may account for the enhanced freezing tolerance of *Arabidopsis* seedlings ectopically expressing AGL15

Over-expression of *GRP2* does not have any detectable effect on the survival rate of freezing stressed *Arabidopsis* plants (K. Hill, unpublished data, Kim *et al.,* 2007a). However over expression of AGL15 does appear to enhance the freezing stress survival rate of *Arabidopsis* plants (Figure 3.9). This is of particular interest because low

temperature is one of the most important environmental factors affecting plant growth. It limits the geographical distribution of plants and reduces the yield of some crops by shortening their growing season of many economically important crops (Mahajan and Tuteja, 2005). Freezing injury in plants results largely from severe cellular dehydration that occurs from ice formation in intercellular spaces. Since the chemical potential of ice is less than that of liquid water, there is a decrease in water potential outside the cell. Consequently, water moves from inside the cell to the intercellular spaces leading to cellular dehydration (recent reviews include Sharma *et al.,* 2005, Mahajan and Tuteja, 2005).

How might AGL15 confer freezing tolerance to the seedling? To addresses this question we must first ask what problems cold shock presents to the cell and how the cell compensates. In contrast to the heat shock, which induces factors required for protein folding (molecular chaperones) and protein degradation (Riezman, 2004), the two most urgent problems faced by cold-shocked cells are:

1. Decreased membrane fluidity, which hampers membrane-associated cellular functions, such as active transport and protein secretion, and is overcome by increased synthesis of unsaturated fatty acids and incorporation into membrane phospholipids (reviewed by Yamanaka *et al,* 1998).
2. Structures of RNA and DNA are stabilized by cold temperature, which is affects the efficiencies of mRNA translation, transcription and DNA replication (Polissi *et al.,* 2003). RNA chaperones may facilitate translation by blocking the formation of secondary structures in mRNAs (recent reviews include Yamanaka *et al,* 1998)

Downstream target genes of AGL15 might function to increase membrane fluidity or otherwise protect or aid recovery of the plant cell from the adverse affects of cold shock and/or injury. An alternative explanation is that morphology of *35S:AGL15* seedlings (shorter roots and hypocotyle, and broader, greener leaves) might render them more protected from the cold. Given the drastic nature of the phenotype, and the fact that the agar was frozen throughout this seems unlikely.

AGL15 may mediate freezing tolerance via downstream target genes other than the CSD-genes. One downstream target of AGL15, *CBF2* is repressed in response to AGL15 (Hill *et al.,* 2007, Chapter 2). The reported function of CBF proteins in cold tolerance (Gilmour *et al.,* 2004, Cook *et al.,* 2004, Vogel *et al.,* 2005) might appear contrary to the observed freezing tolerance of *35S:AGL15* seedlings. However, although over expression of *CBF2* enhances the cold tolerance of Arabidopsis plants (Gilmour *et al.,* 2004) *cbf2* mutant plants also show an increased tolerance to cold (Novillo *et al.,* 2004, Alonso-Blanco *et al.,* 2005). In addition *cbf2* mutant plants exhibited an increased tolerance to dehydration (Novillo *et al.,* 2004), and a consequence of freezing injury is severe cellular dehydration caused by ice formation in intercellular spaces (reviewed by Sharma *et al.,* 2005).

The fact that AGL15 can interact with and regulate the expression of CSD-containing proteins suggests another possible mechanism that might explain the enhanced freezing shock survival of seedlings ectopically expression AGL15. RNA molecules typically form stable secondary structures in response to low temperature (Polissi *et al.,* 2003). According to the current model (reviewed by Graumann and Marahiel, 1998), the role of CSPs is to prevent the formation of secondary structure, thereby keeping RNA in a linear form, which is a prerequisite for efficient initiation of translation in prokaryotes (Gualerzi and Pon, 1990). RNA stability is important in post-transcriptional gene expression. The *E.coli* cold-shock response appears to rely on pools of mRNA present at the time of cold shock, from which mRNAs (including CSP mRNAs) are preferentially translated (reviewed by Graumann and Marahiel, 1998). CSPA, itself a cold-induced protein, stimulates translation of cold-shock and the cold-tolerant mRNAs at low temperature (Giuliodori *et al.,* 2004).

AGL15 binds *in vivo* to the promoter of *GRP2b* and *CSD4*, but not *CSD3* (Figure 3.7b), and represses the transcription of *GRP2b*, but not *CSD3* (Figure 3.8a b). The effect of AGL15 levels on *CSD4* remains to be tested. Heterologous expression of wheat protein CSD-containing protein (Nakaminami *et al.,* 2006) and *Arabidopsis* CSD4, but not GRP2, is able to complement the cold sensitivity of mutant *E. coli* that lack four cold shock proteins (Kim *et al.,* 2007b). mRNAs corresponding to *GRP2*, *CDS3* , and *CDS4* (Karlson and Imai, 2003, Kim *et al.,* 2007b) increase in response to cold, whereas *GRP2b* mRNA decreases (Karlson and Imai, 2003). *GRP2b* is directly bound by AGL15

(Figure 4.7a, b) and transcript accumulation is reduced in response to increased AGL15 levels (Figure 4.8a, b). The effect of temperature on *AGL15* transcription has not been tested directly but according to Genevestigator® (Zimmermann *et al.,* 2004) *AGL15* transcript accumulation increases in response to cold temperature.

3.3.5 The interaction between GRP2 and AGL15 may be relevant in a developmental context

CspA has been identified as the major *E. coli* cold-shock protein whose production reaches more than 10% of total cellular protein synthesis upon temperature downshift from 37°C to 10°C (reviewed by Yamanaka *et al,* 1998). However, CSPs are defined on the basis of their conserved sequences and not all are induced by cold shock. In fact some are involved in other cellular processes. Of the nine CSPs present in *E. coli*, only three (CSPA, CSPB and CSPG) are cold inducible (reviewed by Graumann and Marahiel, 1998). CSPD is induced at the onset of stationary phase and during starvation, and appears to play a role in the nutrient-stress response, and CSPC and CSPE have been implicated in cell division and possibly condensation of the chromosome (reviewed by Graumann and Marahiel, 1998).

Although transcripts corresponding the *GRP2* increase in response to cold (Karlson and Imai, 2003, Kim *et al.,* 2007b), ectopic expression of GRP2 does not appear to enhance the survival of cold shocked seedlings (K. Hill, unpublished data; Kim *et al.,* 2007) or complement the cold sensitive phenotype of *E.coli* deficient in cold inducible proteins (Kim *et al.,* 2007). *AtGRP2* is preferentially expressed in meristematic regions and developing tissues that undergo cell division, namely meristems, carpels, anthers and embryos (Fusaro *et al.,* 2007). Down-regulation of *AtGRP2* gene, using gene-silencing techniques, produces plants that flower early, have altered stamen number and defective seed development (Fusaro *et al.,* 2007). Given the high sequence similarity between the four Arabidopsis CSD-containing proteins it is possible that the phenotypes reported by the aforementioned authors might be due to the silencing of more than one CSD-containing protein. However, it is worth mentioning that an early flowering phenotype is also observed in *agl15/agl18* double mutants (Adamczyk *et al.,* 2007) and plants ectopically expressing *AGL15* are delayed in flowering (Fernandez *et al.,* 2000). No late flowering was apparent in the *35S:GRP2* or *35S:c-myc-GRP2* lines generated in

this study (K. Hill, unpublished observation), however preliminary, unpublished data, discussed in Fusaro *et al.,* 2007, suggests higher levels of GRP2 might result in delayed flowering. AGL15 is believed to function in embryo development, thus a seed phenotype reported for an AGL15-interacting protein is particularly interesting. AGL15 accumulates in the nuclei of cells in the embryo beginning very early in development (by the eight-cell stage for Arabidopsis) and remains at relatively high levels throughout morphogenesis and into maturation stage (Perry *et al.*, 1996). *GRP2* promoter activity can also be observed in the cells of the globular through to torpedo stages embryos (Fusaro *et al.,* 2007). These observations are consistent more with GRP2 playing role in development, perhaps at least partially via its interaction with AGL15.

3.4 Materials and methods

3.4.1Yeast two-hybrid assays

The Clontech™ Matchmaker Library Construction and Screening Kit (Clontech, Mountain View, CA) was used, in accordance with manufacturers instructions, to screen a cDNA expression library, derived from *Arabidopsis* embryonic culture tissue (ECT) (Harding *et al*., 2003) for putative protein-protein interactions involving the MADS-domain transcription factor, AGL15 (At5g13790). For a detailed description see Chapter 4.

For directed tests AH109 cells were co-transformed with specific bait and prey constructs and plated onto SD media (2% dextrose, 0.67% nitrogen base, 1% agar) lacking leucine (-L) and tryptophan (-W), (SD-L/-W). After several days transformed AH109 yeast colonies were transferred onto to plates lacking histidine (-H), and alanine (-A), and supplemented with 0.2mg/ml X-α-gal (SD-L/-W/-H/-A, 0.2mg/ml X-α-gal) to assay for transformants that are able to activate the reporter genes (*HIS3*, *ADE2*, and *MEL1*). For a detailed description see Chapter 4.

Bait constructs used in this study:

DBD-AGL15 (MIKC), DBD-AGL15 (MIK½), DBD-AGL15 (IKC), DBD-AGL15 (KC), DBD-AGL15 (½KC), DBD-AGL15(C), DBD-AGL15 (IK), DBD-SVP (MIKC), DBD-SOC1 (MIKC), DBD-FLC (MIKC), DBD-SHP1 (MIKC), DBD-PI (MIKC), DBD-AGL18 (MIKC), and DBD-SEP3 (MIKC)

The oligonucleotides and restriction sites used generate the pGBKT7 (bait) constructs can be found in Appendix A.

All constructs were verified by DNA sequencing carried out by UK-AGTC. Protein accumulation in yeast cells was determined via Western blotting, using an AGL15 antibody (Heck *et al.,* 1995), c-myc monoclonal antibody (Clontech, Mountain View, CA).

Prey constructs used in this study:

AD-AGL15 (MIKC), AD-AGL15 (IKC), AD-GRP2 (FL), AD-GRP2 (CSD), and AD-GRP2 (Minus CSD)

The oligonucleotides and restriction sites used generate the pGADT7 (prey) constructs used in this study can be found in Appendix B.

3.4.2 Generation of transgenic plants

The plant expression vector, *pBIMC-35Sc-:myc-GRP2*, was generated from pDBD-GRP2 by PCR with oligonucleotides that amplify the c-myc tag and cDNA already cloned into the multiple cloning site of pGBKT7. Restriction enzyme cleavage sites, *Spe*I and *Sst*I*Sac*I (underlined), were added:

Forward 5'GACTAGTATGGAGGAGCAGAAGCTG'3
Reverse 5'CGAGCTCCCTCAAGACCCGTTTAG'3

The c-*myc-GRP2* cDNA was cloned into the binary expression vector, pBIMC (a gift from Dr. D. Falcone, University of Massachusetts) downstream of the *35S* CaMV promoter. The constructs were checked by sequencing and then transformed into the *Agrobacterium tumefaciens* strain GV3101, and used to transform Ws flower buds using the floral dip method (Clough and Bent, 1998). Transgenic T1 seeds were selected on GM plates containing 50ug/ml Kanamycin. *GRP2* transcript levels of putative cmyc-GRP2 over-expressing lines were analyzed by RT-PCR, and protein accumulation verified by Western analysis using an anti-myc monoclonal antibody (Sigma-Aldrich, St. Louis, MO) or a polyclonal antibody raised to GRP2 (Fusaro *et al.,* 2007, generously donated by Professor Gilberto Saccho-Martins, University of Rio de Janeiro, Brazil).

Plants constitutively expressing *myc-GRP2* were cross pollinated to a previously described *35S:AGL15* line (Fernandez *et al.,* 2000, Harding *et al.,* 2003) to produce plants over-expressing both transgenes.

pBIMC-35S:gAGL15 (MIKC)-VP16 encodes a form of AGL15 containing the first three introns and an additional C-terminal viral VP16 (AA 413-490) transcription activation domain (Dalrymple *et al*., 1985) cloned into pBIMC downstream of the *35S* CaMV promoter. *pBIMC-35S:gAGL15 (MIK)-VP16* resembles the previously mentioned construct, only the C domain has been removed.

Plant growth conditions

Seeds of *A. thaliana* ecotype Wassilewskija (Ws) or Columbia (Col) were sterilized by 3-4 brief washes with 70% ethanol containing 0.1% Triton X-100, followed by two rinses with 95% ethanol, then allowed to dry in a sterile laminar flow hood before sprinkling on GM (germination media) plates containing Murashige-Skoog (MS) salts, vitamins, 1% (w/v) sucrose, 0.05% (w/v) MES, and 0.7% (w/v) agar, pH 5.8. Seeds were chilled for several days at 4°C before being grown at 20/18°C under a 16-h light/8-h dark regime. Seedlings were transplanted at 8–10 days to ProMix BX (Premier Brands, Inc., Riviere-du-Loup, Quebec, Canada) and grown in a Conviron growth chamber with fluorescent and incandescent lights set to approximately 200 µmol m^{-2} sec^{-1}.

3.4.3 Embryonic tissue cultures

Generation of *Arabidopsis* embryonic culture tissue (ECT)

35S:AGL15 enhances induction of somatic embryogenesis and maintenance of development in the embryonic mode (Harding *et al.,* 2003). In order to obtain large quantities of *Arabidopsis* embryonic tissue, embryonic culture tissue (ECT) was generated as described by Harding *et al.,* 2003. Briefly, developing zygotic embryos from *35S:AGL15* lines were removed, wounded, and placed on GM plates, containing Murashige-Skoog (MS) salts, vitamins, 1% (w/v) sucrose, 0.05% MES(w/v), and 0.7% agar (w/v), pH 5.6 5.7. 50 µg /ml Kanamycin was included to ensure selection of transgene containing embryos from hemizgous plants. Secondary embryonic tissue, which develops on the cultured zygotic embryos, was sub-cultured at regular intervals of approximately 3 weeks on GM media plates.

In order to obtain sufficient quantities of non-transgenic embryo culture the B5 somatic embryogenesis system was employed as described by Ikeda *et al.,* 2002. In this system green siliques were surface sterilized and immature zygotic embryos isolated and placed on agar-solidified B5 (Gamborg's B-5 Basal with Minimal Organics; CAISSON Laboratories Inc., North Logan, UT) supplemented with 20g/L sucrose and containing 4.5 µM 2,4-D. After two weeks primary somatic embryos were placed into liquid B5 medium containing 9.0 µM 2,4-D. This step induces embryogenic cell clusters. To induce morphologically differentiated somatic embryos, 2-week-old cultures were washed with phytohormone-free liquid B5 medium and transferred to phytohormone-free liquid B5 medium for 1 week.

3.4.4 Co-immunoprecipitation

Plant tissue (5-10g) was fixed in approximately 5ml MC buffer (10 mM potassium phosphate, pH 7.0, 50 mM NaCl and 0.1 M sucrose) per gram of tissue. Protein-protein and protein-DNA complexes were stabilized by addition of 1% formaldehyde and incubation on ice under vacuum. After 1 hr the cross-linking was stopped by the addition of cold glycine to a final concentration of 0.125 M and incubated for a further 10-30 minutes before being thoroughly washed in MC buffer and flash frozen. Nuclei were isolated from fixed tissue as described by Bowler *et al.,* 2004 (for a comprehensive description of the nuclei isolation protocol see Chapter 2). The nuclear pellet was suspended in 0.6 ml of sonication buffer (10 mM potassium phosphate, pH 7, 0.1 M NaCl, 0.5% sarkosyl, 10 mM EDTA) with PMSF added to a final concentration 1mM immediately prior to sonication. The nuclei were sonicated for 10 to 15 second pulses using a probe sonicator (Fisher, Model 300 sonic dismembrator) set to half power. Sonication was repeated four times, interspersed by two minute incubations on ice. The insoluble material was removed by centrifugation at 12000 x g for 5 min at 4°C, and the supernatant transferred to a fresh 1.5ml Eppendorf tube. Prior to immunoprecipitation the DNA was removed so that only proteins in the same complex, and not those bound to adjacent chromatin regions, would co-immunoprecipitated. An equal volume of immunoprecipitation buffer (50 mM HEPES, pH 7.5, 150 mM KCl, 5 mM $MgCl_2$, 10 µM ZnSO4, 1% Triton X-100, 0.05% SDS) was added and a 50 µl aliquot taken as the "pre-DNAse treated" sample. To the remainder, 10 µl (100 units) of DNase (Invitrogen, Carlsad, CA) was added and the sample incubated at room temperature for 10 minutes.

A second aliquot of 50 µl was taken as the "post-DNase treated" sample. To both the pre- and post-DNase treated samples proteinase K was added to final concentration of 0.5 mg/ml and incubated overnight at 37 °C, followed by incubation at 65°C for at least 6 hours to reverse the formaldehyde crosslinks. DNA was extracted by phenol: chloroform extraction (phenol: chloroform: isoamyl alcohol at 25:24:1) and recovered by ethanol precipitation. The samples were electrophoresed on a 1% agarose gel to verify that DNA had been degraded. The DNase treated nuclear extract was divided into equal aliquots for immunoprecipitation with anti-AGL15 specific serum, and for preimmune serum and/or no serum controls, and 30ul of protein A-Sepharose beads (Invitrogen, Carlsbad, CA) added. The samples were incubated for two hours at 4°C with rotation. The beads were pelleted by spinning at top speed for 1-2 min. The supernatant was removed and saved as "post-bind" to check the depletion of the protein from the supernatant. The beads were washed with immunoprecipitation buffer (1 ml each tube) for 10 minutes at room temperature with rotation and pelleted by centrifugation at top speed for 1 min. The wash and centrifugation was repeated for 3-5 times. 100 µl of cold glycine elution buffer (0.1 M glycine, 0.5 M NaCl, 0.05% Tween-20, pH 2.8) was added to the beads. The sample was mixed by vortexing and pelleted by microcentrifugation at top speed for 1 min at room temperature. The supernatant was removed and added to a tube with 50 µl of 1 M Tris, pH 9 to neutralize. The elution and neutralization were repeated twice more to give a 450 µl total volume of the eluted sample. The eluted sample was centrifuged at top speed for 2 min at room temperature and transferred to a new eppendorf tube. Samples were boiled in 1X sample buffer for 5min before being immediately run on a gel or stored at -20ºC.

Proteins were separated on 12.5 % (w/v) polyacrylamide denaturing gels using vertical gel apparatus (Hoefer Scientific Instruments, San Francisco, CA) and then blotted onto Immobilon™ PVDF Transfer Membranes (Millipore, Billerica, MA) using a Genie blotter (Idea Scientific Co., Minneapolis, MN). Blots were blocked with 5% non-fat dry milk, washed in TBST (0.1% Tween 20, 100mM Tris-Cl, 150mM NaCl, pH 7) and probed with either 1:5,000 diluted GRP2 antiserum (Fusaro *et al.,* 2007, generously donated by Professor Gilberto Saccho-Martins, University of Rio de Janeiro, Brazil) or anti-myc (Sigma-Aldrich, St. Louis, MO). A 1:5,000 diluted secondary antibody (HRP-conjugated goat-anti-rabbit or goat-anti-mouse) and the Lumi-Glo System (Kirkegaard and Perry Laboratories, Gaithersburg, MD) were subsequently applied. Blots were exposed to

Kodak XAR5 X-ray films (Eastman Kodak, Rochester, NY) and the films developed in a Konica film processor (SRX-101, Konica Corp., Tokyo, Japan). Blots were exposed to X-ray film (Kodak XAR5) for 0.5-2 minutes.

3.4.5 Bimolecular fluorescence complementation (BiFC)

AGL15 and GRP2 cDNA were cloned into pSAT1-cYFP-1N, and pSAT1-nYFP-N1 respectively (Citovsky *et al*., 2006, generously provided by Dr. Stanton Gelvin, University of Purdue). *GRP2* cDNA was amplified with oligonucleotides containing *EcoR1* and *BamH1* restriction enzyme sites (underlined) and cloned into pSAT1-cYFP-1N, in frame with the c-terminal amino acids of yellow fluorescent protein (YFP):

Forward 5'CCGGAATTCTATGAGCGGAGACAACGGC'3
Reverse 5'CGCGGATCCCACGTCCAAC GCTGGTGC'3

AGL15 cDNA was amplified with oligonucleotides containing *BglII* and *BamH1* restriction enzyme sites (underlined) and cloned into pSAT1-cYFP-1N, in frame with the n-terminal amino acids of yellow fluorescent protein (YFP):

Forward 5'GAAGATCT ATG GGT CGT GGA AAA ATC GAG'3
Reverse 5'GCG GAT CC C AAC AGA GAA CCT TTG TC'3

35S:nYFP, 35S:cYFP, 35S:GRP2-nYFP and 35S:AGL15-cYFP were subsequently sub-cloned into pPZP-RCS2 (Goderis *et el.,* 2002; generously provided by Dr. Michael Goodin, University of Kentucky). All constructs were verified by DNA sequencing.

N. *benthamiana* plants were grown under greenhouse conditions and co-infiltrated with agrobacteria as described by Goodin *et al.,* 2002. Transformed GV3101 agrobacteria colonies were suspended in MES buffer (10 mM MgCl2, 10 mM MES, pH 5.6) and the OD600 adjusted to 0.6. Acetosyringone was added to a final concentration of 150 mM and the bacterial suspensions incubated at room temperature for 2-3 hrs. For co-infiltration of different Agrobacterium transformants, equal volumes of each culture suspension were mixed prior to infiltration. Infiltrations were conducted by nicking the epidermis and gently pressing a 1-ml disposable syringe to the abaxial surface of fully

expanded leaves that were approximately 2.5 cm wide at the midleaf and slowly depressing the plunger. Following agroinfiltration, plants were maintained in the laboratory under continuous fluorescent lighting for approximately 48 hrs. Laser scanning confocal microscopy, using a FluoView™ FV1000 Confocal Microscope (Olympus, Center Valley, PA), was employed to visualize fluorescence caused by the two the halves of YFP being brought together in close proximity.

3.4.6 Chromatin immunoprecipitation

Plant tissue (5-10g) was fixed in approximately 5ml MC buffer (10 mM potassium phosphate, pH 7.0, 50 mM NaCl and 0.1 M sucrose) per gram of tissue. To stabilize protein-protein interactions formaldehyde was added to a final concentration of 1% and the sample incubated on ice under vacuum. After 1 hr the crosslinking was stopped by the addition of cold glycine to a final concentration of 0.125 M and incubated for a further 10-30 minutes before being thoroughly washed in MC buffer and flash frozen. Nuclei were isolated from fixed tissue as described by Bowler *et al.,* 2004 and chromatin immunoprecipitation was performed as described in Wang *et al.,* 2004 using a polyclonal antibody raised against AGL15 (Heck *et al.*, 1995; Perry *et al.*, 1996, 1999, Wang *et al.,* 2000) or preimmune serum as a control. For a more detailed description see Chapter 2.

Multiplex PCR tests were performed on a series of dilutions of total (input) DNA, DNA recovered by immunoprecipitation (I) with an anti-AGL15 antibody, and the preimmune serum (PI). 30-35 cycles of PCR, with an annealing temperature of 52°C were performed using KlenTaq1 (Ab Peptides, St. Louis, MO).

Oligonucleotides specific for house-keeping genes, not believed to be bound by AGL15 were used as internal controls:

EF1α (At1g07920)
Forward 5'ACGCTCTACTTGCTTTCACC-3'
Reverse 5'GCACCGTTCCAATACCACC'3'
TUB2/3 (At5g62690 and AT5g62700)
Forward 5'GTCCTACTTTGTGGAGTGGA3'
Reverse 5'CTGTGTACCAATGCAAGAA'3

Oligonucleotides designed to amplify promoter regions potentially bound by AGL15 *in vivo* are as follows:

p*GRP2b* (At2g21060)
Forward 5'TCGTCATGTCATTTGGACTA'3
Forward 5'CTTCTCAGCCGTTAGATTCA'3
p*CSD3* (At2g17870)
Forward 5'CGTCATCATAATCCATGTGT'3
Reverse 5' CGAAGTTCAATTGTGAAGAA'3
p*CSD4* (At4g36020)
Forward 5'AGTAGTAACCAGCCGAATCGGGAA'3
Reverse 5'ATCCACGTCCGTTTATCAGCTGTG'3

PCR products were analyzed by agarose gel electrophoresis and gel images were captured using a ChemiImager (Alpha Innotech Corporation, San Leandro, CA). Relative Enrichment was calculated as followed:

$[I_{target}/Input_{target}]/[Input_{control}/Input_{control}]$ = Relative enrichment of target gene

I, the intensity of the PCR amplification band from immunoprecipitated DNA; Input, the intensity of the PCR amplification band from total, non- mmunoprecipitated DNA.

3.4.7 Semi-quantitative RT-PCR

Either TRIZOL Reagent (Invitrogen, Calsbad, CA) or the RNeasy® Plant Mini Kit (Qiagen, Valencia, CA) was used to isolate total RNA from ~50 mg of whole 6-8 day old seedlings, grown on GM media. 1.0 µg of total RNA was first treated with DNase I (Invitrogen, Calsbad, CA) and then used for first strand cDNA synthesis. Reverse transcription was performed using A-MLV Reverse Transcriptase System (Promega, Madison, WI) according to the manufacturer's instructions. 1-2 µl aliquot of each first strand cDNA reaction was amplified by specific primer pairs in a reaction containing 1x PCR buffer, dNTPs at 200 µM each, 0.2 µM of each primer and 1.0 unit of KlenTaq1 (Ab Peptides, St. Louis, MO) in a final volume of 20 µl. Amplification reactions were

performed in a PCT-100 (MJ Research Inc., Watertown, MA), under conditions that varied only in the number of cycles of denaturation at 95°C (30 sec), annealing at 55°C (30 sec), and extension at 72°C (30 sec). Control oligonucleotides specific for "house-keeping" genes were used as controls:

EF1α (At1g07920)
Forward 5'ACGCTCTACTTGCTTTCACC3'
Reverse 5'GCACCGTTCCAATACCACC3'
Actin2 (At3g18780)
Forward 5'GAGACCTTTAACTCTCCCGCTATG3'
Reverse 5'GAGGTAATCAGTAAGGTCACGTCC3'

Oligonucleotides specific to the respective target genes are as follows:

AGL15 (At5g13790)
Forward 5'TCCAAGAGGCGTTCTGGGTTACTT3'
Reverse 5'CTGCTCAAGGCTTTGCAGCTCTTT3'
GRP2 (At4g38680)
Forward 5'TGATACCCAGAAGGGTTTCGGCTT'3
Reverse 5' TCAGAACAGTCTCTCGCCATGTGA'3
GRP2b (At2g21060)
Forward 5'CTAGCGGTGGTGCTCGTTGA'3
Reverse 5'AACCAATCCAGTTTCTTTCTC'3
CSD3 (At2g17870)
Forward 5'GTTGCTTAACAAAAAGATGC'3
Reverse 5'TTCAAGTCATACAATCA AC'3

3.4.8 Freezing tolerance assays

7 day old seedlings, grown at 22ºC on GM media under long day conditions, were exposed to a temperature of -20°C for 40-120 minutes, and allowed to recover for another 7 days at 22ºC under long day conditions.

3.5 Summary

An interaction between the MADS-domain transcription factor AGAMOUS-LIKE 15 (AGL15; At5g13790) and the COLD SHOCK DOMAIN (CSD) –containing protein, GLYCINE RICH PROTEIN 2 (GRP2; At4g38680) is reported here. Preliminary data showing enhanced tolerance to freezing stress, conferred by seedlings carrying the *35S:AGL15* transgene, is presented here. AGL15 directly binds and regulates the expression of other CSD-containing proteins, which may act to enhance translation under cold conditions. Unlike other CSD-containing proteins, *35S:GRP2* does not appear to enhance the freezing tolerance of seedling and interaction between AGL15 and GRP2 maybe more developmentally relevant.

4.1 Introduction

The biological question addressed in this study concerns the identification and characterization of proteins able to interact with the MADS-domain transcription factor AGAMOUS-Like 15 (AGL15). There is a vast amount of literature pertaining to the modular nature of MADS-domain proteins and of the combinatorial nature of transcriptional regulation, thus a reasonable assumption is that AGL15 acts in conjunction with other proteins to regulate the expression patterns of various target genes. Yeast two-hybrid screening techniques were initially employed to ask the question, what proteins are able to interact with AGL15?

The yeast two-hybrid system was first described by Song and Fields, 1989, and has since become a routine tool for investigating protein-protein interactions. A review written five years ago made the claim that over 3000 research articles had been published, all of which used the yeast two-hybrid system to explore protein-protein interactions (Toby and Golemis,. 2001). Typing "yeast two-hybrid" into the Pubmed search engine (www.pubmed.gov) yields over 8000 research and 200 review articles, and the exact number rises almost daily. The most widely used yeast two-hybrid systems utilize the reconstitution of an active transcription factor to assay for protein-protein interactions. The most commonly used is the GAL4 system, which exploits the DNA-binding and activation domains of the yeast GAL4 protein (Fields and Song, 1989). While the ability to test whether or not two proteins are able to interact is a valuable tool in of itself, the extension of the technique to allow the screening of expression libraries has enabled the researcher to identify novel interaction partners in a relatively short and non-labor intensive fashion. The Clontech™ Matchmaker Library Construction and Screening Kit (Clontech, Mountain View, CA) was used to screen cDNA expression libraries derived from either *Brassica napus* (B.n) embryos or *Arabidopsis* embryonic culture tissue, using the MADS-factor AGAMOUS-LIKE 15 (AGL15) as bait.

This chapter aims to provide a comprehensive analysis of the yeast two-hybrid data with a view to applying this knowledge to future yeast two-hybrid studies, especially those involving MADS-domain proteins. One of the pitfalls of yeast two-hybrid screens is the high number of false positives, described herein. Indeed a considerable volume of recovered clones were either false positives or unlikely to interact with AGL15 in a

biologically relevant manor. However, a number of interesting and perhaps biologically important interactions were elucidated. In the preceding chapters two AGL15-interaction partners, SAP18 and GRP2, are described in detail. Other putative AGL15-interacting partners that warrant further investigation are presented herein.

4.2 Results

4.2.1 Analysis of yeast two-hybrid screens

Expression libraries, derived from either *Brassica napus* (B.n) embryos or *Arabidopsis* embryonic culture tissue (ECT), were screened for clones coding for polypeptides that are able to interact with AGL15. A total of 12 independent screens were performed using one of two methods; co-transformation of AH109 yeast cells with bait and prey together or mating of independently transformed AH109 and Y187 strains. Either full length or truncated forms of AGL15 were used as bait as indicated (Table 4.1). The number of colonies screened varied greatly, ranging from $2X10^3$ to $2.2X10^7$, and the number of colonies able to activate all three reporter genes (*His^+, Ade^+, and Mel1*) also varied considerably (Table 4.1). The number colonies able to activate the reporter genes, relative to the number of colonies screened was obviously much lower for screens performed with the AGL15-bait that included the MADS domain (Table 4.1; compare screen 2 to screen 3). Liquid cultures of AH109 yeast cells, transformed with bait plasmids encoding for full length AGL15, pDBD-AGL15 (MIKC), grow slower than those transformed with plasmids containing a form of AGL15 minus the MADS domain, pBDB-AGL15 (IKC), and both grow noticeably slower than those carrying just the empty bait vector or the pDBD-Lamin control (data not shown). However, yeast transformed with pDBD-AGL15 (MIKC) show no noticeable growth retardation when plated on solid media compared to those transformed with other forms of AGL15, empty vectors, or pDBD-Lamin (data not shown).

Table 4.1 A comprehensive table of all yeast two-hybrid screens performed

Screen	Bait AGL15:	Library Oligos	Method Temp.	Colonies screened	His+ (Ade+Mel1+) Colonies	His+ Ade+Mel1+ colonies	*Seq.	Unique Clones
1	MIKC	B.n emb. Ran + dT	Co-trans 30°C	$1.5X10^5$	54 (33)	1	11	11
2	IKC	ECT Ran + dT	Co-trans 30°C	$8.9X10^4$	130 (120)	6	67	67
3	MIKC	ECT Ran + dT	Co-trans 30°C	$3.1X10^4$	6 (5)	1	4	4
4	IKC	ECT dT	Co-trans 22°C	$7X10^3$	7 (7)	6	5	4
5	IKC	ECT dT	Co-trans 22°C	$2X10^3$	54 (10)	26	6	6
6	IKC	B.n emb. Ran + dT	Mate 28°C	$2X10^7$	>1000 (500+)	nd	77	5
7	IKC	ECT dT	Mate 28°C	$6.1X10^6$	200 (140)	nd	34	23
8	MI	ECT dT	Mate 28°C	$8X10^6$	nd	6	6	2
9	C	ECT dT	Mate 28°C	$2.2X10^7$	nd	150	10	4
10	IKC	ECT dT	Mate 28°C	nd	nd (128)	nd	79	32
11	IKC	ECT Ran + dT	Co-trans 28°C	$5.5X10^4$	16 (16)	15	45	21
12	IKC	B.n emb dT	Co-trans 28°C	$8.4X10^3$	66 (56)	nd	4	3

Screen, designated number of the independent screens; Bait; the domains of AGL15 that were cloned into pDBD (see appendix A1); The libraries screened were derived either from Brassica napus embryos (B.n emb.) or embryonic culture tissue (ECT); Ran, random oligos were used in library construction; dT, oligo dT was used in library construction; Method: Mate, mating method; Co-trans, co-transformation method; Temp., the temperature transformed colonies were incubated at; His^+, the number of newly transformed colonies that grew on SD/-LWH; (xx) number of colonies able to continue growing when transferred to SD/-LTHW X-α-gal; $His^+Ade^+Mel^+$, newly transformed colonies able to grow on SD/-LTHW X-α-gal; nd, not determined; Seq, number of sequenced. *, Only clones recovered from His+ Ade+Mel1+colonies were sequenced; Unique Clones, the number of clones that, although they may contain sequences coding for the same protein, are not replicates of the same prey plasmid.

A number of clones contained sequences encoding for the same protein, and the number of unique clones encoding for a particular protein, is indicated by roman numbers superscripted above the number assigned to the screen that they were obtained from (Tables 4.2, Appendix C-E). Because it is impossible to discern if two clones containing the exact same sequence were independently recombined or are replication of the same, they were recorded as duplicated clones. When two clones possessed differences in the length or regions of inserted cDNA they were recorded as unique clones derived from independent recombination events. Many of the clones recovered from the yeast two-hybrid screens performed using the mating protocol appeared to duplications of a single progenitor (Table 4.1; Screen 6, 7, 8, 9, 10). These repeat clones were frequently obtained from colonies growing close by on the same plate, suggesting that they were replicated from a single progenitor (data not shown).

Table 4.2 Recovered clones that require AGL15 for activation of reporter genes

Gene	Name	Description	Class	AA	Screens
At5g13790	AGL15	AGL15	A		1^{II}, 2^{IV}, 3^{II}, 7, 10^{VII}, 11
At4g38680	GRP2	Cold shock domain, glycine rich regions and 2 zinc fingers	A	7-203	1, 2^{V}, 4
At2g45640	SAP18	Member of the SIN3/HDAC1 complex	A	1-152	2^{II}
At3g62300	Agenet	Agenet domain-containing	A	263-722	4
At3g12130	KH	K-homology domain and zinc finger (CCCH type)	A	143-248	7
At3g03260	HDG8	homeobox-leucine zipper/lipid-binding START domain	A	568-699	5
At5g49450	bZIP1	bZIP family transcription factor	A	33-145	9^{III}
At1g28520*	AtVOZ1	VASCULAR PLANT ONE ZINC FINGER 1 transcription factor	A	319-486	2
At5g09740*	HAG5	MYST-like histone acetyltransferase	A	1-103 + intron 1	2
At1g67100	LOB40	LATERAL ORGAN BOUNDARIES (LOB) DNA-binding domain	A	73-233	7
At1g29990	Prefoldin	Prefoldin β-domain	C	1-129	2^{II}, 10
At2g01710	DNAJ	DNAJ heat shock N-terminal domain	C	1-206	2, 10
At5g02480	SLT1	SODIUM AND LITHIUM TOLERANT 1: molecular chaperone	C	451-508	7
At5g06460	UBA2	Ubiquitin activating (E1) enzyme	C	867-1077	2
At3g17365	SDM	SAM-dependent Methlytransferase domain	E	1-103	2
At5g43850	ARD4	1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase	E	34-240	2^{II}
At5g20250*	DIN10	DARK INDUCIBLE 10: raffinose synthase	E	623-749	2, 7
At1g29980	-	DUF642 domain: Unknown	X	301-371	7

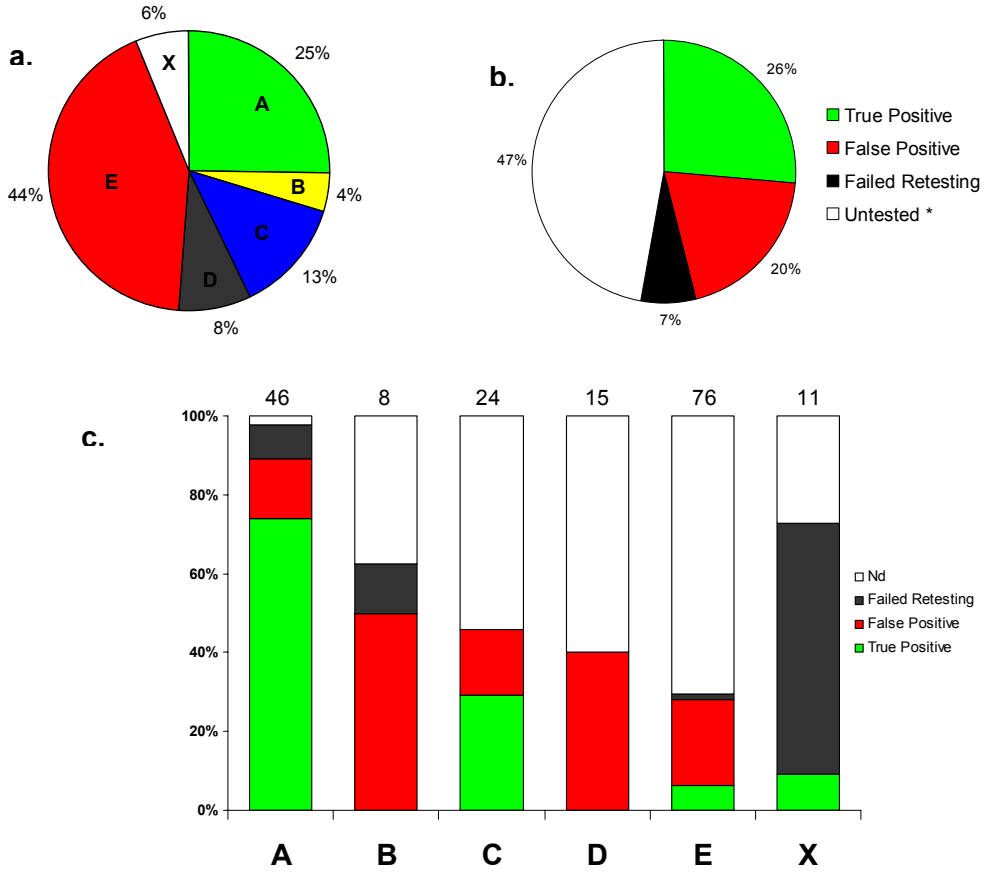
Each putative AGL15-interacting protein listed in Table 4.3 has been retested four or more times. *Transient false positive or inconsistent results; AA, minimum amino acids encoded by the shortest recovered clone/s; Screens, the number designated to the respective independent screen/s said clone was recovered from (see Table 4.1); Roman

numerals above the screen number denote the number of unique clones coding for the same protein that were recovered from that screen. Based upon the known or inferred functions of conserved domains (Marchler-Bauer *et al.,* 2007), or predicted sub-cellular localization (Emanuelsson *et al*., 2000), sequenced clones were categorized into the following classes: A, Transcription factors, chromatin remodeling factors, RNA-binding proteins; B, Cytosolic proteins with a possible role in posttranslational modification or signaling; C, Chaperones, protein folding, degradation, proteases; D, Secreted and structural proteins; E, Ribosomal, chloroplastic, mitochondrial proteins, metabolic enzymes; X, unclassified.

Sequences were obtained for over 300 prey plasmids, which were recovered from colonies able to activate all 3 reporter genes (*His*$^+$*, Ade*$^+$*, and Mel1*$^+$*).* Of these, 182 constituted unique clones (Table 2.1). Based upon known or inferred functions of conserved domains (Marchler-Bauer *et al.,* 2007), or predicted sub-cellular localization (Emanuelsson *et al*., 2000), sequenced clones were categorized into the following classes: A, Transcription factors, chromatin remodeling factors, RNA-binding proteins; B, Cytosolic proteins with a possible role in posttranslational modification or signaling, C; Chaperons, protein folding, degradation, proteases; D, Secreted and structural proteins; E, Ribosomal, chloroplastic, mitochondrial proteins, metabolic enzymes; X, unclassified. Fifty-two percent of recovered clones are likely to be secreted or structural proteins (D), metabolic enzymes, chloroplastic, or ribosomal proteins (E) (Figure 5.1a). AGL15, which is a transcription factor, is unlikely to reside in the same sub-cellular location as some of these, and the majority are probably false positives or biologically irrelevant. Common false positives include ribosomal subunits, heat shock proteins, proteasome subunits and cytoskeletal components, among others (some recent reviews include Serebriiskii *et al.,* 2000, Causier and Davies, 2002). Hence most of the recovered clones that were characterized as metabolic enzymes, or predicted to encode chloroplastic, secreted, or ribosomal proteins (Figure 4.1a) were not retested to determine whether or not they were indeed false positives (Figure 4.1b). However, of those retested the majority of proteins categorized as metabolic enzymes, chloroplastic, ribosomal, secreted, or structural proteins were false positives (Figure 4.1c; Class D and E).

Twenty-five percent of the recovered clones were putative transcription factors or nucleic binding proteins (Figure 4.1a). The majority of these were verified as AGL15 dependent interactions (Figure 4.1c, Table 4.2). A significant number of recovered clones (7% of those retested), when re-transformed with AGL15-containing bait failed to turn on the reporter genes, despite multiple attempts (Figure 4.1b). A number of proteins were uncategorized due to the absence of conserved domains of known or inferred function (Marchler-Bauer *et al.,* 2007), recognizable sub-cellular localization motifs (Emanuelsson *et al*., 2000), or similarity in amino acid sequence to proteins of known function (http://www.ncbi.nlm.nih.gov/BLAST/). It is interesting to note that many of the proteins falling into this category failed to activate the reporter genes when retested (Figure 4.1c).

Figure 4.1 Clones recovered from yeast two-hybrid screens were categorized based on inferred biological function and retested to determine if they specifically interacted with AGL15.



a. Clones recovered from yeast two-hybrid screens were categorized based on inferred biological function. Based upon the known or inferred functions of conserved domains (Marchler-Bauer *et al.,* 2007), or predicted sub-cellular localization (Emanuelsson *et al*., 2000), sequenced clones were categorized into the following classes: Classes: A, Transcription factors, chromatin remodeling factors, RNA-binding proteins; B, Cytosolic proteins with a possible role in posttranslational modification or signaling, C; Chaperons, protein folding, degradation, proteases; D, Secreted and structural proteins; E, Ribosomal, chloroplastic, mitochondrial proteins, metabolic enzymes; X, unclassified. 100% = 182 unique clones

b. Clones recovered from yeast two-hybrid screens were retested to determine if they specifically interacted with AGL15. True Positives, require AGL15-contain bait and do not activate the reporter genes when co-transformed with DBD-GAL or DBD-Lamin (green); False Positives, activate the reporter genes when co-transformed with DBD-GAL or DBD-Lamin (red); Failed Retesting; failed to activate the reporter genes in the presence or absence of AGL15-containing bait (black); Not Tested* (white),* Most of those not tested are likely false positives (based on sequence similarity to reported false positive) or of biological insignificance (based on their known or inferred cellular localization or function).

c. The inferred biological function was correlated with specificity for AGL15 verses auto-activation. The number at the top of the columns indicates the how many clones comprise that category, 100%.

4.2.2 Novel proteins able to interact with AGL15 in yeast two-hybrid assays

Table 4.3 lists the recovered clones that are only able to activate the reporter genes when co-transformed with AGL15-containing bait and not with either the GAL4-DBD alone or bait containing an unrelated protein pDBD-Lamin. All recovered clones were retested a minimum of three times and an asterisk indicates inconsistent results. Where multiple clones were recovered the minimum region of the polypeptide required to mediate an interaction with AGL15 is indicated.

Interesting AGL15 interacting proteins, selected for further study include SAP18 (At2g45640), a member of the SIN3/HDAC1 histone deacetylase complex and GRP2 (At4g38680), a cold-shock domain containing protein, discussed further in Chapters 2 and 3 respectively. Other AGL15-interacting polypeptides, which have been confirmed as AGL15-dependent in yeast two-hybrid assays, include:

Proteins with putative enzymatic functions

Of the proteins predicted to encode metabolic enzymes most of those retested (17 out of 23) were false positives, but AGL15-dependent interactions were verified for a SAM dependent methyltransferase (At3g17365), a 1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase (At5g43850), and a raffinose synthase (At5g20250). Based on inferred roles as metabolic enzymes these interactions were not deemed as likely candidates for biologically significant interactions with a transcription factor *in planta*, and as such were not studied further.

Proteins potentially involved in folding and degradation pathways

DNAJ heat shock N-terminal domain protein (At2g01710), β-Prefoldin subunit (At1g29990), SALT AND LITHIUM TOLERANCE 1 (SLT1; At5g02480), a putative molecular chaperone, specifically interact with AGL15 in yeast. Although the above were tested as specific AGL15-interacting proteins they not further studied as they likely represent interactions mediated by incorrectly folded proteins or protein aggregations. UBIQUITIN ACTIVATING 1 (UBA1; At5g06460) encodes for an ubiquitin activating enzyme (E1) and specifically interacts with AGL15 in yeast, and although it has not been retested, UBA2 was also recovered (At2g30110). However, E1 enzymes activate ubiquitin and represent the first step in the ubiquitination pathway, which associate with

the E2 enzymes rather than specific targets (for review see Vierstra, 2003). It is perhaps worth noting, although they have not been re-tested, a number of putative proteases were recovered from the yeast two-hybrid screens. These include a trypsin inhibitor (At1g47540), a subtilisin-like serine protease (At3g14240), an aspartic proteinase (At1g11910), an aspartyl protease family protein (At1g62290), and pepsin A (At1g62290).

Transcription factors

Interesting transcription factors include VASCULAR PLANT ONE-ZINC FINGER 1 (VOZ1; At1g28520), which encodes a putative one zinc finger transcription factor, bZIP1 (At5g49450), a basic leucine zipper, HOMEODOMAIN GLABROUS 8 (HDG8) (At3g03260), A homeobox-leucine zipper protein, and LOB40 (At1g67100), which contains a conserved N-terminal DNA -binding LATERAL ORGAN BOUNDAIRES (LOB) domain. Multiple assays demonstrated that VOZ1 exhibits a weak and transient ability to activate the reporter genes in the absence of AGL15 and as a consequence were not followed up on. However, quantitative β-galactosidase yeast two-hybrid assays did demonstrate a significant increase in reporter gene activity when co-expressed with AGL15-containing bait relative to the empty bait control (data not show). Likewise, LOB40 is also able to moderately activate the reporter genes in yeast two-hybrid assays in the presence of the DBD-GAL4 domain alone or the DBD-Lamin control. However, there is clearly a strong induction of reporter genes when AGL15 is present in the bait, indicating that a specific interaction between AGL15 and LOB40 is occurring, and preliminary β-galactoside assays support this (data not shown). LOB41 (At3g02550), was also recovered from three independent screens, but because of its auto-activation of reporter genes was not further analyzed.

Putative chromatin remodelling factors

Proteins with inferred roles in chromatin remodeling include a member of the SWI-INDEPENDENT 3/HISTONE DEACETYLASE (SIN3/HDAC) complex (see Chapter 2), a putative histone acetyltransferase, HAG5 (At5g09740), and an Agenet domain-containing protein (At3g62300). A clone encoding for HISTONE 4 (H4) (At1g07820) was also recovered, but failed to activate the reporter genes when tested. Multiple assays demonstrated that the HAG5 clone, which encodes for the proline-rich first exon, exhibits a weak and transient ability to activate the reporter genes in the absence of AGL15

(Appendix C) and as a consequence was not studied further. AGL15 interacts with amino acids 263-722 of an Agenet domain containing protein (At3g62300). Three conserved Agenet domains are found between amino acids 98-145, 161-226, and 228-284. Another conserved domain, excusive to plants and often found in association with the Agenet domain, lies between amino acids 115-178 (Marchler-Bauer *et al.,* 2007).

Putative RNA binding proteins

A number of putative RNA-binding proteins were isolated from yeast two-hybrid screens. Those that specifically interact with AGL15 in yeast include, GLYCINE RICH PROTEIN 2 (GRP2; see Chapter 3) and a K-Homology (KH) domain/CCCH type zinc finger containing protein (At3g12130). Two more potentially interesting RNA binding proteins, which failed to activate the reporter genes when retested code for a RNA helicase, HELICASE IN VASCULAR TISSUE AND TAPETUM 1 (HVT1; At2g30800), and predicted RNA methylase (At4g28830). The later was recovered from two independent screens using full-length (MIKC) AGL15 and again with truncated (IKC) AGL15 as bait.

Proteins of unknown function

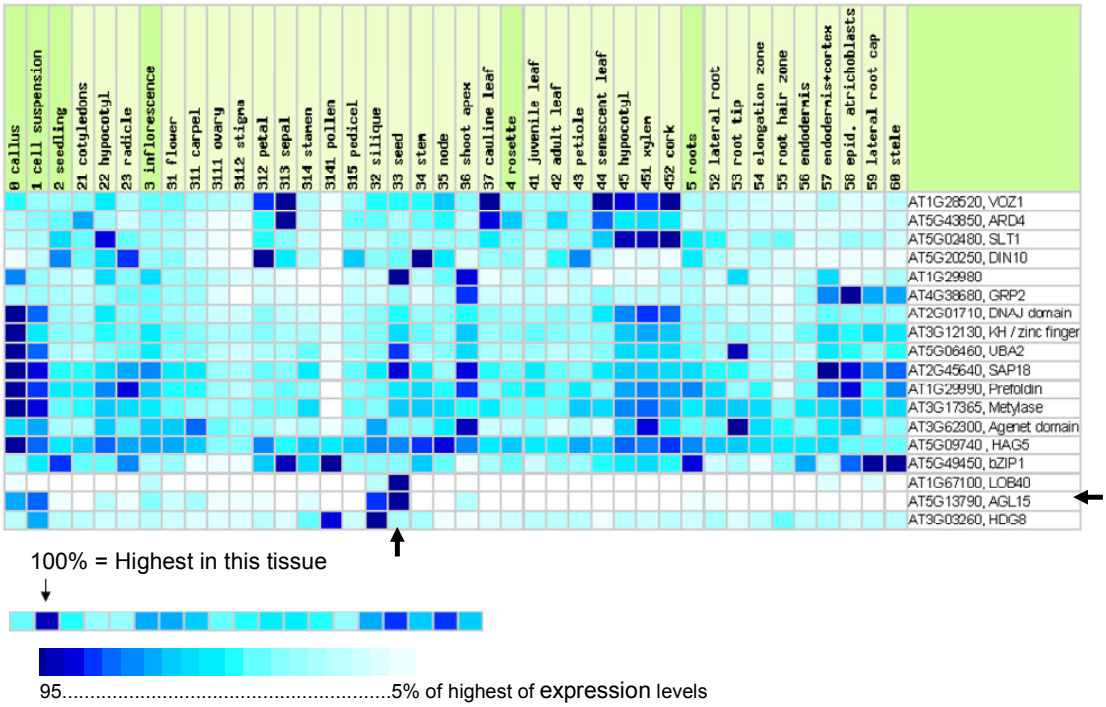
An unknown protein (At1g29980) that contains a conserved DUF642 domain, but bears no resemblance to any protein of known function was identified as an AGL15-interacting protein. TAPETUM DETERMINANT PROTEIN 1 (TD1) (At4g24973) was recovered from three independent screens, thus it was surprising that the interaction was never reproducible in directed tests (Appendix D). EARLY RESPONSIVE TO DEHYDRATION 15 (ERD15), a novel mediator of stress-related ABA signaling involved in freezing and pathogen resistance (Kariola *et al.,* 2006) was recovered as a putative AGL15-interacting protein. However, it too failed to activate the reporter genes when retested (Appendix D).

4.2.3 Genes coding for AGL15-Interacting proteins have overlapping expression patterns

All the AGL15-interacting clones were recovered from libraries derived from embryonic tissue. However, AGL15 expression is not limited to tissues developing in the embryonic mode (Heck, *et al.,* 1995, Fernandez *et al.*, 2000) and its ability to associate with other factors outside of the embryo may be of biological significance.

Transcripts corresponding to all AGL15-interacting partners were expressed in seeds and siliques, consistent with the fact that they were derived from embryonic expression libraries. The expression of *LOB40* and the unknown protein (*At1g29980*) is higher in seeds and siliques than in any other tissues. In fact, *LOB40* expression appears to be almost exclusively in the seed (Figure 4.2). *SAP18*, *UBA2*, *HDG8* and *HAG5* also show moderately level of expression in seeds and siliques. *AGL15* expression is highest in the embryo, but is not restricted to the embryo and is expressed in the shoot apex and floral buds (Heck, *et al.,* 1995, Fernandez *et al.,* 2000, Figure 4.2). Expression of all AGL15-interaction partners (Table 4.2) except *LOB40*, *bZIP1*, *DIN10*, and *HDG8* are also detected in the shoot apex. The unknown protein (*At1g29980*), the Agenet domain containing protein, *GRP2*, *SAP18*, and *HAG5* all show a relatively high expression in the shoot apex. Expression of all the AGL15-interacting partners are detected in the inflorescence, along with *AGL15* (Figure 4.2)

Figure 4.2 *In silico* mRNA expression patterns of AGL15-interactng proteins[1]



[1]Genvestigator ® (Zimmermann *et al.,* 2004).

Expression patterns determined using the Genvetigator® (version 2) microarray database and analysis toolbox (https://www.genevestigator.ethz.ch/at/). The darkest blue indicates the tissue in which the highest level of said transcript is found and expression in other tissues is displayed as a percentage of this (Zimmermann *et al.,* 2004).

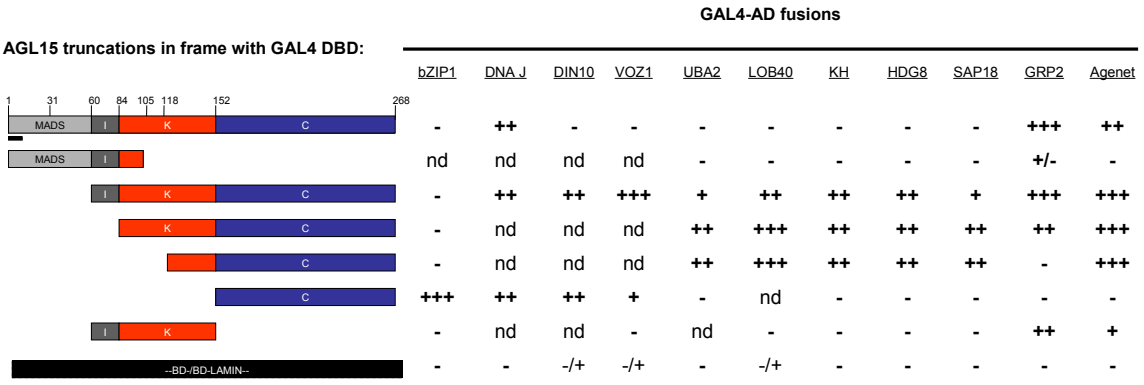
4.2.4 Elucidation of the Regions of AGL15 that Mediate Various Protein-Protein Interactions

Truncation tests were performed to determine the regions of AGL15 essential to its association with its various interaction partners. Two AGL15-interacting proteins, GRP2 and the Agenet domain containing protein, required only the K-domain (AA 80-152; Figure 4.3b), a region where the first two alpha helices of MIKC^c MADS-domain proteins are predicted to reside (Yang and Jack, 2004). Another group of proteins were able to interact with AGL15 solely through its C-terminal domain (AA 152-268; Figure 2.2b), VOZ1, bZIP1, DIN10, and the DNA J domain containing protein. A third group of proteins comprised of SAP18, LOB40, the KH-domain/CCCH type zinc finger protein, and HDG8, were unable to interact with the K-domain (AA 80-152) or C-domain (AA 152-268) alone. However, they were able to interact with the C-domain together with the later half of the K-domain (AA 118-268; Figure 4.3b), a region containing the second and third predicted helices of the K-domain (reviewed by Kaufmann *et al*., 2005).

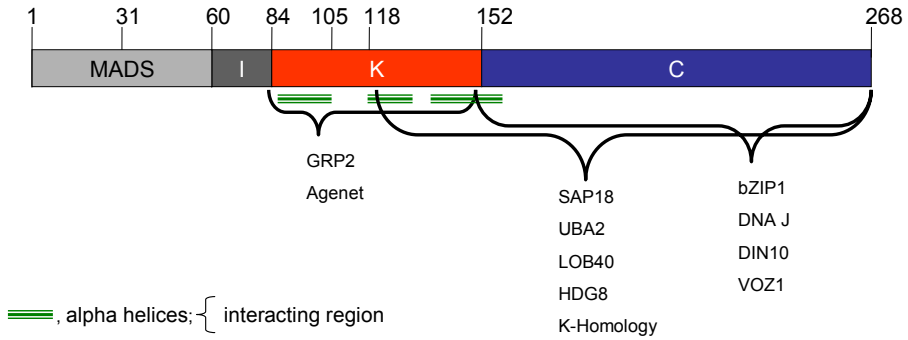
Those proteins whose interaction was mediated solely via the K domain (AA 80-152; Figure 2.3b), GRP2 and the Agenet domain containing protein, are able to interact with full-length AGL15, whereas none of the interactions requiring the K domain along together with C-domain interact with full-length AGL15-GAL4 fusion in yeast (Figure 4.3a). Three out of four of those proteins whose interaction is via the C-terminal domain alone (AA 152-268; Figure 4.3b) do not activate the reporter genes in the yeast cells when the MADS domain is present, the exception being the DNA J domain containing protein (Figure 4.3a).

Figure 4.3 Elucidation of the regions of AGL15 that mediate protein-protein interactions in yeast two-hybrid assays

a.



b.



a. Yeast two-hybrid assays

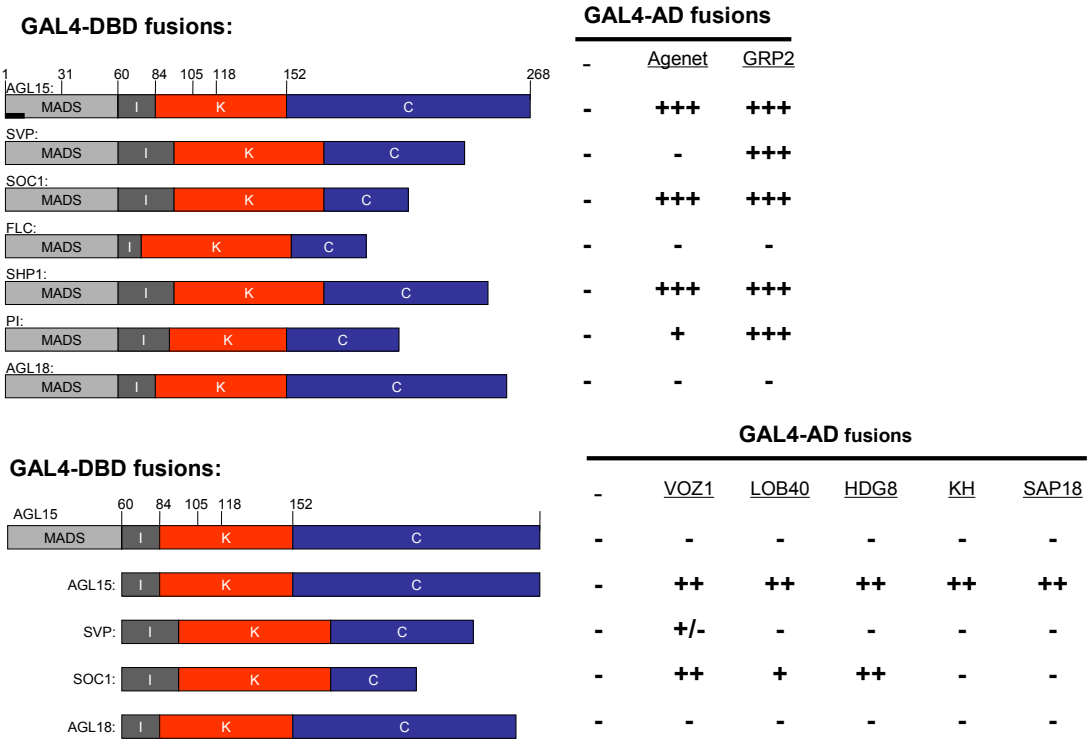
Black bar, 10 aa; nd, not determined; +++, strong activation of reporter genes; ++, moderate activation of reporter genes; +, weak activation of reporter genes; +/-, very weak/transient activation of reporters genes; -, no activation of reporter genes (for a visual representation of how growth was scored see Figure 4.8). GAL4-AD fusions; see Table 4.3 for accession numbers and amino acids encoded. Results represent a minimum of three independent assays.

b. Schematic depicting the regions of AGL15 that mediate protein-protein Interactions in yeast two-hybrid assays

4.2.5 Some, but not all AGL15 interacting proteins are able to associate with other MADS-domain proteins in yeast-2-hybrid assays

Yeast two-hybrid assays were performed using other MADS domain proteins as bait to address the question of whether the AGL15-interacting proteins were specific to AGL15 or able to interact with other MADS domain proteins. Not surprisingly, those proteins whose interaction with AGL15 involved the structurally conserved K domain (reviewed by Kaufmann *et al*., 2005), interacted with other MADS domain proteins, whereas those whose interaction was dependent on the divergent C-domain were more specific to AGL15 (Figures 4.3a, 4.3b, 4.4). None of the proteins recovered as AGL15-interacting partners were able to interact with AGL18, either full-length or truncations lacking the MADS or the MADS and I domains (pDBD-AGL18 (IKC) and pDBD-AGL18 (KC)). AGL18 does not interact with any of the AGL15-interacting proteins tested (Figure 4.4 and data not show) or any other MADS-domain proteins tested (Chapter 5), which might indicate it was not expressing or folding correctly, or is perhaps forming aggregates. However, Western analysis detected protein, corresponding to the predicted size of the GAL4-AGL18 accumulating in AGL18 transformed yeast. Unlike AGL15, AGL18 does not interact with the Prefoldin or DNAJ protein (data not shown).

Figure 4.4 Some AGL15-interacting partners can interact with other MADS-domain proteins in yeast two-hybrid assays



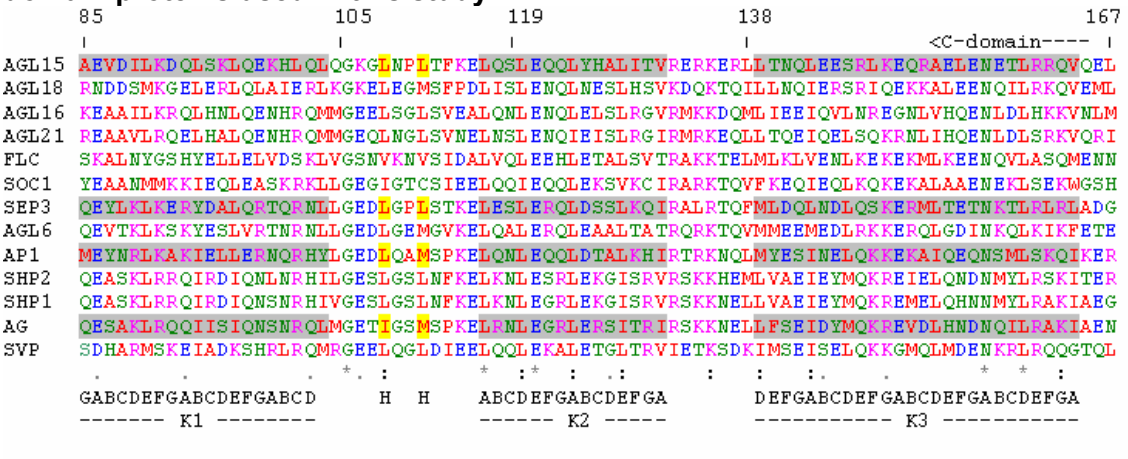
Black bar, 10 aa; nd, not determined; +++, strong activation of reporter genes; ++, moderate activation of reporter genes; +, weak activation of reporter genes; +/-, very weak/transient activation of reporters genes; -, no activation of reporter genes (for a visual representation of how growth was scored see Figure 4.8). GAL4-AD fusions; see Table 4.3 for accession numbers and amino acids encoded. Results represent a minimum of three independent assays.

4.2.6 Three putative α-helices reside in the regions of AGL15 that mediate protein-protein interactions

From the yeast two-hybrid truncation studies it is apparent that three distinct regions of AGL15 mediate interactions with other proteins. Some proteins can interact with the K-domain alone (AA 84-152), and some can interact with the C-domain (AA 152-268). A third group requires the later half of the K-domain along with the C-domain (AA 118-268).

The K domain of $MIKC^C$ MADS is predicted to form three α-Helices (Yang and Jack, 2004). The first two predicted helices lie within the K-domain of AGL15, as defined in this study, but the third spans into the C-domain (Figure 4.5). It could be hypothesized that those proteins that require the later half of the K-domain along with the C-domain (AA 118-268) might be interacting via the third, or second and third α-helices. Although the experiment has not been performed, it could be hypothesized that this group of interacting proteins only require amino acids 118 to 167 of AGL15.

Figure 4.5 Predicted positions of the three α-Helices of AGL15 and other MADS-domain proteins used in this study



Clustal W software (Chenna *et al*., 2003) was used to align AGL15 with other MADS-domain proteins.

Numbering refers to amino acid positions of AGL15. The positions of the three alpha helices of AG, SEP3, and AP1 as defined by Yang and Jack, 2004, along with the corresponding regions of AGL15 are highlighted grey. The conserved hydrophobic positions between K1 and K2 are highlighted yellow.

4.3 Discussion

4.3.1 Analysis of yeast two-hybrid screens performed

The number of colonies screened (Table 2.1) varied greatly, ranging from $2X10^3$ to $2.2X10^7$. Similar studies that used yeast two hybrid assays to screen for protein-interactions involving plant MADS-domain proteins ranged between $1.8\ X10^4$ to $3.9X10^9$ (Fan *et al*., 1997, Davis *et al*., 1999, Immink *et al*., 2002, Pelaz *et al*., 2001, Homo & Goto, 2001, Masiero *et al*., 2002, Moon *et al*., 2002, and others).

4.3.1.1 False positives

The large number of false positives is a pitfall of the yeast two-hybrid system (discussed by Serebriiskii *et al.,* 2000, Toby and Golemis, 2001, and many others). The Clontech™ Matchmaker Library Construction and Screening Kit (Clontech, Mountain View, CA) utilizes four reporter genes (*HIS3*, *ADE1*, *MEL1*, and *LACZ*) under the control of three distinct promoters (*pMEL1*, *pGAL1*, and *pGAL2*), which reduces the number of false positives compared to older systems (reviewed by Causier and Davies, 2002). However, false positives, as exemplified in Figure 2.1b, are still an issue. In theory there are two types of false positives: those that promiscuously bind chromatin and due to the GAL4-activation domain fusion activate the reporter genes when bound close by, and those that promiscuously interact with proteins in general, such as the GAL4-DBD domain itself. The former can be determined by transforming yeast with the AD-fusion alone and assaying for activation of the HIS^+, ADE^+ or *MEL1* reporter genes. A third type of false positive may include prey fusions that alter the metabolism of yeast cells, indirectly creating a bias towards activation of the reporter genes (discussed by Toby and Golemis, 2001). This might account for the high number of recovered clone coding for proteins with putative metabolic roles.

One obvious trend, made apparent in Table 4.1, is that screens using full-length AGL15 as bait yielded a much lower numbers of positive colonies relative to the number of clones screen. The reason for this became apparent when a good portion of the AGL15-interacting proteins, identified using the truncated form of AGL15 as bait failed to interact with full length AGL15 bait. It was also found that some false positives, isolated from

AGL15-IKC screens no longer activated the reporter genes when full-length AGL15 was cloned into the bait (data not shown). It would be interesting to determine if there was a correlation between the type of false positive and the ability of AGL15 to override it, especially given the function of AGL15 as a transcriptional repressor (see Chapter 2).

Twelve out of the ninety-six clones that were retested failed to activate the reporter genes (Appendix D). Many of these, because they were deemed interesting, were tested several times. Although the colonies were re-streaked on selective media 4-5 times in order to select against multiple clones not involved in activation of reporter genes, it is possible that some were still retained. This might be true if there was a selective pressure to retain two clones, such as an interaction mediated by three or more proteins. However, only one clone was ever recovered in each of these cases. If transformed cells are plated too densely problems can arise in interpreting reporter activity (reviewed by Causier and Davies, 2002). What is puzzling is that three out of the twelve were isolated from multiple screens.

4.3.1.2 Full length AGL15-GAL4 fusion protein might obscure one or more protein-protein interaction surface

It is interesting to note that those proteins whose interaction with AGL15 is mediated solely via the K domain (AA 80-152; Figure 4.3a), GRP2, the Agenet domain containing protein, and AGL15 itself, are able to interact with full-length AGL15, whereas those proteins requiring at least part of the C-terminal domain (AA 152-268), do not activate the reporter genes in the yeast cells when the MADS domain is present (Figure 4.4). AGL15 minus the MADS domain, but not full-length AGL15 is able to interact with SEP3 in the yeast system (Chapter 5). However In vitro co-immunoprecipitation experiments confirmed that full-length AGL15 and SEP3 were able to interact with each other in the absence of the GAL4 fusions (Chapter 5). What is more, when SAP18 is expressed as a DBD-fusion and full-length AGL15 as an AD-fusion an interaction is observed (Chapter 2). This suggests that GAL4 DBD fusions containing full length AGL15 are not correctly folded, and that perhaps only the C-terminal domain is affected. The fact that proteins predicted to function as part of the unfolded protein response (Marchler-Bauer *et al*., 2007) were recovered and verified as AGL15-dependent interacting proteins (Table 4.3) supports this hypothesis. The literature contains a number of examples where a full-

length MADS-domain protein, when fused to the GAL4-DBD are unable to interact with partners that the truncated forms is able to associate with (Yang *et al*., 2003a, Yang *et al*., 2003 b, Yang & Jack, 2004, Fujita *et al*., 2003) or where an interaction shown to occur between full-length MADS *in planta* does not occur in the yeast system (Immink *et al*., 2002), thus it is likely a conformation effect brought about by the artificial nature of the GAL4-DBD fusion. It is well known that use of protein fusions can cause the site of interaction to be occluded by one of the transcription factor domains (for review see Phizicky and Fields, 2005). This may perhaps affect only one of several interaction surfaces that are present in correctly folded AGL15. The literature contains examples of yeast two-hybrid interactions that were reported by another group as being negative. Indeed, interactions found to occur by this study (FLC-FLC, SVP-SVP, SHP1-FLC; Chapter 5) have been reported as negative (de Folter *et al.,* 2005). In addition some interactions such as that between FLC and SHP1 are only observed in one orientation (in this case when SHP1 is the bait). This might be due to the differences in bait constructs. The pGBT7 vector (Clontech, Mountain View, CA) used in this study contains a c-myc epitope tag between the GAL4-DBD and multiple cloning site, whereas the pDEST^{TM}32 vector (Invitrogen, Carlsbad, CA) used by de Folter *et al*., 2005 does not. Even using the same vector, extra amino acids between the GAL4 DBD and the fusion protein are encoded depending on the multiple cloning sites used in the construction of the bait vector. This might theoretically have a steric effect on the resultant fusion proteins.

4.3.1.3 Full length AGL15-GAL4 fusion protein may act as a repression domain in yeast two hybrid assays

An enigma related to our yeast two-hybrid studies has been the observation that full-length AGL15, although permissive to yeast growth, is able to override the activation of reporter genes normally activated by known false positives, isolated using AGL15 IKC as bait (data not shown). While this is not true of all false positives, it does hint that AGL15 might possess a transient ability to repress transcription, even in the presence of the GAL4 activation domain. The fact that a number of false positives appear not to activate reporter genes in the presence of bait containing full length AGL15 explains the reduced number of recovered clones from screens where full-length AGL15 was used as bait, despite comparable transformation efficiencies (Table 2.1; compare screen 1 to screen

2), and growth on plasmid selective media (SD/–LW). Another observation is that no AGL18-interacting partners have been demonstrated using the yeast system in our lab or have been reported in the literature, despite such experiments having been performed (de Folter *et al*., 2005, Verelst *et al*., 2007). However fusion protein accumulates in yeast cells (data not shown). This lack of interaction could be due to steric restraints imposed by the nature of the fusion protein, either obscuring the interaction or exposing a strong repression domain. The transcription factor OCT-2 contains an inhibitory domain which is able to override acidic activations domains, including the GAL4 activation domain, but has little or no effect on proline-rich or glutamine-rich activation domains (Liu *et al.,* 1996).

4.3.1.4 The limitations of yeast two-hybrid screening and suggestions of future strategies

Many interesting AGL15-interacting partners were isolated only once and from only one screen, thus it is likely that many interactions have been overlooked. Indeed, directed tests revealed interactions between MADS, whose transcripts were present in mRNA used to make the libraries screened (Chapter 5). Transcription factors tend to be low abundance transcripts. It was found that freshly transformed yeast, containing AGL15 bait and SEP3 prey do not grow when plated directly onto SD/-LWH. Neither will freshly transformed yeast containing AGL15 bait and SAP18 grow directly on SD/-LWA (data not shown). Thus it is likely that other interactions are not able to establish themselves enough to maintain survival of newly transformed cells. Therefore, it is likely that many AGL15-interacting factors were not detected in the yeast two-hybrid screens. Newer additions of the Matchmaker 3 Library Construction and Screening Protocols (Clontech, Mountain View, CA) suggest a third approach whereby newly transformed cells are first selected for on SD/-LW before being transferred to media selective for protein-protein interactions. This is especially recommended if the bait is slightly toxic, or slows the growth of yeast cells as is the case with AGL15.

A number of similar studies that have adopted yeast two-hybrid based strategies to screen for protein-interactions involving plant MADS-domain proteins have found more interaction partners when screens were performed at 22-25ºC rather than 30ºC (Pelaz *et al*., 2001, Honma & Goto, 2001, Shchennikova *et al*., 2004). Although interactions

involving AGL15 and other proteins, including MADS-domain proteins were observed at 28-30ºC, the growth higher temperature may have adversely affected protein folding enough to prevent newly transformed, nutrient deprived cells from activating high enough levels of HIS3 or ADE1 to survive. Thus screening at a lower temperature as well as recovery of newly transformed cells on SD/-LW, before selecting for reporter gene activity, are highly recommended for future studies.

4.3.2 Types of AGL15-Interacting proteins recovered by yeast two-hybrid screens

A number of clones encoded for regions that are conserved among other members of the same family. AGL15 could potentially interact with these closely related proteins as well. An emerging picture from extensive analyses of protein interaction surfaces is that usually only a few strongly conserved residues, contribute dominantly and cooperatively to the stability and specificity of protein–protein interactions (reviewed by Uhrig, 2006). Amino acid residues that constitute the interfaces protein-protein are under strong evolutionary pressure and are often identifiable conserved patches of amino acids shared among orthologs and homologs.

4.3.2.1 Proteins with putative enzymatic functions

AGL15-dependent interactions were verified for a SAM dependent methyltransferase (At3g17365), a 1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase (ARD4) (At5g43850), and a raffinose synthase (At5g20250). At3g17365 contains a conserved SAM dependent methyl transferase domain, amino acid 52-155 (Marchler-Bauer *et al*., 2007). Based on inferred roles as metabolic enzymes these interactions were not deemed as likely candidates for biologically significant interactions with a transcription factor *in planta*, and as such were not studied further. However, transcriptional repressors involved in chromatin are known to posses histone methlytransferase (HMT) activity, and use S-adenosyl-L-methionine (SAM) as the methyl group donor to methylate either lysine or arginine residues present on amino-terminal histone tails or in the nucleosome core (reviewed by Berger and Gaudin, 2003), thus the SAM dependent methlytransferase isolated from the yeast two-hybrid screen may play an interesting and significant role in chromatin remodelling. Chromatin immunoprecipitation experiments show that two well characterized mitochondrial enzymes involved in arginine

biosynthesis, ARG5 and ARG6, are associated with specific nuclear and mitochondrial loci in vivo and deletion of either gene causes altered transcript levels of both nuclear and mitochondrial target genes (Hall et el., 2004). GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE (GAPDH) serves as a co-activator to regulate the expression of histone H2B, although it does not directly associate with DNA and may have an indirect role (Zheng, 2003). Therefore these seemingly irrelevant yeast two-hybrid interactions may prove to be of biological importance to the functioning of the plant cell.

4.3.2.2 Proteins with inferred roles in folding and degradation pathways

A protein containing a predicted DNAJ heat shock N-terminal domain (At2g01710), a β-Prefoldin subunit (At1g29990), and SALT AND LITHIUM TOLERANCE 1 (SLT1; At5g02480), a putative molecular chaperone, were excluded from further study, because although the AGL15 mediated interaction appears to be real in the yeast their inferred biological role suggests that they might not be recognising correctly folded AGL15. Molecular chaperones serve to prevent protein misfolding and aggregation. Nascent chain-binding chaperones, including trigger factor, Hsp70, and Prefoldin, stabilize elongating chains on ribosomes in a non-aggregated state (for review see Hartl and Hayer-Hartl, 2002). The DNAJ domain containing protein (At2g01710) is predicted to localize to the nucleus (Mierny, 2001), but so far the only known role for J-domain proteins is in association with the 70-kDa stress proteins, which act as molecular chaperones (Green *et al.,* 1998). *SLT1* encodes a protein implicated in salt tolerance in tobacco and Arabidopsis (Matsumoto *et al*., 2001). SLT1 contains a conserved IbpA domain, which is found in small heat shock proteins and is predicted to function in protein turnover (Marchler-Bauer *et al*., 2007). However, SLT may play a role in transcriptional regulation. Tobacco SLT1 enhances transcription of the CaN-dependent ENA1 gene and compensates the salt sensitivity of a mutant deficient in a transcription factor that normally induces ENA1 expression (Matsumoto *et al*., 2001).

A number of proteins were recovered appeared to be components of the ubiquitination pathways. The ubiquitin/26S proteasome pathway targets proteins for degradation (reviewed by Vierstra, 2003). Ubiquitination has also been implicated in transcriptional regulation. Histone ubiquitination is a reversible covalent modification of histone

residues leading to the formation of an isopeptide bond with ubiquitin, and the E2 ubiquitin-conjugating enzyme UBC2/RAD6 mediates ubiquitination of the H2B in yeast (reviewed by Berger and Gaudin, 2003). UBA1 (At5g06460) encodes for Ubiquitin activating enzyme (E1) and specifically interacts with AGL15 in yeast two-hybrid assay, and although it has not been retested, UBA2 was also recovered (At2g30110). Arabidopsis UBA1 and UBA2 can activate ubiquitin in an ATP-dependent manner and transfer it to a variety of E2s (Hatfield *et al*., 1997). The Arabidopsis genome encodes only two E1s (UBA1 and UBA2), at least 45 E2 or E2-like proteins and almost 1200 E3 components which allow for specificity and targeted regulation (Vierstra, 2003). Because UBA1 and UBA2 are E1 enzymes and more general to the pathway it is not likely that their interaction with AGL15 is of biological significance *in planta*. More than likely they are artificial to the yeast two-hybrid system. It is also interesting to note that a number of putative proteases were recovered from the yeast two-hybrid screen (Appendix E)

4.3.2.3 Proteins that may play a role in signaling or post-translational modification

One putative phosphatase (At4g14930) recovered from a yeast two-hybrid screen tested as a false positive, but a different putative phosphatase (At1g73010) and a protein phosphatase 2C family protein (At4g03415) remain untested. Likewise, a putative Protein Kinase C (At1g70810) and kinase family protein (At4g01330) remains untested (Table 4.2c). These are potentially interesting because the MADS-domain proteins AG and AGL24 can interact with a putative phosphatase and MERISTEMATIC RECEPTOR-LIKE KINASE (MRLK), respectively (Gamboa *et al*., 2001, Fujita *et al*., 2003). In the absence of MLRK, AGL24 is expressed in the cytoplasm and nucleus, whereas it is exclusively nuclear when MRLK is expressed (Fujita *et al*., 2003). AGL15 is initially present in the cytoplasm of egg cells and localizes to the nucleus just before or soon after the first embryogenic cell divisions (Perry *et al*., 1996). Perhaps nuclear localization of AGL15 is mediated by an interaction with a protein kinase?

4.3.2.4 Transcription factors

Interesting transcription factors include VASCULAR PLANT ONE-ZINC FINGER 1 (VOZ1; At1g28520), which encodes a putative one zinc finger transcription factor, bZIP1

(At5g49450), a basic leucine zipper, HOMEODOMAIN GLABROUS 8 (HDG8; At3g03260), A homeobox-leucine zipper protein, LOB40 (At1g67100), which contains a conserved N-terminal DNA -binding LATERAL ORGAN BOUNDAIRES (LOB) domain,

AtVOZ1 and VOZ2 bind to the 38-bp pollen-specific cis-regulatory region of the V-PPase gene (AVP1) and function as transcriptional activators in Arabidopsis (Mitsuda *et al*., 2004). AGL15 interact with amino acids 319-486 of VOZ1, which is a highly conserved region between VOZ1 and VOZ2, and their ortholog in other plants (Mitsuda *et al*., 2004). Domain-B (amino acid 207-423 in VOZ1) is comprised of the zinc-coordinating motif and the basic region, and responsible for the DNA binding and dimerization in AtVOZ2, which has been shown to specifically bind to the palindrome sequence, GCGTNx7ACGC (Mitsuda *et al*., 2004). AtVOZ1 is specifically expressed in the phloem tissue whereas AtVOZ2 is strongly expressed in the root (Mitsuda *et al*., 2004) and according to Genvestigator® (Zimmermann *et al.,* 2004) VOZ1 is expressed highest in curling leaves, senescent leaves, sepals and xylem, and intermediate levels are indicated in the seed, where AGL15 expression is greatest. VOZ1 also interacts with SOC1 (Figure 4.4), whose expression pattern according to Genvestigator® (Zimmermann *et al.,* 2004) (Figure 4.7) more closely resembles that of VOZ1.

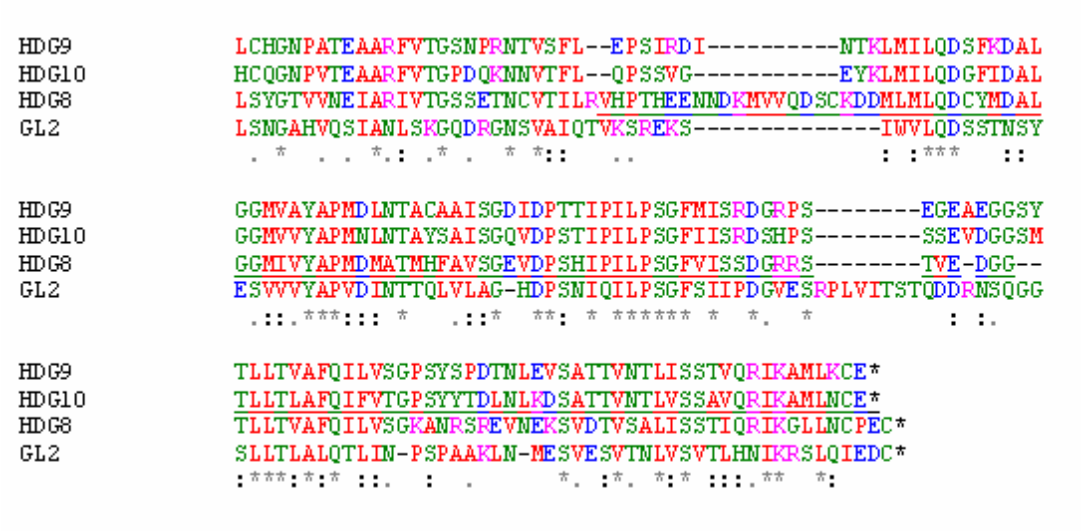
bZIP1 belongs to the S-group whose characterized members include ATBZIP11/ATB2 (reviewed by Jackoby *et al*., 2002), and is localized to the nucleus in plant cells (Satoh *et al*., 2004). Plant bZIP proteins preferentially bind to DNA sequences with an ACGT core (reviewed by Jakoby *et al*., 2002), and ATB2, which belongs to the same group as bZIP1, has been shown to specifically bind ACTCAT in vitro (Satoh *et al*., 2004). bZIP1 interacts with all 4 members of group C (bZIP9, bZIP10, bZIP25 and bZIP63), and with other group S members (bZIP44 bZIP11/ATB2, bZIP42, bZIP58, and bZIP4) in yeast two-hybrid assays (Ehlert *et al*., 2006). Two members of bZIP group C, bZIP9 and bZIP10 show activation activity in yeast-hybrid assays, but none of S group bZIP members' exhibit auto-activation (Ehlert *et al*., 2006). In yeast bZIP1 does not form homodimers, but it does interact with bZIP9 and bZIP10 (Ehlert *et al*., 2006), suggesting that heterodimer might function as transcriptional activators. Indeed, transient expression analyses reveal that four members of the Group S bZIPs, (AtbZIP11/ATB2, AtbZIP44, AtbZIP2/GBF5 and AtbZIP53) activated expression of the GUS reporter gene driven by the ACTCAT sequence while other bZIPs and different families of plant

transcription factors did not (Satoh *et al*., 2004), suggesting that *in planta* co-factors confer activation activity not seen in yeast cells.

HDG8 is a homeobox-leucine zipper protein with a lipid-binding StAR-related lipid transfer (START) domain and a conserved C-terminal region, which is present in other HD-LZ proteins (Marchler-Bauer *et al*., 2007). Although most START domains appear to have diverged in the evolution of plants and animals, one notable exception is the preservation of PCTP-like sequences, which binds phosphatidylcholine (Schrick *et al*., 2004). The homeodomain (HD)-START proteins are unique to plants, suggesting a mechanism by which lipid/sterol ligands can directly modulate transcription in plants (Schrick *et al*., 2004). The Arabidopsis genome contains 16 genes belonging to the class IV homeodomain-Leucine zipper gene family, to which *HGD8* belongs (Nakamura *et al*., 2006). Other members of this family are *GLABRA 2* (*GL2*), *ANTHOCYANINLESS 2* (*ANL2*), *ARABIDOPSIS THALIANA MERISTEM LAYER 1* (*ATML1*), *PROTODERMAL FACTOR 2* (PDF2), *FWA* (aka *HDG6*) and *HDG 1* through *HDG 12* (Nakamura *et al*., 2006). HDG8 interacts with AGL15 via this C-terminal region, which is conserved among members of the class IV HD-LZ (Nakamura *et al*., 2006). The C-terminal region of HDG8 is highly conserved among its four closest relatives. However, an extra 12 amino acids, within the AGL15-interaction domain are unique to HDG8. Database searches of flowering plant ESTs (http://www.ncbi.nlm.nih.gov/BLAST/) failed to uncover any putative orthologs of *HDG8*, and all the similar ESTs deposited into the database to date share a higher identity with *HDG9* or *GL2* than with *HDG8* and do not appear to encode for the extras amino acids observed in HDG8. Whether or not this region is important or if AGL15 can interact with other members of the class IV HD-ZIP family remains to be determined. *HDG8* and other members from the same family have expression patterns consistent with a role in embryo devolvement (Abe *et al*., 2003, Ikeda *et al*., 2007, Nakamura *et al*., 2006, Takada *et al*., 2007). Roles in gene regulation have been assigned to several members of the class IV HD-LZ (Ohashi *et al*., 2003, Ikeda *et al*., 2007, Abe *et al*., 2001, 2003, Shen *et al*., 2006). ATML1 and PDF2 can bind to the L1 box [5'-TAAATG(C/T) A-3'], within the PDF1 promoter *in vitro* (Abe *et al*., 2001, Abe *et al*. 2003) and no PDF1 expression is detectable in *pdf2/atml* double mutants (Abe *et al*. 2003). However, PDF2 transcript is up in atml1 mutant seedling and ATML1 is up in pdf2 mutant seedlings (Abe *et al*. 2003). Mutant alleles of *HDG8* have no mutant phenotype (K. Hill, unpublished observation). However, mutant alleles of

atml1 and *pdf2* have no obvious phenotype either, but the double mutants exhibit severe defects in cotyledon development and fail to produce flowers (Abe *et al*., 2003).

Figure 4.6 Multiple alignments of the C-terminal region of Class IV HD-LZ proteins



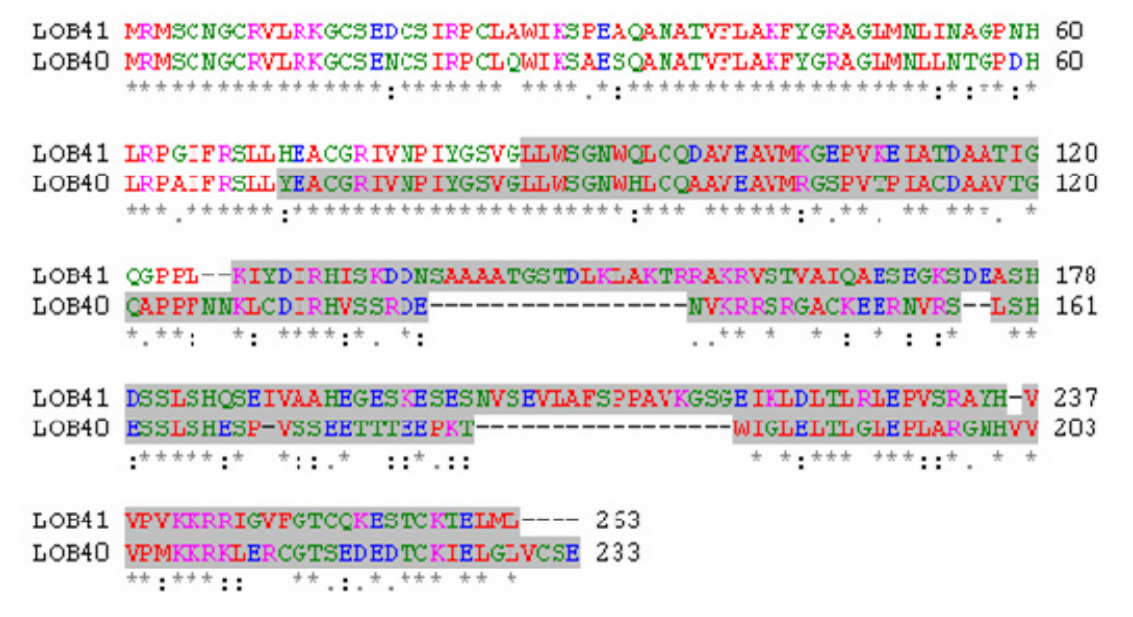
Clustal W software (Chenna *et al*., 2003) was used to align the C-terminal, AGL15-interaction region of HDG8 (amino acids 541-699) with four closest members of the IV class of HD-LZ family, HDG9 (At5g17320), HDG10 (At1g34650), and GL2 (At1g79840).

LOB40 (At1g67100) contains a conserved N-terminal DNA-binding Lateral Organ Boundaries (LOB) domain, AA 1-100 (Marchler-Bauer *et al*., 2007), and interacts with AGL15 via its less conserved C-terminal domain, AA 73-233 (Table 2.3). *In vivo* LEC2, a promoter of somatic embryogenesis, induces the expression of both AGL15 and LOB40 and *in vitro* LEC2 can bind the RY motifs (CATGCA) present in AGL15 and LOB40 promoter regions (Braybrook *et al*., 2006). According to Genevesigator® (Zimmermann *et al*., 2004), LOB40, along with AGL15, is expressed highest in seeds (Figure 4.2). The conserved LOB domain is approximately 100 amino acids in length and is present in 42 other Arabidopsis proteins, which can be divided into two classes (Shuai *et al*., 2002). Members of class I include 36 Arabidopsis genes that are predicted to encode proteins that are similar to LOB (25%–82% identity) throughout the LOB domain, whereas class II consists of six Arabidopsis genes that encode deduced proteins that are less similar to LOB (28%–33% identity) and the other class I proteins (Shuai *et al*., 2002). LOB40 is in class II and shares highest homology with LOB41 (Shuai *et al*., 2002). LOB41 (At3g02550), was also recovered from 3 independent screens, but because of the strong auto activation exhibited by first two recovered clones it was not further analyzed. Yeast assays show the rice ADVENTITIOUS ROOTLESS1 (ARL1) and that the C-terminal domain can act as a transcriptional activator, and because full-length protein failed to activate the reporter genes in the same way as the truncated protein, this activity might be dependent on the unmasking of a repressor domain (the N-terminal LOB domain; Liu *et al*., 2005). The LOB40 clone (amino acids 73-233), recovered in the yeast two-hybrid screen codes for part of the LOB domain (1-100), whereas the false positive LOB41 clones all, or almost the entire LOB domain (87-263, 127-263). A full-length LOB41 was recovered from screen 11, but because previously recovered LOB41 clones behaved as false positives it was not tested for auto-activation. The regions outside of the LOB domain are less conserved and there are patches of amino acids in LOB41, not present in LOB40, which may account for its strong auto-activation. Full-length LOB40 might also repress the slight auto-activation observed.

Although no class II LOB-domain proteins have yet to be assigned functions, class I LOB-domain proteins function as transcription factors and play important roles in development. Loss-of-function LOB mutants have no detectable phenotypes under standard growth conditions, but ectopic expression of LOB cause curled up leaves,

tightly packed cluster of flowers, and abnormal floral organs, along with male and female sterility (Shuai *et al*., 2002). Likewise the *ASYMMETRIC LEAVES2-LIKE 1* (*ASL1*)/*LATERAL ORGAN BOUNDARIES DOMAIN 36* (*LBD36*) gene loss-of-function mutant, *asl1/lbd36* exhibits no morphological aberration (Chalfun-Junior *et al*., 2005), but overexpression of *ASL1* results in phenotypes similar to those of LOB (Ueno *et al*., 2007). However, the *asymmetric leaves2* (*as2*) mutant of generates leaf lobes and leaflet-like structures from the petioles of leaves in a bilaterally asymmetric manner (Semiarti *et al*., 2001). The rosette leaves of *as1* and *as2* single mutants have a higher potential for regeneration of shoots in vitro without exogenous hormones (Semiarti *et al*., 2001). Braybrook *et al*., 2006 make the statement *"LEC2 may also induce somatic embryo development by increasing tissue competency to undergo somatic embryogenesis through AGL15"* and it will be interesting to determine if, like AGL15 (Harding *et al*., 2003), over-expression of LOB40 also promotes somatic embryogenesis or if over-expression of both LOB40 and AGL15 enhance this.

Figure 4.7 Multiple alignments of LOB40 and LOB41



Clustal W software (Chenna *et al*., 2003) was used to align LOB40 (At1g67100) and LOB41 (At3g02550). The amino acids shaded grey show those encoded by the clones that were retested in yeast two-hybrid assays.

4.3.2.5 Putative chromatin remodelling factors

Proteins with inferred roles in chromatin remodeling include a member of the SIN3/HDAC complex (see Chapter 3), a putative histone acetyl transferase (HAG5; At5g09740), and an Agenet domain-containing protein (At3g62300). A clone encoding for HISTONE 4 (H4; At1g07820) was also recovered, but failed to activate the reporter genes when tested.

Arabidopsis *GCN5/HAG1* is a member of the GNAT-MYST HAT super-family (Pandey *et al*., 2002) and mutations in this gene of result in a long-hypocotyl phenotype and reduced light-inducible gene expression, whereas mutation of *HDA19* induced opposite effects (Benhamed *et al*., 2006). GCN5 directly associates with the light-responsive promoters and is required for acetylation on the target promoters, whereas HDA19 cause a decrease in acetylation (Benhamed *et al*., 2006). Given the suggested antagonist role of GCN5, which is similar to HAG5, a putative AGL15 interacting protein, and HDA19 (an AGL15-interacting protein, see chapter 3), further study regarding the association between AGL15 and HAG5 may be warranted.

AGL15 interacts with amino acids 263-722 of an Agenet domain containing protein (At3g62300). Three conserved Agenet domains are found between amino acids 98-145, 161-226, and 228-284. Another conserved domain, excusive to plants and often found in association with the Agenet domain, lies between amino acids 115-178 (Marchler-Bauer *et al.,* 2007). Twenty eight Agenet domain-containing genes are found in Arabidopsis and some of these co-occur with acetyltransferase, plant homeodomain (PHD), and BROMO-ADJACENT HOMOLOGY (BAH) domains (Maurer-Stroh *et al.,* 2003).

4.3.2.6 Putative RNA binding proteins

The literature contains limited example of plant MADS-domain proteins interacting with RNA binding proteins, such as a poly(A)-binding protein II-like (At5g65260) being recovered from a yeast two-hybrid screen using AP1 as bait (Pelaz *et al*., 2001). However, a picture is emerging relating the importance of transcription factors in RNA-processing mechanism (reviewed by Kornblihtt *et al.,* 2004)

A number of putative RNA-binding proteins were isolated from yeast two-hybrid screens. Those that specifically interact with AGL15 in yeast include, GRP2 (see Chapter 3) and a KH domain/CCCH type zinc finger containing protein (At3g12130). Amino acids 143-248 are encoded by the clone recovered from the yeast two-hybrid a screen, which partially overlaps the K homology RNA-binding domain, amino acids 115-178, and the second CHHH zinc finger amino acids, 213-239 (Marchler-Bauer *et al*., 2007). CHHH zinc fingers are characterized by three cysteine residues and one histidine residue that coordinate the zinc ion and form, in general, a Cys-X8-Cys-X5-Cys-X3-His sequence. The prototypic CCCH zinc finger protein is tristetraprolin (TTP), an RNA-binding protein that binds to AU-rich elements in the 3'untranslated regions (UTRs) of certain oncogenes and induces the degradation of the messages and so regulates the level of protein expression (Hall, 2005). Solution structure of the tandem zinc fingers of Tis11d in complex with AU-rich element RNA reveals that each zinc finger module binds to the sequence UAUU; thus, the tandem zinc fingers bind to UAUU-UAUU, although a single CCCH zinc finger can bind weakly but specifically to AU-rich element RNA (Hall, 2005). Twenty six proteins containing KH domains are present in the Arabidopsis genome, but only one other, At5g06770, contains both the KH- and CHHH-domains (Lorkovic and Barta 2002). *HUA ENHANCER 4* (*HEN4*) codes for a K-homology (KH) domain-containing, putative RNA binding protein that interacts with HUA ENHANCER 1 (HUA1), a CCCH zinc finger RNA binding protein in the nucleus (Cheng *et al*., 2003). Mutations in HUA1 and HEN1 result in *agamous*-like phenotypes and compromise *AG* pre-mRNA processing as evidenced by an increase in *AG* RNA retaining the second intron (Cheng *et al*., 2003). The CCCH zinc finger protein, HUA1 binds *AG* RNA in vitro (Cheng *et al*., 2003), and the first intron of *AG* contains a CArG motif which may serve as a MADS-binding site. The flowering-time gene, *FLOWERING LOCUS K* (*FLK*) encodes a protein with three KH domains and functions as a repressor of *FLOWERING LOCUS K* (*FLC*; Mockler *et al*., 2004). FLK is able to bind to FLC RNA *in vitro* (Mockler *et al*., 2004), and FLC, like AG also contains a CArG motif in the first intron which may serve as a MADS-binding site. Higher levels of AGL15 transcription are observed in transgenic plants carrying a form of AGL15 containing the first three introns compared to those transformed with the cDNA version (Fernandez *et al.,* 2000), suggesting that these intronic sequences are important in transcriptional regulation. AGL15 binds its own promoter and regulates transcription (Zhu and Perry, 2005). Therefore, it would be

interesting to test if there was a connection between transcription and RNA-processing, mediated by the interaction of AGL15 with the KH-domain protein. Work in our lab has identified *in vivo* AGL15-binding sites, and in some case AGL15 does not seem to affect transcript levels of the gene to which it binds. Although there are other explanations, one might be that AGL15 is mediating mRNA processing through its recruitment of members of the RNA-processing machinery, such as the KH-domain protein. Depending upon the region of a transcript oligonucleotides were designed to amplify, such changes might have been overlooked. It would be interesting to determine if some of the direct targets of AGL15 retain their intron in a KH-mutant background.

4.3.2.7 Proteins of unknown functions

An unknown protein (At1g29980) contains a conserved DUF642 domain, but bears no resemblance to any protein of known function. This according to Genevestigator® (Zimmermann *et al.,* 2004) this gene is predominantly expressed in seeds (Figure 2.2).

TAPETUM DETERMINANT PROTEIN 1 (TD1; At4g24973) was recovered from three independent screens, thus it was surprising that the interaction was not reproducible in directed tests. Another potentially interesting AGL15-interacting protein that failed subsequent retesting was EARLY RESPONSIVE TO DEHYDRATION 15 (ERD15). ERD15, a small acidic protein, is a novel mediator of stress-related ABA signaling involved in freezing and pathogen resistance (Kariola *et al.,* 2006). ERD15 activates reporter genes when fused to the GAL4-DBD and interacts with a Poly(A)-binding protein (PABP) in yeast two-hybrid assays (Wang and Grumet, 2004). Poly(A)-binding proteins (PABPs) are multifunctional proteins that play important roles in mRNA stability and protein translation. ERD15 possesses a motif common to four other PABP-binding proteins SxLNxxAxxFxP, which is necessary for the interaction yeast (Wang and Grumet, 2004).

5.3.3 Three putative α-helices reside in the regions of AGL15 that mediate protein-protein interactions

The K domain of the higher plant MIKC^c types MADS is characterized by 3 separate strings of heptad repeats (abcdefg)n with hydrophobic amino acids predominantly in a and d positions and is assumed to generate an interaction surface that consists of amphipatic α-helices, potentially forming coiled coils (Yang and Jack, 2004; Figure 2.11a). K1, K2 and the region between K1 and K2 but not the K3 helices (Figure 2.19) are important for strong PI-AP3 and PI-SEP3/1interaction (Yang & Jack, 2004) and the majority of amino acids for critical to the AP3/PI interaction are located on the putative hydrophobic faces of these first two α-helices (Yang *et al*., 2003a). A point mutation in the *PI* gene, *pi-5*, which causes a glutamic acid to be replaced by lysine within the K-domain (E125K), which abolishes the ability of PI to dimerize with AP3 or SEP3 in yeast (Yang *et al*., 2003 b). Weak *ap3* and *pi* alleles have defects in whorls 2 (petals to sepals) and whorls 3 (stamen to carpels), whereas *pi-5* has defects only in whorl 2 (petaloid sepals, rather than petals to sepals), (Yang *et al*., 2003 b) and flowers resemble those of anti-sense *35S:SEP3* (Pelaz *et al*., 2001). These phenotypes likely represent reduced capacity for interactions with member of a particular complex.

4.4 Materials and methods

4.4.1 Yeast two-hybrid library construction and screening

cDNA encoding for full-length, or various truncations of AGL15, were cloned into pGBKT7 to produce the GAL4 DNA Binding Domain-AGL15 fusion protein expression constructs, DBD-AGL15 (MIKC), DBD-AGL15 (IKC), DBD-AGL15 (C), DBD-AGL15 (MI), that were used as "bait" in the yeast two-hybrid screens (see Appendix A).

4.4.1.1 Expression library construction

The Clontech™ Matchmaker Library Construction and Screening Kit (Clontech, Mountain View, CA) was used, in accordance with manufacturers instructions, to screen a cDNA expression library, derived from Arabidopsis embryonic culture tissue (ECT; Harding *et al*., 2003) or Brassica napus embryos for putative protein-protein interactions involving the MADS-domain transcription factor, AGL15 (At5g13790). TRIZOL Reagent (Invitrogen, Calsbad, CA) was used to isolate total RNA from ~50mg of *Arabidopsis thaliana* ECT or 5-6 day old *Brassica napus* embryos. Double-stranded cDNA was generated in accordance with the Clontech™ Matchmaker Library Construction and Screening Kit Instruction Manual (protocol PT3024-1, version PR13103, 2001, Clontech, Mountain View, CA). Unless otherwise stated all components used to generate a ds-cDNA library were supplied in The Clontech™ Matchmaker Library Construction and Screening Kit (Clontech, Mountain View, CA). For first strand synthesis 1 µg of poly A^+ RNA and 1 µl CDS III (random or oligo dt) primer were incubated together in a total volume of 4 µl RNAase free deionized H_2O for 2 minutes at 72 ºC, followed by 2 minutes incubation on ice. First-strand buffer, 2 mM DTT, 1mM dNTPs, and 1µl MMLV Reverse were added to a final volume of 9 µl, incubated at 42 ºC for 10 minutes (oligo dt) or 25ºfor 10 minutes followed by another 10 minute incubation (random oligo), 1 µl SMART III was added before 1 hour incubation at 42ºC. The reaction was terminated by 10 minutes incubation at 75ºC. Once the sample had cooled to room temperature 2 units of RNaseH was added. Either 2 µl of synthesized first strand generated using the oligo d(T) or 1 µl of the oligo d(T) generated plus 1 µl or the random oligo generated first-strand was used in the Long Distance-PCR (LD-PCR) reaction, which was performed as directed using Clontech™ Advantage 2 Polymerase Mix, and minimum number of

amplification cycles (15-21). The sample was purified using CHROMA SPIN+ TE-400 columns (Clontech, Mountain View, CA), and suspended in a final volume of 20 µl. A 2µl aliquot was run on a 1.2% agarose/EtBr gel to verify product ranging in size from 0.1 to 4 kb.

4.4.1.2 Preparation of yeast competent cells

Yeast competent cells were prepared using the LiAc method, as described in The Clontech™ Yeast Protocols Handbook (version PR13103, 2001, Clontech, Mountain View, CA). Briefly, 50ml of YPDA (tryptone 20g/L, yeast extract 10g/L, 2% dextrose, 0.003% adenine) was inoculated with fresh (2-5 day old) AH109 (MATa, trp1-901, leu2-3, 112, ura3-52, his3-200, gal4Δ, gal80Δ, LYS2::$GAL1_{UAS}$-$GAL1_{TATA}$-HIS3, $GAL2_{UAS}$ - $GAL2_{TATA}$-ADE2, URA3::$MEL1_{UAS}$-$MEL1_{TATA}$-lacZ, MEL1) or Y187 (MATα, ura3-52, his3-200, ade2-101, trp1-901, leu2-3, 112, gal4Δ, met, gal80Δ, URA3::$GAL1_{UAS}$-$GAL1_{TATA}$-lacZ, MEL1) colonies, incubated overnight on a shaker (~200 rpm) and 28ºC. 300 ml of YPDA was inoculated from the overnight culture to an OD600 of 0.2-0.3, and grown for a further 3 hrs, at which point the OD600 reading was 0.4 to 0.6. Cells were harvested by room-temperature centrifugation for 5 min at 1,000 xg, washed once in TE (10 mM TRIS, 1 mM EDTA, pH 8), and suspended in 5 ml 1XTE/1XLiAc (TRIS 10 mM, LiAc 100 mM. Competent cells were used immediately for transformation.

Two protocols were used to screen the cDNA libraries for AGL15-interacting proteins: Co-transformation or mating of independently transformed yeast strains.

4.4.1.3 Co-transformation method

The 600 µl of AH109 competent cells co-transformed with 5 µg DBD-AGL15 (MIKC, IKC, C, or MI), 3 µg pGADT7-Rec, 18 µl double-stranded cDNA, and 20 µl denatured Herring Testes Carrier DNA. 2.5 mL of PEG/LiAc solution (TRIS 10 mM, LiAc 100 mM, 40% PEG-400) was added, and samples incubated at 30 ºC for 45 minutes, with gently mixing at 15 minute intervals. 160 µl DMSO (Sigma-Aldrich, St. Louis, MO) was added, the tubes inverted several times, and heat shocked at 42 ºC for 20 minutes. Cells were centrifuged for 5 minutes at 700 xg, suspended in 3 ml of YPD Plus Liquid Media™

(Clontech, Mountain View, CA), and allowed to recovered at 30 ºC for 90 minutes with gentle shaking. The cells were harvested by centrifugation at 700 xg for 5 minutes and re-suspended in 6 ml of 0.9% NaCl.

To determine the transformation efficiency and calculate the number of colonies screen, a 30 µl aliquot was removed and diluted in 720 µl 0.9% NaCl. 150 µl was spread on to SD/-L (2% dextrose, 0.67% nitrogen base, 1% agar, all amino acids except leucine) and SD/-LW (2% dextrose, 0.67% nitrogen base, 1% agar, all amino acids except leucine and tryptophan), and the number of colonies appearing after 4-5 days was recovered.

Transformation Efficiency (number of transformants/3 µg pGADT7-rec) =

Number of colonies appearing on SD/-L x 1000

Number of clones screened =

Number of colonies appearing on SD/-LW x 1000.

Transformed AH109 yeast colonies, able to activate the reporter genes (*HIS3*, *ADE2*, and *MEL1*), were selected onto either SD/-LWH (2% dextrose, 0.67% nitrogen base, 1% agar, all amino acids except lacking leucine, tryptophan, and histidine) or SD/-LWHA X-α-gal (all amino acids except leucine, tryptophan, histidine, and alanine, supplemented with 0.2mg/ml X-α-gal). The plates were incubated at 25-30º C as indicated (Table 4.1). Colonies appearing on SD/-LWH were subsequently re-streaked onto SD/-LWHA X-α-gal. Colonies were re-streaked onto selective media several times to reduce the number of colonies carrying multiple preys.

4.4.1.4 Mating method

AH109 competent were cells transformed with 18 µl double-stranded cDNA (derived from *Arabidopsis* embryonic tissue culture or *Brassica napus* embryos) and 3 µg pGAD7-rec as described above, and tranformants selected for on SD media (2% dextrose, 0.67% nitrogen base, 1% agar) lacking leucine (-L). After several days the transformed colonies were harvested by chilling the plates for several hours at 4 ºC and

scraping the cells into 5ml of Freezing Media (SD media lacking leucine, 65% glycerol, 0.1 M MgSO4, 25 mM Tris-HCl pH 7.4). A hemocytometer was used to verify that the cell density was $>10^7$ cells/ml. 1 ml aliquots (10^7 cells/ml) were stored at -80 ºC until ready to use.

Y187 competent cells were transformed with 500 ng DBD-AGL15 (MIKC, IKC, C, or MI) as, and tranformants selected for on SD media (2% dextrose, 0.67% nitrogen base, 1% agar) lacking tryptophan (-W). 50ml of SD/-W was inoculated with freshly (2-5 day old) transformed colony and incubated overnight on a shaker (~200 rpm) and 28ºC. Cells were harvested by centrifugation at 600 xg for 5 minutes and suspended in ~5 ml SD/-W to a density of $\geq 1x10^9$ cells/ml and added to 45 ml 2X YPDA along with 1 ml of library. The cells were incubated in a 2 L flask at 28 ºC for 20 hours with gently swirling (30-50 rpm). After 20 hours a drop of the mating mix was examined under a phase-contrast microscope (400x) and if zygotes were still present the mating was allowed to continue for a further 4 hours. The cells were harvested by centrifugation at 1,000 xg for 10 minutes, and re-suspended in 10 ml 0.5X YPDA.

To select for diploids able to activate the reporter genes (*HIS3*, *ADE2*, and *MEL1*), the yeast cells were spread onto plates as described for the co-transformation approach (see above). Diploid colonies AH109, able to activate the reporter genes (*HIS3*, *ADE2*, and *MEL1*), were selected for on SD media (2% dextrose, 0.67% nitrogen base, 1% agar) lacking leucine (-L), tryptophan (-W), histidine (-H), and adenine (-A) and supplemented with 0.2mg/ml X-α-gal (SD-L/-W/-H/-A, 0.2mg/ml X-α-gal). The prey plasmids were recovered using a protocol described by Rose, 1987, and sequenced by The University of Kentucky Advanced Genetic Technologies Center, UK-AGTC.

4.4.1.5 Plasmid recovery

The prey plasmids were recovered using a protocol described by Rose, 1987. The following modifications were made to the original protocol: Yeast colonies were grown overnight to late log phase in SD/-Leu media to retain selective pressure on the prey plasmids. 1.5 ml of cells were harvested by centrifugation and washed in 1 ml of 0.9 M sorbitol, 0.1 M Na_2EDTA (pH 7.5). The cells were re-suspended in 0.4 ml of 0.9 M sorbitol, 0.1 M Na2EDTA (pH7.5), 14 mM 2-mercaptoethanol, and 0.1 ml zymolyase

(2mg/ml) made up in the 0.9M sorbitol solution. The reaction was incubated at 37 ° for 2hrs. The spheroplasts were centrifuged and gently re-suspended in 0.45 ml of 50 mM Tris-HCl (pH 8.0), 50 mM Na_2EDTA, 2% SDS, and incubated at 65 °C for 30 minutes. 80 µl 5 M potassium acetate was added and the samples prior to a minimum of 60 minutes incubation on ice. The precipitate was removed by 15 minutes centrifugation at maximum speed and the supernatant transferred to a fresh tube. The DNA was precipitated by addition of 1 ml of room temperature ethanol, immediately followed by a brief centrifugation. The pellet was rinsed with cold 70% ethanol, air dried, and re-suspended in 0.5 ml TE (10 mM Tris-HC1 at pH 8.0, 1 mM Na2EDTA). The insoluble material was removed by centrifugation for 15 minutes and the supernatant transferred to a fresh tube, 25µl of RNase at 1 mg/ml was added and the sample incubated at 37 ° for 30 min. Finally the DNA was precipitated by adding an equal volume of 2-propanol, mixing gently, and spinning in centrifuge for 10 min at maximum speed. The pellet was washed once and re-suspended in 30µl of sterile ddH_2O. 2 µl was used to transform highly competent (>10^6) *E.coli* (DH5α) cells, which typically take up only one plasmid, allowing those colonies having taken up the prey plasmid to be selected on solid media containing 5mg/ml ampilicin.

The recovered plasmids were sequenced by The University of Kentucky Advanced Genetic Technologies Center, UK-AGTC.

4.4.2 Yeast two-hybrid directed tests

For directed tests AH109 100 µl of competent cells were co-transformed with ~500 ng of specific bait and prey constructs, along with 10 µl denatured Herring Testes Carrier DNA and plated onto SD-L/-W media. After several days transformed colonies were re-streaked onto SD plates selective for protein-protein interactions (SD-L/-W/-H/-A, 0.2mg/ml X-α-gal).

Bait constructs used in this study:

DBD-AGL15 (MIKC), DBD-AGL15 (MIK½), DBD-AGL15 (IKC), DBD-AGL15 (KC), DBD-AGL15 (½KC), DBD-AGL15(C), DBD-AGL15 (IK), DBD-SVP (MIKC), DBD-SVP (IKC),

DBD-SOC1 (MIKC), DBD-SOC1 (IKC), DBD-FLC (MIKC), DBD-SHP1 (MIKC), DBD-PI (MIKC), DBD-AGL18 (MIKC), DBD-AGL18 (IKC), and DBD-SEP3 (MIKC)

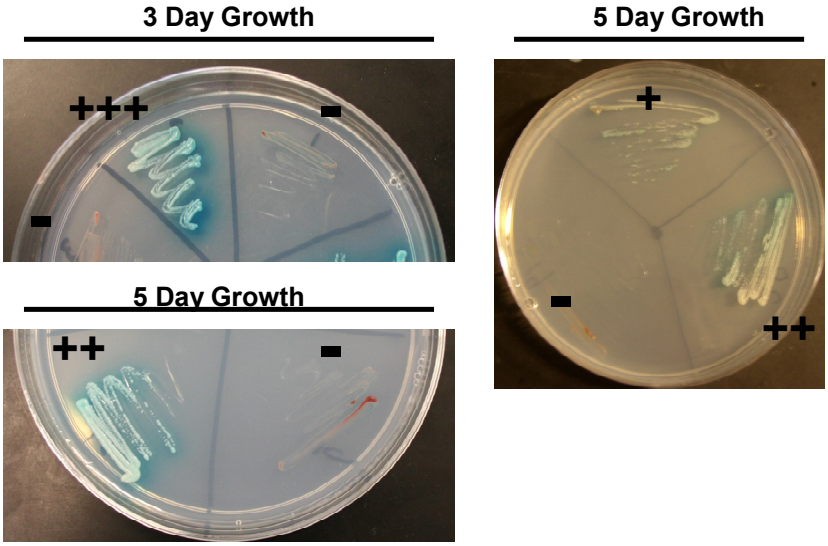
The oligonucleotides and restriction sites used generate the pGBKT7 (bait) constructs can be found in Appendix A

Prey constructs used in this study:

For details on recovered prey see Table 4.2 and Appendices C-E

All directed tests were repeated at least three times, and were performed along with positive and negative controls. Several colonies, growing on SD-L/-W, were heavily streaked onto SD-L/-W/-H/-A, 0.2mg/ml X-α-gal. Only AH109 colonies able to activate the nutritional reporter genes His^+ and Ade^+ can grow on this media, and blue coloration indicates expression of Mel1 reporter gene, thus controlling for contamination by wild-type yeast strains. Growth was scored as follows: +++, strong activation of reporter genes determined by heavy growth after 2-3 days; ++, moderate activation of reporter genes determined by heavy growth after 4 days; +, weak activation of reporter genes determined by light but clearly visible growth by 5 days; +/-, very weak/transient activation of reporter genes; -, no activation of reporter genes determined by the absence of any growth after 5 days (Figure 4.8). A pink growth is sometimes observed near the top of the plate, where the streaking is most heavy (Figure 4.8). This is due to dying cells and disappears after several days.

Figure 4.8 Scoring of yeast growth on SD-L/-W/-H/-A, 0.2mg/ml X-α-gal



+++, strong activation of reporter genes; ++, moderate activation of reporter genes; +, weak activation of reporter genes; -, no activation of reporter genes.

4.5 Summary

Yeast two-hybrid screens were undertaken to identify novel proteins able to interact with AGAMOUS-LIKE 15 (AGL15). Despite the high volume of false positive and seemingly irrelevant associations, a number of interesting and potentially biologically important AGL15-interacting partners were discovered. These include a member of the SIN3 histone deacetylase complex, SAP18, and a CSD-containing protein, GRP2, which are described Chapters 2 and 3 respectively. Other interesting AGL15-interacting proteins include a K-homology domain/CCCH type zinc finger containing protein, a homeobox-leucine zipper protein, a bZIP transcription factor, a putative histone acetyl transferase, a LOB-domain containing protein, and an Agenet domain containing protein. Regions of AGL15 mediating these interactions were mapped to one of three regions: the K-domain, the C-domain, or the K-and C-domains together. Some, but not all of the AGL15-interaction proteins were also able to associate with other MADS-domain proteins in yeast two-hybrid assays.

5.1 Introduction

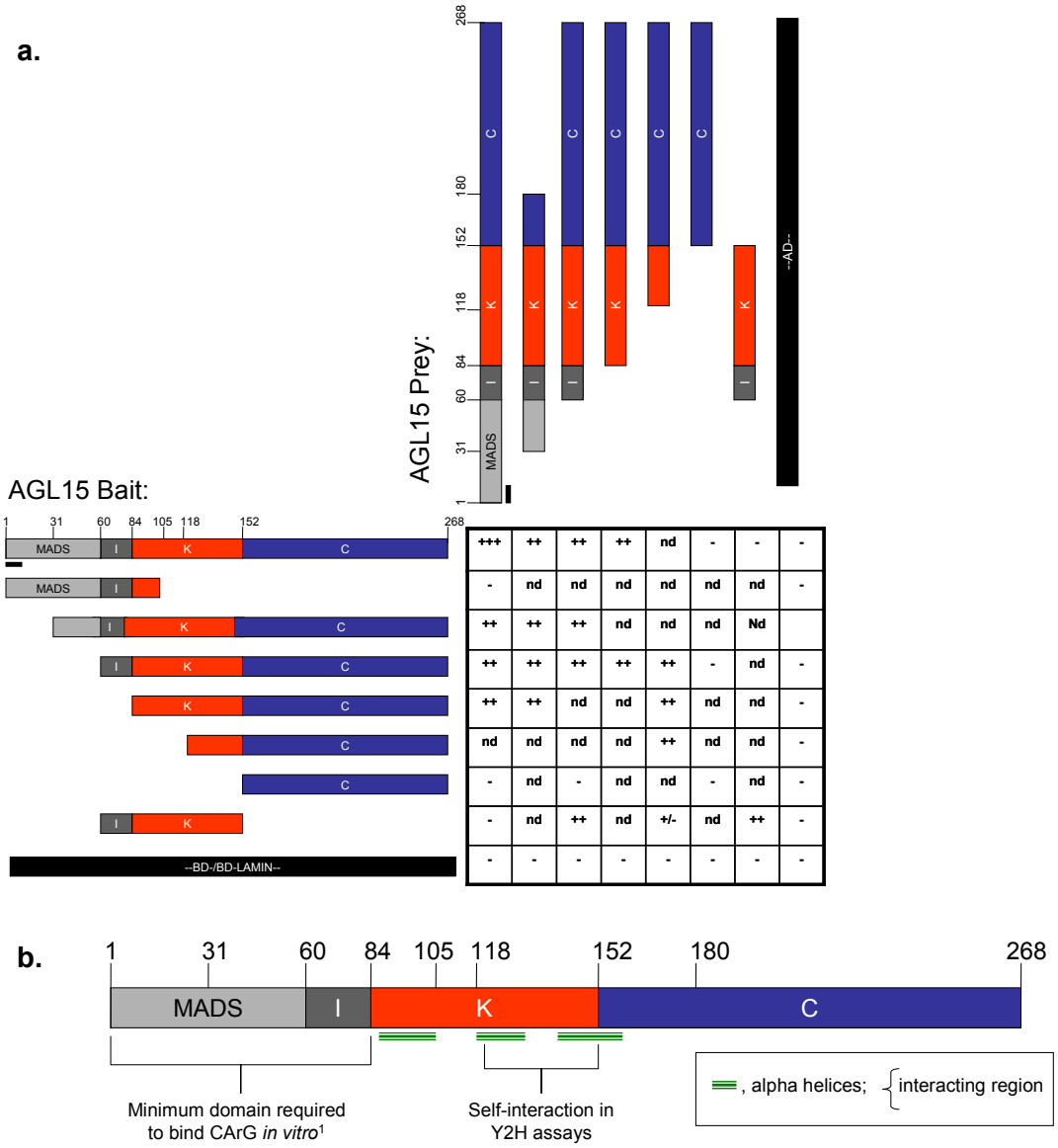
Upon completion of the first yeast two-hybrid screens one of the questions that arose was the fact that, with the exception of AGL15 itself, no other MADS-domain proteins were recovered. This was a surprise given the considerable volume of literature pertaining to the plant MIKC MADS-domain proteins interacting with one another (Mizukami *et al*., 1996, Fan *et al.,* 1997, Egea-Cortines *et al.,* 1999, Moon *et al*., 1999, Lim *et al*., 2000, Honma and Goto, 2001, Immink *et al.,* 2002, Jang *et al.,* 2002, Favaro *et al.,* 2002, 2003, Causier *et al.,* 2003, Yang *et al.,* 2003a, b, Yang and Jack, 2004, Shchennikova *et al*., 2004, de Folter *et al.,* 2005, Cseke *et al.,* 2007, and others), and the driving force behind the initiation of directed yeast two-hybrid assays. Because transcription factors are generally low abundance transcripts it is possible that they were not detected in yeast two-hybrid screens. Indeed, directed tests revealed interactions between AGL15 and other MADS domain proteins present in embryonic tissue. At the time this study was initiated no other MADS-domain proteins had been reported in the literature as interacting with AGL15. However, that has changed and as a consequence most of the MADS-MADS interactions presented herein have since been reported by other groups. One exception is the interaction between AGL15 and SEPALLATA 3 (SEP3), which has been expounded upon and its potential role in somatic embryogenesis is discussed here.

5.2 Results

5.2.1 Elucidation of the regions of AGL15 that mediate self interaction

Based on a variety of studies, including crystal structures, MADS domain proteins are inferred to bind DNA as dimers (for reviews see Riechmann and Meyerowitz, 1997, Messenguy and Dubois, 2003). Because the MI domains of AGL15 were sufficient to mediate DNA-binding *in vitro* (W. Tang and S. Perry, unpublished data) it seems likely that AGL15 is able to interact with itself to form a homodimer. Therefore AGL15 homodimerization was used as a positive control in the yeast two-hybrid assays. Full-length AGL15 has since been reported to interact with itself and other MADS-domain proteins in yeast-two-hybrid studies (de Folter *et al.,* 2005). In this study a series of yeast two-hybrid assays were performed, whereby truncated forms of AGL15 were assayed to determine the regions of AGL15 necessary for self-interaction. The latter half of the K domain of AGL15 (AA 118 – 152) was sufficient to mediate a protein-protein interaction in yeast (Figure 5.1a, b). Taken together these data suggest that AGL15 contains two distinct self interacting regions, one involving the MADS plus I domains and another involving the K domain.

Figure 5.1 Elucidation of the regions of AGL15 that mediate self interaction in yeast two-hybrid assays



+++	++	++	++	nd	-	-	-
-	nd	nd	nd	nd	nd	nd	-
++	++	++	nd	nd	nd	Nd	
++	++	++	++	++	-	nd	-
++	++	nd	nd	++	nd	nd	-
nd	nd	nd	nd	++	nd	nd	-
-	nd	-	nd	nd	-	nd	-
-	nd	++	nd	+/-	nd	++	-
-	-	-	-	-	-	-	-

a. Yeast two-hybrid assays

Black bar, 10 aa; nd, not determined; +++, strong activation of reporter genes; ++, moderate activation of reporter genes; +, weak activation of reporter genes; +/-, very weak/transient activation of reporters genes; -, no activation of reporter genes (for a visual representation of how growth was scored see Figure 4.8). Results represent a minimum of three independent assays.

b. Schematic depicting regions of AGL15 that mediate self interactions

[1]Although not detected in the yeast two hybrid studies, pervious data suggests that DNA-binding homodimerization involving the MADS and I domain are able to form *in vitro* (Tang and Perry, unpublished data).

5.2.2 MADS-domain proteins form hetero- and homo-dimers in yeast two-hybrid assays

To address the question of why no MADS domain proteins, other than AGL15 itself, were recovered from the yeast two-hybrid screens, a series of directed tests were performed using MADS expressed in embryonic tissue culture (Lehti-Shiu *et al.,* 2005). In Table 5.1 an interaction, regardless of the truncations used is indicated by a [+] symbol, and a [–] symbol indicates that no interaction was ever observed. Because some interactions appear to be obscured by the presence of the MADS domain, and because truncations were not made for all the MADS-domain proteins tested, some combinations were more thoroughly tested than others, and further examination might reveal an interaction previously overlooked. The left-hand column assayed for the ability of full-length DBD-fusions to activate reporter genes in the presence of the GAL4-AD alone. Of those tested (AGL15, AGL18, SVP, SHP1, FLC, SOC1, PI, SEP2, and SEP3) only SEP3, as previously reported (Homo and Goto, 2001, de Folter *et al.,* 2005) activated the reporter genes and this was barely detectable when activation of all three reporter genes was assayed.

Table 5.1 Results of yeast two-hybrid assays testing for interactions between MADS domain proteins whose mRNA is detectable in ECT[1]

	AD	AGL15	AGL18	SVP	SHP1	FLC	SOC1	SEP3	SEP2	SEP1	PI
DBD	-	-	-	-	-	-	-	-	nd	-	-
AGL15	-	+ 2	-	+ 2	+/- 2	-	+ 2	+	nd	-	-
AGL18	-	-	- 2	-	-	-	-	-	nd	nd	-
SVP	-	+ 2	-	+ 2*	nd	nd	+ 2	+/- 2	nd	+ 2	nd
SHP1	-	+ 2	-	+ 2*	- 2	+ 2*	- 2*	+ 2,3	nd	nd	nd
FLC	-	-	-	+	- 2	+ 2*	nd	+ 2*	nd	nd	nd
SOC1	-	+/- 2	-	+ 2	nd	-	+ 2	- 2*	nd	nd	nd
SEP3	+/-(C)	+	-	nd	nd	nd	nd	+	nd	+ 2*	nd
SEP2	-	-	-	+	nd	nd	nd	nd	nd	nd	nd
SEP1	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
PI	-	-	-	-	nd	nd	nd	+/- 4	nd	nd	- 4

DBD, GAL4-DNA Binding Domain; AD, GAL4-Activation Domain; Left hand column, DBD-fusions; Top row, AD-fusions. (C), indicates that auto-activation occurred only when the C-terminal domain was present. nd, not determined; +, activation of reporter genes; +/-, transient, or inconsistent activation of reporter genes; -, no activation of reporter genes. Because various truncated forms of the proteins listed were tested a + indicates any interaction that was observed, regardless of the inferred strength of the interaction of domains tested. [1]Lehti-Shiu *et al.,* 2005, [2]Interaction/non-interaction confirmed by de Folter *et al.,* 2005, 2* Opposite result confirmed de Folter *et al.,* 2005, [3]Interaction/non-interaction confirmed Favaro *et al*., 2003, [4]Interaction/non-interaction confirmed by Yang & Jack, 2004

Some but not all MADS-domain proteins are able to form homodimers

AGL15 and SOC1 are able to form homodimers (Table 5.1, de Folter *et al.,* 2005). No AGL18, SHP1 or PI homodimers were detected using the yeast two-hybrid assay (Table 5.1, de Folter *et al.,* 2005). This study demonstrated that full-length SVP, SEP3, and full-length FLC were able to interact with themselves in yeast two hybrid assays (Table 5.1), although no such homodimerization was reported for either of these MADS-domain proteins in a similar study (de Folter *et al.,* 2005).

AGL15 is able to interact with other MADS-domain proteins

Full-length AGL15 interacts with MIKC versions of SVP, SOC1, SHP1 (Table 5.1, de Folter *et al.,* 2005), SHP2, AP1, AGL6, AG, STK, AGL16, AGL21 (de Folter *et al.,* 2005). Truncated forms of AGL15, lacking the MADS domain, were able to interact with SEP3, but not SEP1 or SEP2 (Table 5.1), although there is no detectable interaction between full-length AGL15 and SEP3 in yeast two-hybrid assays (Figure 5.3a, de Folter *et al.,* 2005).

MADS-domain proteins that interact with SEP3 (an interacting partner of AGL15)

SEP3 can interact with AGL15 in yeast two-hybrid assays, although SEP2 and SEP1 do not (Table 5.1). SEP3 also interacts with SHP1 (Table 5.1, Favaro *et al*., 2003, de Folter *et al.,* 2005), AGL24, AP1, AGL6, AGL16, AG (de Folter *et al.,* 2005), SHP2 (Favaro *et al*., 2003, de Folter *et al.,* 2005), and STK (Favaro *et al*., 2003, de Folter *et al.,* 2005). This study demonstrated an interaction in yeast between full-length SEP3 and FLC (Table 5.1), although full-length SEP3 is reported as not interacting with full-length FLC in a yeast two-hybrid assay performed by another group (de Folter *et al.,* 2005). Conversely, in this study no interaction was observed between full-length SEP3 and SOC1 (Table 5.1), although a positive interaction was previously reported (de Folter *et al.,* 2005).

MADS-domain proteins that interact with SVP (an interacting partner of AGL15)

SVP can interact with AGL15 in yeast two-hybrid assays (Table 5.1, de Folter *et al.,* 2005). SVP also interacts with SOC1, SEP3 and SEP1 (Table 5.1, de Folter *et al.,* 2005), AP1 (Pelaz *et al*., 2001, de Folter *et al.,* 2005), AG, AGL6, AGL21 (de Folter *et al.,* 2005). In addition this study demonstrated an association between full length SVP

and SHP (Table 3), although no interactions between these partners were reported in a previous study (de Folter *et al.,* 2005).

MADS-domain proteins that interact with SOC1 (an interacting partner of AGL15)

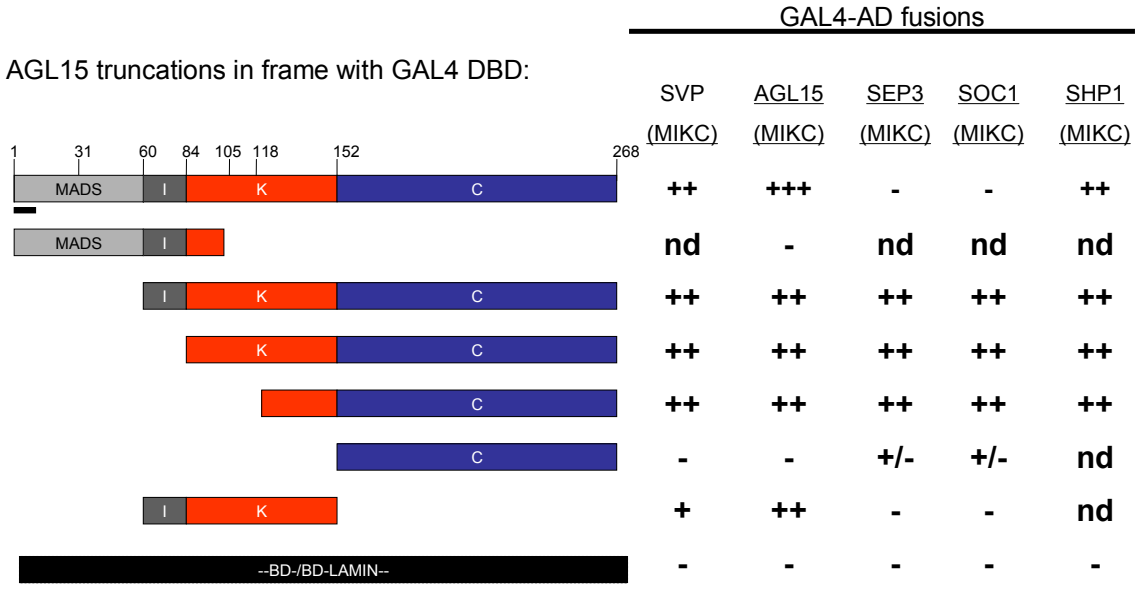
SOC1 can interact with AGL15 in yeast two-hybrid assays (Table 5.1, de Folter *et al.,* 2005). In this study no interaction was observed between full-length SOC1 and SEP3 or SHP1 (Table 3), although such an interactions have been reported by de Folter *et al.,* 2005. SOC1 interacts with SVP (Table 5.1, de Folter *et al.,* 2005), CAL, AP1 (Pelaz *et al*., 2001, de Folter *et al.,* 2005), SEP2, SEP1, AGL16, AGL14, AGL19, AGL42, AGL12, ANR, AGL17, AGL13, AGL71, AGL24, SEP4, FUL, AGL6, SHP2 (de Folter *et al.,* 2005).

5.2.3 Elucidation of the regions of AGL15 that mediate interactions with other MADS-domain proteins in yeast two-hybrid assays

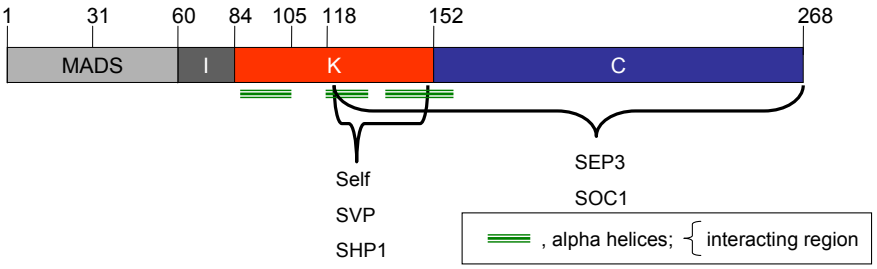
To elucidate the regions of AGL15 that mediates protein-protein interactions with other MADS-domain proteins, a series of AGL15 truncations were tested for their ability to interact with full-length MADS domain proteins in yeast two-hybrid assays (Figure 5.2a). The C-domain was required for AGL15 to interact with SEP3 and SOC1, whereas the K-domain alone permitted self-interaction and interactions with SVP and SHP1. In fact only the latter half of the K-domain (AA 118-152) seemed to be required for an interaction between AGL15 and itself, AGL15 and SVP, or AGL15 and SHP1 (Figure 5.2b). Intriguingly those able to associate with AGL15 via the K-domain alone were able to interact with full-length AGL15, whereas those whose interaction was dependent on the C-domain were not (Figure 5.2a). This pattern mirrors that observed for non-MADS interacting partners of AGL15 (Chapter 4).

Figure 5.2 Elucidation of the regions of AGL15 that mediate interactions with other MADS-domain proteins in yeast two-hybrid assays

a.



b.



a. yeast two-hybrid assays
Black bar, 10 aa; nd, not determined; +++, strong activation of reporter genes; ++, moderate activation of reporter genes; +, weak activation of reporter genes; +/-, very weak/transient activation of reporters genes; -, no activation of reporter genes (for a visual representation of how growth was scored see Figure 4.8). MIKC, full-length proteins encoded for. Results represent a minimum of three independent assays.

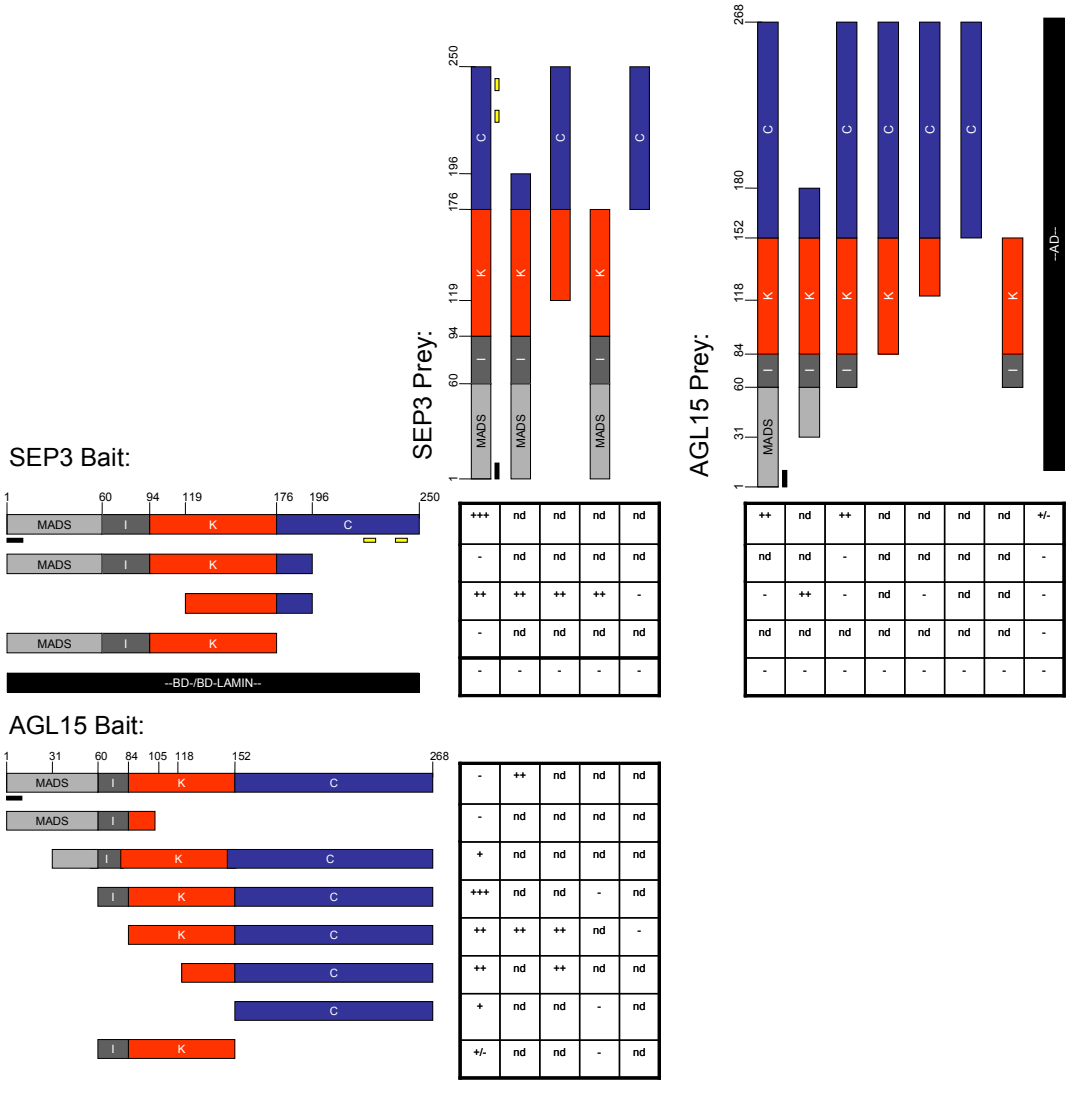
b. Schematic depicting regions of AGL15 that mediate interactions with other MADS-domain proteins in yeast two-hybrid assays

Although full-length AGL15 cannot, a truncated form of AGL15 that lacks the 60 amino acid MADS-domain can interact with full length SEP3 in yeast two-hybrid assays (Figure 5.3a). The full-length proteins can also interact with each *in vitro* when the GAL4 domains are not present (Figure 5.4). SEP3 has been previously reported to auto-activate when expressed as a GAL4-DBD fusion (Homna and Goto, 2001, de Folter *et al.,* 2005), and removal of the C-terminal domain alleviates this (Homna and Goto, 2001). Here only a weak auto-activation was observed, which was almost undetectable when all three reporter genes were selected for.

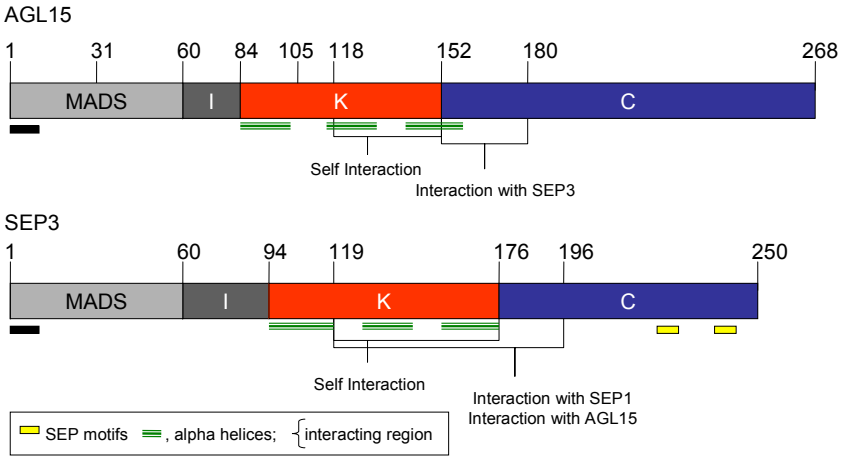
When a truncated form of SEP3 (AA 119-196) is used as bait it is able to interact with a truncated form of AGL15 (AA 31-180), but not full-length AGL15 or the IKC domains of AGL15. However, when AGL15 is used as bait, the C-terminal domain alone is enough to mediate an interaction with SEP3. This discrepancy could be due to the different differences in GAL4-AD versus GAL4-DBD fusions somehow obscuring the binding surface. However, it can be deduced that the region important for AGL15 ability to interact with SEP3 resides within the first half of the C-domain (AA 152-180), where part of the third α-helix is predicted lies (Chapter 4, Figure 4.5). An interaction between AGL15 and either SEP1 or SEP2 has not been detected in yeast two-hybrid assays.

Figure 5.3 Elucidation of the regions of SEP3 and AGL15 that meditate their interactions with each other

a.



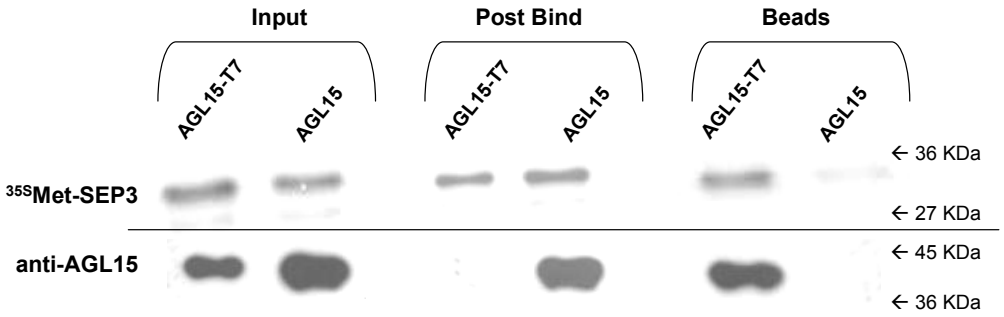
b.



a. Yeast two-hybrid assays. Black bar, 10 aa; nd, not determined; +++, strong activation of reporter genes; ++, moderate activation of reporter genes; +, weak activation of reporter genes; +/-, very weak/transient activation of reporters genes; -, no activation of reporter genes (for a visual representation of how growth was scored see Figure 4.8). Results represent a minimum of three independent assays.

b. Schematic depicting the regions of SEP3 and AGL15 that meditate their interactions with each other. The numbers at the top correspond to amino acids of AGL15. The black bar denotes 10 amino acids, the green bars the positions of the 3 α-helices (K1, K2, and K3; Yang and Jack, 2004), and the yellow bars the location of the conversed “SEP” motifs (Malcomber and Kellogg, 2005).

Figure 5.4 AGL15 and SEP3 Co-immunoprecipitate *in vitro*



Labeled SEP3 was incubated with AGL15 or AGL15-T7 and immunoprecipitated with anti-T7-agarose. Western analysis (bottom) was performed with an antibody specific to AGL15 to verify that only T7-tagged AGL15 remained with the washed beads. Radio-labeled SEP3 was co-immunoprecipitated along with T7-AGL15 (top). The sizes of the bands were determined by a standard ladder, which was run concurrently, and is indicated by the right hand arrows.

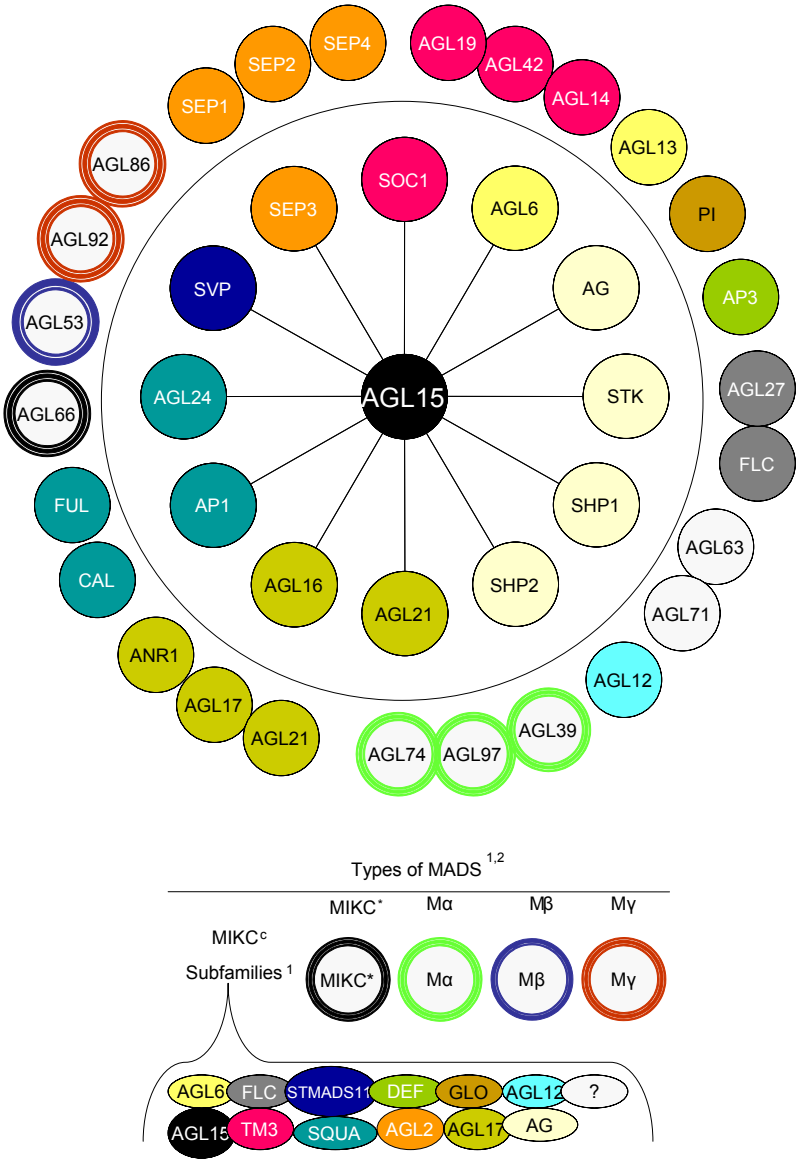
Table 5.2 displays the MADS domain proteins, including AGL15 itself, reported here and elsewhere (de Folter *et al.,* 2005) that interact with AGL15 in yeast two-hybrid assays. Of these twelve MADS box genes, transcripts corresponding to *AGL15*, *SEP3, SHP1, AGL16, SOC1* and *SVP* have been detected in embryonic culture tissue (Lehti-Shiu *et al.,* 2005). Figure 5.5 illustrates the possible higher order complexes, involving AGL15 that could theoretically occur. MADS directly interacting with AGL15 are represented in the inner circle and MADS directly interacting with any one or more MADS present in the inner circle are represented by the outer circle. Figure 5.6 combines protein-protein interaction data with expression data. MADS that are expressed in embryonic tissue culture (Lehti-Shiu *et al.,* 2005) and that interact directly or indirectly with AGL15 (this study, de Folter *et al.,* 200) are shown.

Table 5.2 AGL15 interacting MADS

AGL15 interacting MADS	MIKC^c Subfamily[3]	Function[3]	Expression in ECT[4]	Expression in Seeds[1]
AGL15[1,2]	AGL15		+	+
AGL9/SEP3[1]	AGL2/SEP	Class E floral homeotic	+	+
AGL7/AP1[2]	SQUA	Class A floral homeotic & floral meristem identity	-	nd
AGL6[2]	AGL6		-	nd
AG[2]	AG	Class C floral homeotic	-	nd
AGL11/STK[2]	AG	Ovule identity	-	nd
AGL1/SHP1[1,2]	AG	Control of fruit dehiscence zone development	+	nd
AGL5/SHP2[2]	AG	Control of fruit dehiscence zone development	-	nd
AGL16[2]	AGL17	Control of lateral root development	+	nd
AGL21[2]	AGL17		-	nd
AGL20/SOC1[1,2]	TM3	Floral promoter	+	nd
AGL24[2]	STMADS11	Floral repressor	-	nd
AGL22/SVP[1,2]	STMADS11	Floral repressor	[+]	+

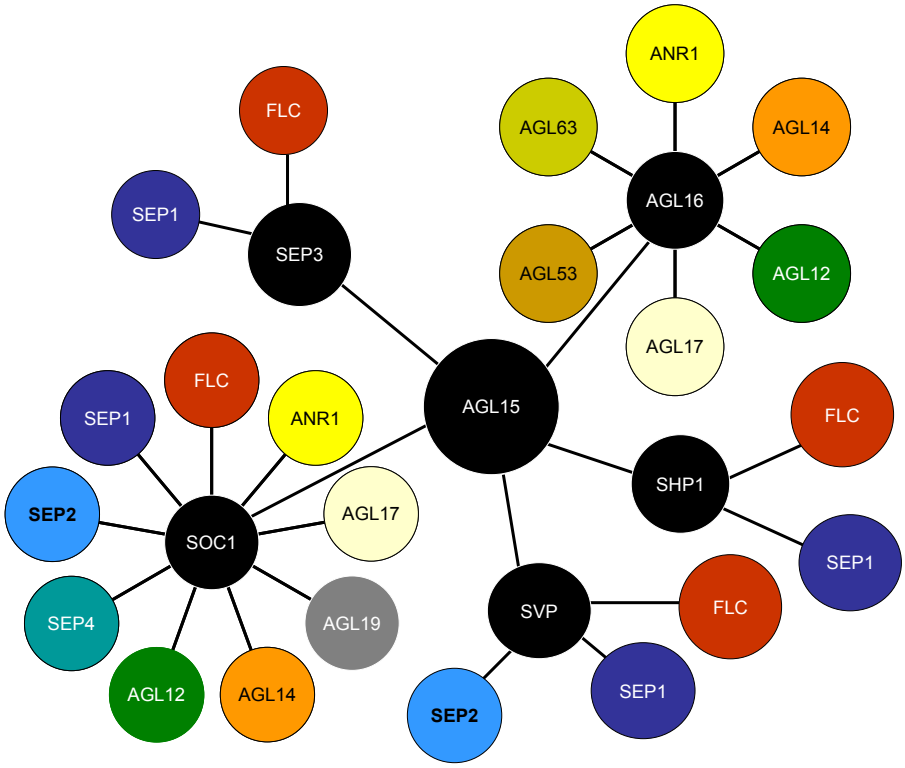
ECT, embryonic culture tissue; nd, not determined; +, transcript expressed; [+] barely detectable; -, transcript not detectable. [1]This study, [2]de Folter *et al.,* 2005, [3]Becker and Theiben, 2003, [4]Lehti-Shiu *et al.,* 2005

Figure 5.5 Interactions involving AGL15 and other MADS-domain proteins



Interactions involving AGL15 and other MADS-domain proteins presented in this study and elsewhere (de Folter *et al.,* 2005) are shown here. The inner circle represents MADS-domain proteins shown to interact with AGL15 in yeast two hybrid studies, and the outer represents MADS-domain proteins that can interact with any one or more of the inner circle AGL15-interacting MADS. [1]The types of classes and subclasses of MIKC MADS are as described by (Becker and Theiben, 2003, and [2]Parenicová *et al.,* 2003).

Figure 5.6 Interactions involving AGL15 and other MADS-domain proteins who's mRNA can be detected in embryonic tissue

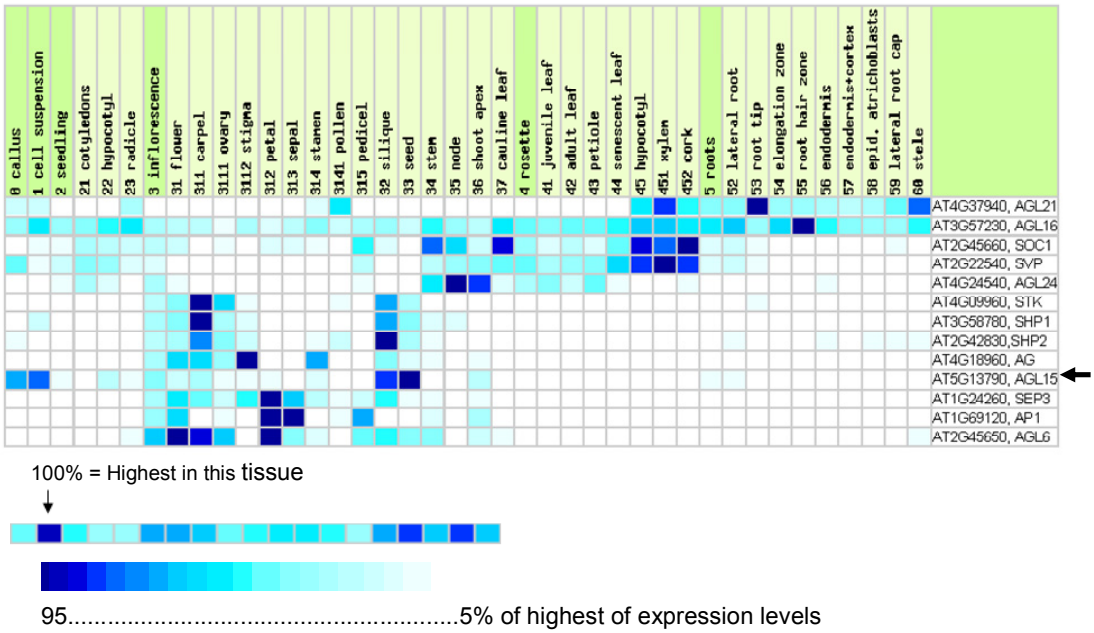


Interactions involving AGL15 and other MADS-domain proteins presented in this study and elsewhere (de Folter *et al.,* 2005) that are also expressed in embryonic tissue (according to Lehti-Shiu *et al.,* 2005) are shown. MADS domain proteins that interact with AGL15-binding MADS domain proteins (black) are color-coded to highlight those that interact with multiple partners.

5.2.4 MADS-domain proteins able to interact with AGL15 share overlapping mRNA expression patterns

AGL15 preferentially accumulates in tissue developing in an embryonic mode (Heck *et al.,* 1995; Rounsley *et al*., 1995; Perry *et al*., 1996, 1999). *SOC1, SVP1, SEP3, SVP*, and *AGL16* are all expressed in embryonic culture tissue (Lehti-Shiu *et al.,* 2005), and transcript can be detected in seed and silique (Figure 5.7). However, *AGL15* is not exclusively expressed during the embryonic phase of development but is expressed at lower levels after completion of germination in restricted sets of cells including the vegetative shoot apical meristem, leaf primordia, young flower buds, and in the base of expanding lateral organs (rosette and cauline leaves, and floral organs; Fernandez *et al.*, 2000). Expression patterns determined using the Genvetigator® (version 2) microarray database and analysis toolbox (Zimmermann *et al.,* 2004) demonstrate overlapping expression patterns between AGL15 and its interaction partners (Figure 5.7).

Figure 5.7 *In silico* mRNA expression patterns of AGL15-interacting MADS-box genes[1]



[1]Genvestigator ® (Zimmermann *et al.,* 2004).

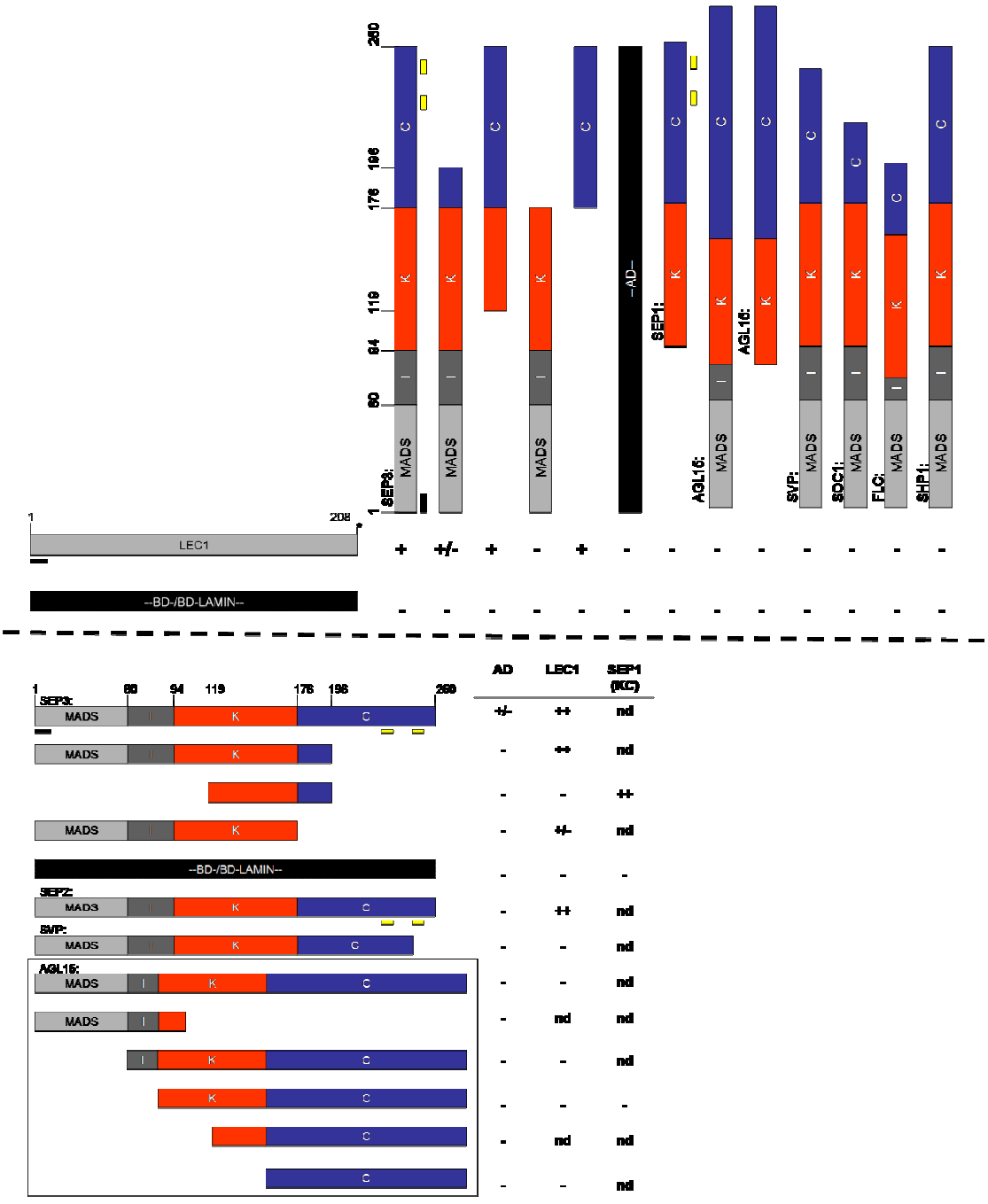
Expression patterns determined using the Genvetigator® (version 2) microarray database and analysis toolbox (https://www.genevestigator.ethz.ch/at/). The darkest blue indicates the tissue in which the highest level of said transcript is found and expression in other tissues is displayed as a percentage of this (Zimmermann *et al.,* 2004).

5.2.5 SEP3 and SEP2, but not other MADS-domain proteins, can interact with LEC1 in yeast two-hybrid assays

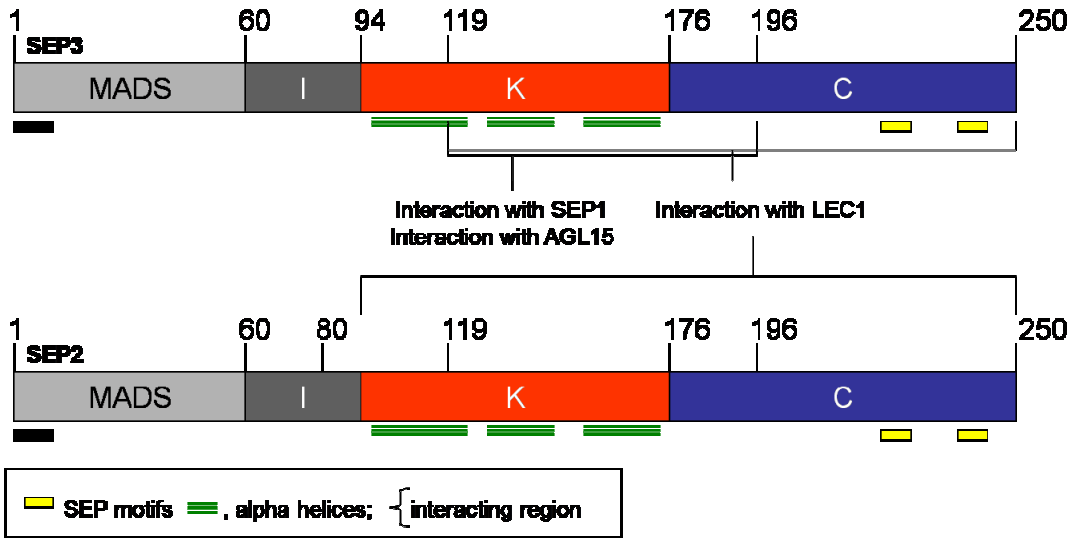
LEAFY COTYLEDON 1 (LEC1) is an important regulator of both early and late embryogenesis and is also required for somatic embryogenesis (Gaj *et al*., 2005). In rice a seed-specific NF-YB (a LEC1-like homolog) was identified as partner of OsMADS18 by two-hybrid screening (Masiero *et al.,* 2002). Therefore yeast two-hybrid assays were performed to determine if AGL15 was able to interact with LEC1. Although no interaction between LEC1 and AGL15 was detected, despite extensive testing of alternate orientations and truncations, an interaction between LEC1 and SEP3 was discovered. LEC1 interacts with SEP3 and SEP2, but not SEP1 in yeast two-hybrid assays. Neither does LEC1 interact with any of the other MADS-domain proteins tested, SVP, SOC1, FLC, and SHP1 (Figure 5.8a). The interaction between LEC1 and SEP3 is via the C-terminal domain (Figure 5.8a, b), which contains the conserved "SEP" motif (Malcomber and Kellogg, 2005)

Figure 5.8 Elucidation of the regions of SEP3 that mediate interactions with LEC1

b.



b.



a. Yeast two-hybrid assays. Black bar, 10 aa; nd, not determined; +++, strong activation of reporter genes; ++, moderate activation of reporter genes; +, weak activation of reporter genes; +/-, very weak/transient activation of reporters genes; -, no activation of reporter genes (for a visible representation of how growth was scored see Figure 2.18). Results represent a minimum of three independent assays.

b. Schematic depicting the regions of SEP that mediate interaction with LEC1.

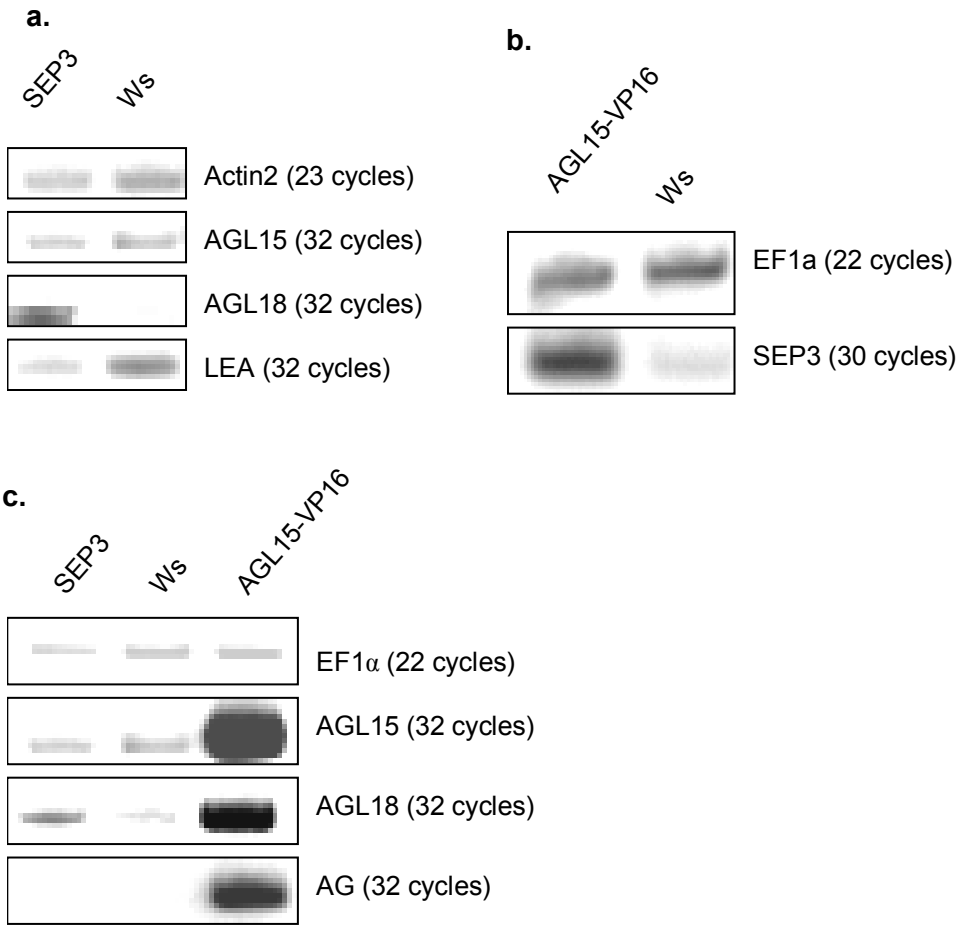
5.2.6 Phenotypes associated with ectopic expression of SEP3 resemble those caused by ectopic expression of an activated from of AGL15, AGL15-VP16

Transgenic plants carrying the *35S:SEP-c-myc* transgene were generated. Thirteen out of twenty plants displayed phenotypes previously reported for overexpression of *SEP3* (Pelaz *et al*., 2001, Homna and Goto, 2001 Castillejo *et al*., 2005), suggesting that the c-myc tag does not effect the proteins function. Phenotypes associated with plants ectopically expressing an activated form of AGL15 (*AGL15-VP16*) are small, flower early, and have narrow upward curling rosette leaves (data not shown). These phenotypes are similar to those resulting from ectopic expression of *SEP3* (Pelaz *et al*., 2001, Homna and Goto, 2001 Castillejo *et al*., 2005), and opposite to those associated with over-expression of *AGL15* (Fernandez *et al*., 2000). Research in the lab of Dr. Perry has identified a number of downstream targets of AGL15 (Wang *et al.,* 2002, Tang and Perry, 2003, Wang *et al.,* 2004, Zhu and Perry, 2005, Hill *et al*., 2007, unpublished data), and while some of these target genes are induced in response to AGL15, others are repressed. The activated form of AGL15 theoretically induces the transcription of all its direct downstream targets, including those normally repressed. However, for some direct down stream targets of AGL15 this does not appear to be the case (K. Hill, H. Wang, and S. Perry, unpublished data).

5.2.7 *AG*, *SEP3*, and *AGL18* transcript accumulates response to ectopic AGL15-VP16 and SEP3

AGL18, SEP3, and *AG* transcript abundance was increased in *35S:AGL15-VP16* seedlings (Figure 5.9b, c). *AGL18*, but not *AGL15* or *AG* transcript abundance was induced by ectopic expression *SEP3* (Figure 5.9a, c). *SEP3* has been previously reported as acting upstream of *AG* (Castillejo *et al*., 2005), so the induction of *AG* in *35S:AGL15-VP16* seedlings could be an indirect result of *SEP3* induction. Ectopic expression of *SEP3* results in the induction of *AP3* and *AG* expression in rosette leaves (Castillejo *et al*., 2005). However, activation of AG was not detected in 6 day *35S:SEP3-c-myc* seedlings (Figure 5.9a) as previously reported for *35S:SEP3* rosette leaves (Castillejo *et al*., 2005) suggesting the presence or absence of a factor in seedlings that is present or absent in leaves. *LEA76* is a direct downstream target of AGL15 (see Chapter 2). Ectopic expression of *SEP3* also causes a reduction in *LEA76* transcript accumulation, perhaps via induction of AGL18.

Figure 5.9 Changes in *AGL18*, *AGL15* and *AG*, transcript levels in response to SEP3 and AGL15-VP16



RT-PCR was performed on 6 day old seedlings. Parenthesis indicate the number of PCR cycles.

a. *AGL18 and AG* transcript abundance was induced by *AGL15-VP16*. *AGL18, but not AGL15 or AG* transcript abundance was induced by *SEP3*.
b. *SEP3* transcript abundance was induced by *AGL15-VP16*.
c. *AGL18, but not AGL15,* transcript abundance was induced by *SEP3. LEA, but not AGL15,* transcript abundance was repressed by *SEP3.*

5.2.8 Novel phenotypes are seen in plants carry both *35S:SEP3* and *35S:AGL15* transgenes

Pollen from a severe *SEP3* over-expressor line was dusted onto a plant hemizygous for the *35S:AGL15* transgene, because the homozygous plants do not set seeds. As expected, all the F1 progeny were Kanamycin resistant. One out of six displayed the typical phenotype associated with *SEP3* over-expression and PCR analysis confirmed it carried only the *35S:SEP3* transgene. Four out of six plants looked like typical *AGL15* over-expressers. One out of the 6 displayed a phenotype not seen in either of the single over-expressors (Figure 5.10), and PCR analysis confirmed that this plant carried both transgenes (data not shown). This plant, although it was larger and flowered later than its *35S:SEP3* sibling was considerably shorter in stature and flowered much earlier than either Ws or *35S:AGL15* plants. Like *35S:SEP3* plants it had curled up leaves, and like *35S:AGL15* plants its sepals and petals remained attached to the base of the siliques. Like *35S:SEP3* plants, the inflorescence terminated in a single, or "bunch" of siliques, but more siliques than were produced by the *35S:SEP3* plant.

The progeny of plants hemizygous for *35S:AGL15* and *35S:SEP3* segregate into five distinct phenotypes, perhaps reflecting the relative dosage of the respective transgenes. The phenotypes observed were "SEP3-Like", resembling the phenotype of *35S:SEP3* parent, a severe *35S:SEP3* phenotype, that failed to set seed, and *"AGL15-Like"*, resembling *35S:AGL15* hemizygous plants; a severe *35S:AGL15* phenotype, resembling a *35S:AGL15* homozygous plant, and an intermediate phenotype typified in Figure 5.10.

Figure 5.10 Phenotype of a *35S:SEP3* x *35S:AGL15* F1 Plant

35S:SEP3 35S:AGL15

F1 Plant

F1 plants carrying both *35S:SEP3-myc* and *35S:AGL15* transgenes. Parent plants exhibited typical phenotypes associated with ectopic expression of the SEP3 (Pelaz *et al.,* 2001, Honma & Goto, 2001 Castillejo *et al*., 2005) and AGL15 (Fernandez *et al.,* 2000).

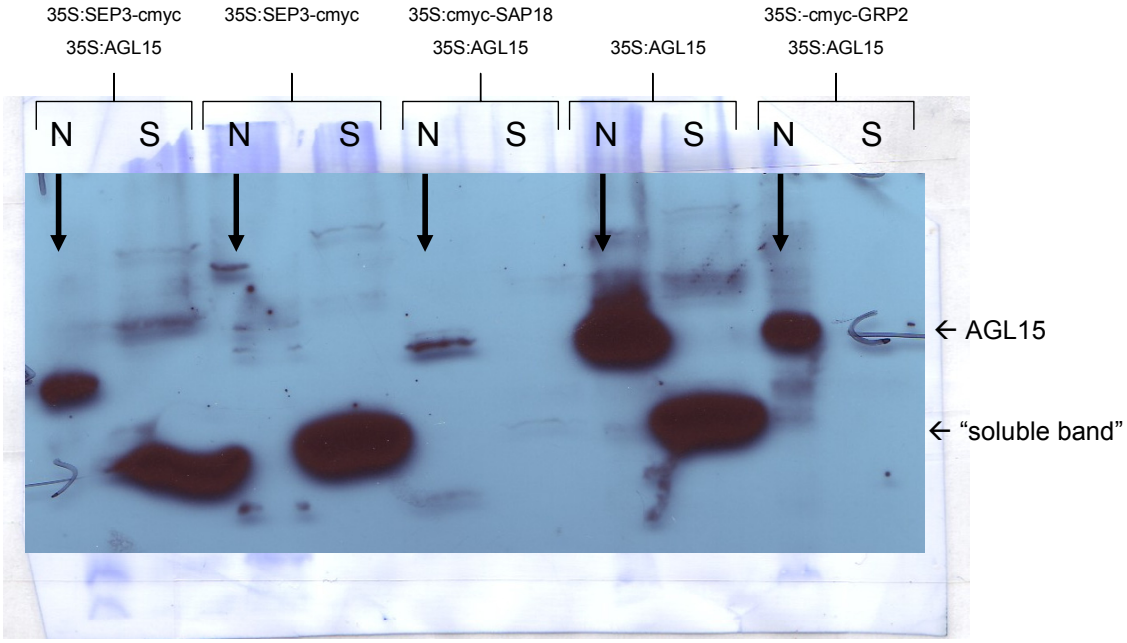
5.2.9 Embryonic culture tissue (ECT) carrying the *35S:SEP3-myc* transgene do not require a high level of AGL15

In order to generate enough tissue for future co-immunoprecipitation experiments, embryonic culture tissue (ECT) was generated from plants carrying both the *35S:AGL15* transgene, which increases induction and maintenance of development in an embryonic mode (Harding *et al.,* 2003), and *35S:SEP3*. Surprisingly the *35S:SEP3* transgene by itself was found to be sufficient for induction and maintenance of ECT. Although ECT can be initiated from wild type plants, it cannot be maintained beyond a few months (data not shown).

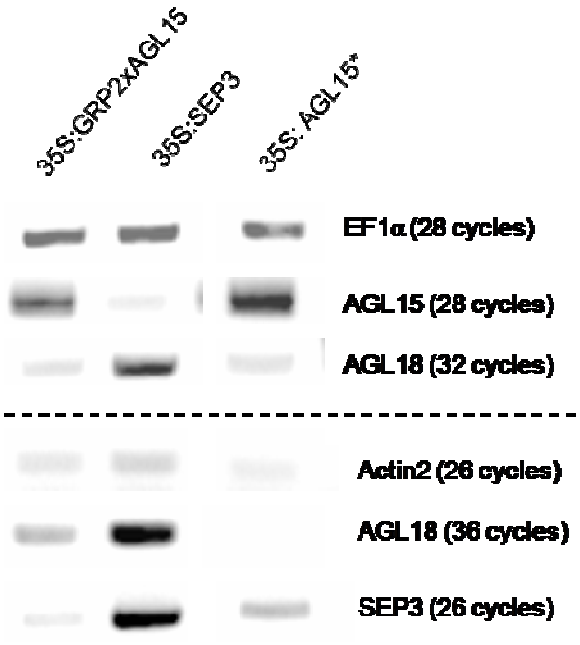
Like *35S:AGL15*, *35S:SEP3* ECT maintains development in the embryonic mode (nine months to date) without the addition of hormones. AGL15 protein accumulation was not detected in nuclear preps from *35S:SEP3* ECT (Figure 5.11a). However, protein accumulation could be detected with another AGL15-antibody, known to cross-react with AGL18 (data not shown). RT-PCR experiments reveal that, compared to established *35S:AGL15* ECT (Harding *et al.,* 2003) and non-*35S:SEP3* ECT started at the same time, *35S:SEP3* ECT accumulates *AGL18* transcript. Ectopic SEP3 also causes an accumulation of *AGL18*, but not *AGL15* transcript in seedlings (Figure 5.11b). AGL18 is closely related to AGL15 and also increases induction and maintenance of development in the embryonic mode (D. Fernandez, personal communication).

Figure 5.11 AGL15 RNA and protein accumulation in embryonic culture tissue (ECT) initiated from *35S:SEP3* plants

a.



b.



a. Western analysis of nuclear preps made from embryonic culture tissue.
N, Nuclear sample S; sample from the nuclear depleted “soluble” faction;
The “AGL15” band corresponds to AGL15, and the “Soluble” band is caused by a cross reaction with an unrelated protein present in the nuclear depleted fraction.

b. RT-PCR on embryonic culture tissue
Cultures were started on the same day from plants carrying the *35S:GRP2 and 35S:AGL15* transgenes or the *35S:SEP3-cymc* transgene. AGL15* is established ECT, started at a much earlier date.

5.3 Discussion

5.3.1 Putative higher order MADS-domain protein complexes involving AGL15

All AGL15 interacting partners are MIKC^c type (Table 5.2). AGL15 likely participates in a number of complexes in a redundant fashion with AGL18 and other MADS-domain proteins. Some of the phenotypes associated with ectopic expression AGL15 (Fernandez *et al.,* 2000) might also be due to its associating with complexes that might not form under native expression.

The most studied example of the combinatorial nature of MADS-domain protein pertains to the floral organ identity. The quartet model is the revised "ABC" model of floral development, which makes predictions about the composition of the tetramers in the four whorls of the flower (for review see Jack, 2004). In the quartet model a combination of AP3/PI-SEP/AP1 is postulated to specify petals in whorl 2, a combination of AP3/PI-SEP/AG to specify stamens in whorl 3, and a combination of AG/AG-SEP/SEP to specify carpels in whorl 4 (Figure 1.4). Figure 5.5 illustrates the possible higher order complexes involving AGL15 that could theoretically occur. MADS-domain proteins that can interact with AGL15 are represented in the inner circle and MADS-domain proteins that can interact with any one or more MADS present in the inner circle are represented by the outer circle. Figure 5.6 combines the protein-protein interaction data, presented here and elsewhere (de Folter *et al.,* 2005), and shows only those MADS-box genes known to be expressed in embryonic tissue culture (Lehti-Shiu *et al.,* 2005). Based on Figure 5.6, it could be hypothesized that higher order complexes involving AGL15, FLC, and SEP, might be important for embryonic development. It would be interesting to determine what embryogenesis-related phenotypes, if any, triple *agl15/agl18/sep* or quadruple *agl15/agl18/flc/sep* mutants might exhibit. However, there is a high level of redundancy among the four SEP proteins (Pelaz *et al.,* 2000), and null alleles of *sep3* have very subtle phenotypes (Ditta *et al*., 2004). Thus it seems unlikely that *sep3* alone will contribute a phenotype in the aforementioned higher order mutants.

AGL15 preferentially accumulates in a wide variety of tissues that are developing in an embryonic mode (Heck *et al.,* 1995; Rounsley *et al*., 1995; Perry *et al*., 1996, 1999). AGL15 is not exclusively expressed during the embryonic phase of development but is

expressed at lower levels after completion of germination in restricted sets of cells including the vegetative shoot apical meristem, leaf primordia, young flower buds, and in the base of expanding lateral organs (rosette and cauline leaves, and floral organs; Fernandez *et al.*, 2000). Expression patterns determined using the Genvestigator® (version 2) microarray database and analysis toolbox (Zimmermann *et al.,* 2004) demonstrate overlapping expression patterns between AGL15 and its interaction partners (Figure 5.7). Diagrams similar to Figure 5.6 could be drawn from expression data relating to any developmental stage or environmental response and higher order complexes hypothesized.

5.3.2 SEP3 can interact with LEC1 in yeast two-hybrid assays

SEP3 exhibits transactivation activity in yeast assays (Honma and Goto, 2001, supplementary data de Folter *et al.,* 2005, K. Hill, unpublished observation) and could explain how AGL15, a protein not known to contain an activation domain, might activate transcription of target genes *in vivo*. SEP3 might also function as a bridge, recruiting AGL15 higher order complexes. SEP3 is able to bind to and LEC1 in yeast two-hybrid assays (Figure 5.8a), and AGL15 in yeast and *in vitro* (Figure 5.3a, Figure 5.4), and all three genes are expressed in embryonic tissue culture (Lehti-Shiu *et al.*, 2005) suggesting that complex containing AGL15 and LEC1 might exist *in vivo*. LEC1 is an important regulator of both early and late embryogenesis and is also required for somatic embryogenesis (Gaj *et al*., 2005). LEC1 interacts with SEP3 and SEP2, but not SEP1 in yeast two-hybrid assays (Figure 5.8a). Neither does LEC1 interact with any of the other MADS-domain proteins tested, SVP, SOC1, FLC, and SHP1 (Figure 5.8a). OsNF-YB, a rice LEC1-like protein, which contains a histone fold motif and is part of the trimeric CCAAT-binding NF-Y complex, does not directly bind to the MADS-domain protein, OsMADS6, but does interact, via OsMADS18, in yeast three-hybrid assays (Masiero *et al.,* 2002). The interaction between LEC1 and SEP3 is via the C-terminal domain (Figure 5.8b), which contain conserved “SEP” motifs (Malcomber, and Kellogg, 2005).

In *lec1* mutants, immature embryos precociously enter a germination pathway after the torpedo stage of and acquire characteristics normally restricted to vegetative parts of the plant (Meinke *et al.,* 1994). *Ectopic* expression of AGL15 and *LEC1* enhances the induction of somatic embryogenesis in the absence of hormones (Harding *et al.,* 2003,

Lotan *et al*., 1998). No direct interaction between AGL15 and LEC1 was observed in yeast two-hybrid assays, but the possibility of a complex involving SEP, AGL15 and LEC1, forming in planta and regulating embryogenesis is particularly exciting.

5.3.3 Phenotypes associated with ectopic expression of SEP3 and AGL15-VP1 may be due to regulation of a similar subset of genes

Phenotypes associated with plants ectopically expressing an activated form of AGL15 (AGL15-VP16) are small, flower early, and have narrow upward curling rosette leaves (data not shown). These phenotypes are similar to those resulting from ectopic expression of SEP3 (Pelaz *et al*., 2001, Homna and Goto, 2001 Castillejo *et al*., 2005) and opposite to those associated with over-expression of AGL15 (Fernandez *et al.,* 2000). Research conducted in the lab of Dr. Perry has identified a number of downstream targets of AGL15 (Wang *et al.,* 2002, Tang and Perry, 2003, Wang *et al.,* 2004, Zhu and Perry, 2005, Hill *et al.,* 2007, unpublished data), and while some of these target genes are induced in response to AGL15, others are repressed. The activated form of AGL15 theoretically activates the transcription of all its direct downstream targets, including those normally repressed. However, this appears not to always be the case (K. Hill and S. Perry, unpublished observation). The phenotype of the *AGL15-VP16* plants could be due to up regulation of *SEP3* (Figure 5.9b), which may be direct or indirect, such as via up regulation of *AG* (Figure 5.9c). *SEP3* likely acts upstream of *AG*, because *35S:SEP3* is unable to rescue the *ag* mutant phenotype, and ectopic expression of *SEP3* results in the induction of *AP3* and *AG* expression in rosette leaves (Castillejo *et al*., 2005). However, activation of *AG* was not detected in 6 day *35S:SEP3-c-myc* seedlings (Figure 5.9c).

A complex picture of MADS-domain proteins regulating the transcription of MADS-box genes is emerging and activation of *AG* could account for a number of phenotypes associated with *35S:SEP3* and *35S:AGL15-VP16*. *AG* up regulates itself (Yanofsky *et al*., 1990, Gómez-Mena *et al*., 2005), and the MADS-box genes, *SEP3*, *SEP2*, *SEP1* (Pelaz *et al*., 2000, Gómez-Mena *et al*., 2005) *AP3* (Jack *et al*., 1992, Gómez-Mena *et al*., 2005), *PI* (Goto and Meyerowitz, 1994, Gómez-Mena *et al*., 2005), *SHP1*, *SHP2* (Liljegren *et al*., 2000, Gómez-Mena *et al*., 2005), *AGL15* (Gómez-Mena *et al*., 2005), and binds the promoter of *AG*, *SEP3* and *AP3 in vivo* (Gómez-Mena *et al*., 2005). AG

also up regulates *SUP* (Sakai *et al*., 1995, Gómez-Mena *et al*., 2005) and *CRABS CLAWS* (*CRC*; Bowman and Smyth, 1999, Gómez-Mena *et al*., 2005), and directly binds the promoter of *CRC* (Gómez-Mena *et al*., 2005). *AG* and *SHOOTMERISTEMLESS* (*STM*) are direct targets of, and repressed by the plant Polycomb-group (Pc-G) protein *CURLY LEAF* (*CLF*; Schubert *et al.,* 2006). *AG*, *AP3*, *PI* and *SEP3* are highly expressed in *terminal flower 2 (tlf2)* mutants that miss-express *FT* (Kotake *et al*.,2003) and expression of *SEP3* in leaves is dependent on FT levels (Teper-Bamnolker and Samach, 2005). *35S: FT* plants also flower early and exhibit leave curling when grown under blue light, and a much higher degree of leave curling occurs under long day conditions compared to short day (Teper-Bamnolker and Samach, 2005). *SEP3* expression correlates with leaf curling in transgenic *35S:FT/TFT* plants, and a mutation in *SEP3*, but not *AP1*, suppress leaf curling in *35S:FT* plants, even when the mutant allele, *sep3-2* is heterozygous (Teper-Bamnolker and Samach, 2005). High levels of SEP3 alone are not sufficient for severe curling under normal growth conditions, suggesting that *SEP3* is not the only component required for leaf curling (Teper-Bamnolker and Samach, 2005). However, introducing even higher levels of SEP3 into *35S :FT* plants increases the degree of curling, suggesting that SEP3 levels are rate-limiting in leaf curling (Teper-Bamnolker and Samach, 2005).

No obvious phenotype results from plants carrying null alleles of *agl15*, but *agl15/agl18* double mutant flower early under short days (Adamczyk *et al.,* 2007). While no late-flowering phenotype was described for the *sep3-2* mutant in a wild-type background, loss of SEP3 caused a slight yet significant delay in the *35S:FT* background (Teper-Bamnolker and Samach, 2005). *FT* mRNA is expressed in the leaves in response to photoperiod and expression is increased in long days (Corbesier *et al.,* 2007), which might account for the early flowering of *agl15/agl18* in short days but not long days (Adamczyk *et al.,* 2007).

Plants expressing the *35S:SEP3* transgene (Pelaz *et al*., 2001, Homna and Goto, 2001 Castillejo *et al*., 2005) or the *35S:AP1* transgene flower considerably earlier than wild-type plants and these early flowering phenotypes are dramatically enhanced in plants containing both the *35S:SEP3* and *35S:AP1* transgenes (Pelaz *et al*., 2001, Homna and Goto, 2001 Castillejo *et al*., 2005). An activated form of AGL15, fused to the strong VP16 activation domain also flowers early. In contrast, ectopic expression of *AGL15*

causes a late flowering phenotype (Fernandez *et al.,* 2000), and the double mutant *agl15/18* is early flowering under short days (Adamczyk *et al.,* 2007). *35S:SEP3/35S:AGL15* plants flower early, but later than *35S:SEP3*. *svp* mutant plants are severe early flowering but floral development is normal (Hartmann *et al*. 2000). agl24-2 mutant plants are late flowering floral development is normal (Michaels *et al.,* 2003). *svp/agl24* double mutants are also severe early flowering, and display mild floral defects at 20ºC and strong floral defects at 30ºC (Gregis *et al*., 2006). Loss of *SVP* activity significantly decreases the plants response to lower temperature (Lee *et al.,* 2007). SVP controls flowering time by negatively regulating the expression of a floral integrator, *FLOWERING LOCUS T* (*FT*), via direct binding to the CArG motifs in the *FT* sequence.

It would be interesting to determine the flowering phenotype of an agl15/agl18/sep triple or quadruple mutant, and if the AGL15 OE phenotype was as pronounced in sep3, or sep1/2/3 background, or indeed if genes normally activated in response to AGL15 still accumulated.

5.4 Materials and methods

5.4.1 Yeast two-hybrid assays

For directed tests AH109 cells were co-transformed with specific bait and prey constructs and plated onto SD media (2% dextrose, 0.67% nitrogen base, 1% agar) lacking leucine (-L), tryptophan (-W) (SD-L/-W). After several days transformed AH109 yeast colonies were transferred onto to plates lacking histidine (-H), and alanine (-A) and supplemented with 0.2mg/ml X-α-gal (SD-L/-W/-H/-A, 0.2mg/ml X-α-gal) to assay for transformants that are able to activate the reporter genes (*HIS3*, *ADE2*, and *MEL1*). For a detailed description see Chapter 4.

Bait constructs used in this study:

DBD-AGL15 (MIKC), DBD-AGL15 (MIK½), DBD-AGL15 (½MIKC), DBD-AGL15 (IKC), DBD-AGL15 (KC), DBD-AGL15 (½KC), DBD-AGL15 (C), DBD-AGL15 (IK), DBD-SEP3 (MIKC), DBD-SEP3 (MIKC½), DBD-SEP3 (½KC½), DBD-SEP3 (MIK), DBD-SEP2 (MIKC), DBD-SVP (MIKC), DBD-SVP (IKC), DBD-AGL18 (MIKC), DBD-AGL18 (IKC), AGL18 (KC), DBD-SHP1 (MIKC), DBD-FLC (MIKC), DBD-FLC (IKC), DBD-SOC1 (MIKC), DBD-SOC1 (IKC), DBD-PI (MIKC), and DBD-LEC1.

The oligonucleotides and restriction sites used generate the pGBKT7 (bait) constructs can be found in Appendix A.

Prey constructs used in this study:

AD-AGL15 (MIKC), AD-AGL15 (½MIKC½), AD-AGL15 (IKC), AD-AGL15 (KC), AD-AGL15 (½KC), AD-AGL15 (C), AD-AGL15 (IK), AD-SEP3 (MIKC), AD-SEP3 (MIKC½), AD-SEP3 (½KC), AD-SEP3 (MIK), AD-SEP3 (C), AD-SEP1 (KC), AD-SVP (MIKC), AD-SVP (IKC), AD-SOC1 (MIKC), AD-SHP1 (MIKC), AD-PI (MIKC), AD-AGL18 (MIKC), AD-AGL18 (IKC), AD-AGL18 (KC), AD-FLC (MIKC), AD-FLC (IKC) and AD-LEC1.

The oligonucleotides and restriction sites used generate the pGADT7 (prey) constructs used in this study can be found in Appendix B.

All constructs were verified by DNA sequencing carried out by UK-AGTC. Protein accumulation in yeast cells was determined via Western blotting, using AGL15 antibody (Heck *et al.,* 1995), c-myc monoclonal antibody or HA polyclonal antibody (Clontech, Mountain View, CA).

5.4.2 *In vitro* co-immunoprecipitation

Full length AGL15 and Full-length AGL15 carrying a C-terminal T7 tag (MASMTGGQQMG) were cloned into an expression vector pET-15b (Novagen, San Diego, CA). The *E.coli* strain BL21$^{(DE3)}$ (Novagen, San Diego, CA) was used to express the proteins, and inclusion bodies were harvested and solublized in 8 M urea, in wash-bind buffer (4.29mM Na_2HPO_4, 1.47mM KH_2PO_4, 2.7mM KCl, 0.137M NaCl, 0.1% Tween-20, 0.002% sodium azide, pH 7.3) and dialyzed in 2 M increments (6M, 4M, 2M, 0M urea) for two hours each. Total protein concentration was determined from a standard curve, which was generated by measuring the OD595 of known concentrations of BSA (1-10 µg/,ml) incubated for 5 minutes at room temperature in Bradford Reagent (Biorad, Hercules, CA).

Radiolabeled SEP3 was generated from 1 µg of RNAase free pAD-SEP3 (MIKC) using the TNT®Quick Coupled Transcription/Translation System (Promega, Madison, WI) as directed. SEP3 was translated in the presence of 10-40 µ Ci of Amersham Biosciences Redivue™ L-[^{35}S] methionine (Amersham Biosciences, Pittsburg, PA) in a total volume of 50 µl, at 30 °C for 1 hour.

To determine the percentage of label incorporation 2 µl aliquots of translation mix were dropped onto blotting paper (~ 2 cm^2), and placed directly into scintillation cocktail to be ("unwashed"). A second 2 µl aliquot was dropped onto another piece of blotting paper and washed for several minutes in cold 10% TCA (trichloroacetic acid), washed twice in boiling 10% TCA for 2 minutes, followed by a final wash in 95% EtOH. The dried sample was and place into scintillation cocktail ("washed"). The samples were read in a Packard Tri-Carb model 1500 liquid scintillation counter (PerkinElmer, Waltham, Massachusetts)

The percent incorporation calculated by dividing the cpm of the washed by that of the unwashed [(cpm of washed filter/cpm of unwashed filter) x 100 = percent incorporation].

The efficiency of incorporation for the experiment displayed in Figure 2.10 was 9.8%, but anywhere between 4-18% is typical. TNT® Quick Master Mix contains roughly 100-200mg/ml endogenous proteins

1 µg of AGL15 (MIKC), or tagged AGL15 and 0.4 µg of ^{35}S-metionine labeled SEP3 (~ 40 ng labeled protein) were incubated together in a total volume of 120 µl 2 M urea in wash-bind buffer (4.29mM Na_2HPO_4, 1.47mM KH_2PO_4, 2.7mM KCl, 0.137M NaCl, 0.1% Tween-20, 0.002% sodium azide, pH 7.3). The samples were placed on a rota for 30 minutes at room temperature. For the "Total" sample 5 µl of sample loading buffer (2% SDS, 0.001% bromophenol blue, 0.2M DTT, 20% glycerol, 20mM sodium phosphate, pH7.2) was added to 5µl of sample and boiled for 5 minutes. The samples were centrifuged for 5 min at 16,000 xg and the supernatant transferred to a fresh Eppendorf tube containing 20 µl T7 Tag Antibody Agarose Beads (Novagen, Madison, WI). Following 1 hour incubation on a rota at room temperature, the beads were spun down at 500xg for 10 minutes. For the post bind sample 5 µl of 4XSB was added to 5µl of supernatant and boiled for 5 minutes. The beads were washed four times with 1X wash bind buffer (4.29mM Na_2HPO_4, 1.47mM KH_2PO_4, 2.7mM KCl, 0.137M NaCl, 0.1% Tween-20, 0.002% sodium azide, pH 7.3) and suspended in 100µl 1XSB and boiled for 5 minutes.

Proteins were separated on two 12.5 % (w/v) polyacrylamide denaturing gels using vertical gel apparatus (Hoefer Scientific Instruments, San Francisco, CA). To verify that the T7 Tag Antibody Agarose Beads pulled down only T7 tagged AGL15 on of the gels was blotted onto Immobilon™ PVDF Transfer Membranes (Millipore, Billerica, MA) using a Genie blotter (Idea Scientific Co., Minneapolis, MN). The membrane was blocked with 5% non-fat dry milk, washed in TBST (0.1% Tween 20, 100mM Tris-Cl, 150mM NaCl, pH 7) and probed with 1:1,000 diluted AGL15 antiserum (Heck *et al.*, 1995; Perry *et al.*, 1996, 1999, Wang *et al.,* 2000), followed by 1:5,000 diluted secondary antibody (HRP-conjugated goat-anti-rabbit). Finally the membrane was incubated for 1 minute in substrate (0.5M Tris, 0.03% H_2O_2, 0.1mg/ml Luminol, 0.4 mg/ml P-idophenol) before being exposed (0.5-5 minutes) to Kodak XAR5 X-ray films (Eastman Kodak, Rochester,

NY). To detect radio-labeled SEP3 the gel was dried for ~2hrs at 80ºC under vacuum and exposed to X-ray film for 24-48hrs at -80ºC. The films were developed in a Konica film processor (SRX-101, Konica Corp., Tokyo, Japan).

5.4.3 Generation and Analysis of 35S:SEP3-myc *Arabidopsis* plants

cDNA coding for full-length *SEP3* (At1g24260) was PCR amplified with oligonucleotides containing *Xba1* and *Xho1* restriction enzyme sites (underlined) and a c-terminal c-myc epitope (MEEQLISEEDLHM, double underlined):

Forward 5'GCTCTAGA ATG GGA AGA GGG AGA GTA GAA TTG AAG3'
Reverse
5'CCTCGAGGCAGGTCCTCCTCTGAGATCAGCTTCTGCTCCTCAATAGAGTT GGTGTC3'

35S:SEP3-c-myc was cloned into the MCS site of pBIMC (a gift from Dr. D. Falcone, University of Massachusetts) downstream of the *35S* CaMV promoter.

The construct was checked by sequencing and then transformed into the *Agrobacterium tumefaciens* strain GV3101, and used to transform Ws flower buds using the floral dip method (Clough and Bent, 1998). Transgenic T1 seeds selected on GM plates (Murashige-Skoog (MS) salts, vitamins, 1% (w/v) sucrose, 0.05% MES (w/v), and 0.7% agar (w/v), pH 5.6-5.7) containing 50ug/ml Kanamycin. Putative SEP3-cymc over-expressing lines were indentified by their typical *SEP3* over-expression phenotype (Pelaz *et al.,* 2001, Homna and Goto, 2001, Castillejo *et al.,* 2005). *SEP3* transcript levels were also analyzed by RT-PCR, and protein accumulation verified by Western analysis using an anti-c-myc monoclonal antibody (Sigma-Aldrich, St. Louis, MO).

Plant growth conditions

Seeds of *A. thaliana* ecotype Wassilewskija (Ws) or Columbia (Col) were sterilized by 3-4 brief washes with 70% ethanol containing 0.1% Triton X-100, followed by two rinses with 95% ethanol, then allowed to dry in a sterile laminar flow hood before sprinkling on GM (germination media) plates containing Murashige-Skoog (MS) salts, vitamins, 1%

(w/v)sucrose, 0.05% MES(w/v), and 0.7% agar (w/v), pH 5.8. Seeds were chilled for several days at 4°C before being grown at 20/18°C under a 16-h light/8-h dark regime. Seedlings were transplanted at 8–10 days to ProMix BX (Premier Brands, Inc., Riviere-du-Loup, Quebec, Canada) and grown in a Conviron growth chamber with fluorescent and incandescent lights set to approximately 200 µmol m^{-2} sec^{-1}.

Crosses

Stamens were removed from unopened flower buds of the maternal plant (35S:AGL15) and pollen from the paternal plant (35S:SEP3) dusted onto the exposed carpel twice a day for the next 3 days.

Genotyping of transgenic plants: DNA isolation and PCR analysis:

One young rosette or cauline leaf was ground in liquid nitrogen and suspended in 500 µl shorty buffer (0.2M Tris-HCL pH9, 0.4M LiCl, 25mM EDTA, 1% SDS). Following extraction with phenol: chloroform: isoamyl alcohol (25:24:1) an equal volume of 100% isopropanol was added and the sample incubated for 10-15 minutes. The sample was then centrifuged at 16,000xg for 10 minutes at room temperature. The pellet was washed once with 75% ethanol, thoroughly dried, and re-suspended in 30 µl mili-Q H_2O. 1µl was taken per PCR reaction and added to 1x PCR buffer, dNTPs at 200 µM each, 0.2 µM of each primer and 1.0 unit of KlenTaq1 (Ab Peptides, St. Louis, MO) in a final volume of 20 µl. Amplification reactions were performed in a PCT-100 (MJ Research Inc., Watertown, MA), or a AB 2720 (Applied Biosciences, Oakville, Ontario, Canada) or 32-40 cycles of denaturation at 95°C (30 sec), annealing at 55°C (30 sec), and extension at 72°C (30 sec). Oligonucleotide pairs contained one oligonucleotide specific the *35S* promoter of the transgene:

Forward 5’ TCGGATTCCATTGCCCAG 3’

and another specific to the respective transgenes:

SEP3, Reverse, 5’ GCT TCA AGG AAG AAT CAA GC3’
AGL15, Reverse, 5’ TTC AAG TTG GTT AGT CAG CAA TCG’3

5.4.4 Semi-quantitative RT-PCR

Either TRIZOL Reagent (Invitrogen, Calsbad, CA) or the RNeasy® plant mini kit (Qiagen, Valencia, CA) was used to isolate total RNA from ~50 mg of whole 6-8 day old seedlings, grown on GM media. 1.0 µg of total RNA was first treated with DNase I (Invitrogen, Calsbad, CA) and then used for first strand cDNA synthesis. Reverse transcription was performed using A-MLV Reverse Transcriptase System (Promega, Madison, WI) according to the manufacturer's instructions. 1-2 µl aliquot of each first strand cDNA reaction was amplified by specific primer pairs in a reaction containing 1x PCR buffer, dNTPs at 200 µM each, 0.2 µM of each primer and 1.0 unit of KlenTaq1 (Ab Peptides, St. Louis, MO) in a final volume of 20 µl. Amplification reactions were performed in a PCT-100 (MJ Research Inc., Watertown, MA), under conditions that varied only in the number of cycles of denaturation at 95°C (30 sec), annealing at 55°C (30 sec), and extension at 72°C (30 sec).
Control oligonucleotides specific for "house-keeping" genes were used as controls:

EF1α (At1g07920)
5'ACGCTCTACTTGCTTTCACC3'
5'GCACCGTTCCAATACCACC3'
Actin2 (At3g18780)
5'GAGACCTTTAACTCTCCCGCTATG3',
5'GAGGTAATCAGTAAGGTCACGTCC3'

Oliogonucleotides specific to the respective target genes are as follows:

SEP3 (At1G24260)
Forward 5'TCAAGAGAGGCCTTAGCAGT3'
Reverse 5'GCTTCAAGGAAGAATCAAGC3'
AGL15 (At5g13790)
Forward 5'TCCAAGAGGCGTTCTGGGTTACTT3'
Reverse 5'CTGCTCAAGGCTTTGCAGCTCTTT3'
AGL18 (At3g57390)
Forward 5'ACACTACTGCGTCCACTGAGCATA3'

Reverse 5'AGAAGCCACTTGACTCCCAGAGTT3'

AG (At4G18960)

Forward 5'ATGCTGAAGTCGCACTCATCGTCT3'

Reverse 5'TGCCTTCCAAGTTCCTGAGCTCTT3'

LEA (At1g52690)

Forward 5'TAGGGCTTCGCACTGATGAAGGAA3'

Reverse 5'GGCATAACCTCACGAACGCAACAA3'

5.4.5 Generation of *Arabidopsis* embryonic culture tissue (ECT)

35S:AGL15 induces and maintains development in the embryonic mode without hormones (Harding *et al.,* 2003). In order to obtain large quantities of *Arabidopsis* embryonic tissue, embryonic culture tissue (ECT) was generated as described by Harding *et al.,* 2003. Briefly, developing zygotic embryos from the siliques of plants carrying the 35S:AGL15 or 35S:SEP3 transgene were removed, wounded, and placed on GM plates, containing Murashige-Skoog (MS) salts, vitamins, 1% (w/v) sucrose, 0.05% MES (w/v), and 0.7% agar (w/v), pH 5.6 5.7, containing 50 µg /ml Kanamycin to ensure selection of transgene containing embryos from heterozygous plants. Secondary embryonic tissue, which develops on the cultured zygotic embryos, was sub-cultured at regular intervals of approximately 3 weeks on GM media plates.

5.4.6 Nuclear prep and Western analysis

Nuclei were isolated from fixed tissue as described by Bowler *et al.,* 2004. Approximately 5g of tissue was ground to powder in liquid nitrogen and suspended in 25ml of EB1 (0.4 M Sucrose 10 mM Tris–HCl, pH 8.0 5 mM b-ME 0.1 mM PMSF). The resulting tissue slurry was filtered through Miracloth and centrifuged at 5000 x g for 20 min at 4°C in a (the brand of centrifuge). A fraction of the supernatant (700 µl) was precipitated with acetone, suspended in 100 µl of sample loading buffer (2% SDS, 0.001% bromophenol blue, 0.2M DTT, 20% glycerol, 20mM sodium phosphate, pH7.2), and boiled for 5 minutes. This sample is referred to a "nuclear depleted" or "soluble" and saved for Western analyses. The pellet was re-suspended in 1ml of EB2 (0.25 M Sucrose 10 mM Tris–HCl, pH 8.0 10 mM MgCl2 1% Triton X-100 5 mM b-ME 0.1 mM PMSF), transferred to a 1.5ml Ependorf tube and centrifuged at 12000 x g for 10 min at

4°C. The pellet was re-suspended in 400 ul of EB3 (1.7 M Sucrose 10 mM Tris–HCl, pH 8.0 0.15% Triton X-100 2 mM MgCl2 5 mM BME 0.1 mM PMSF) and overlayed on top of another 400 ul of EB3 and centrifuged at 16000 x g for 30 min at 4°C. The pellet was re-suspended in 200 ul sample loading buffer.

Proteins were separated on 12.5 % (w/v) polyacrylamide denaturing gels using vertical gel apparatus (Hoefer Scientific Instruments, San Francisco, CA) and then blotted onto Immobilon™ PVDF Transfer Membranes (Millipore, Billerica, MA) using a Genie blotter (Idea Scientific Co., Minneapolis, MN).

Blots were blocked with 5% non-fat dry milk, washed in TBST (0.1% Tween 20, 100mM Tris-Cl, 150mM NaCl, pH 7) and probed with 1:1,000 diluted AGL15 antiserum (Heck *et al.*, 1995; Perry *et al.*, 1996, 1999, Wang *et al.,* 2000), followed by 1:5,000 diluted secondary antibody (HRP-conjugated goat-anti-rabbit). Finally the membrane was incubated for 1minute in substrate (0.5M Tris, 0.03% H_2O_2, 0.1mg/ml Luminol, 0.4 mg/ml P-idophenol) before being exposed (0.5-5 minutes) to Kodak XAR5 X-ray films (Eastman Kodak, Rochester, NY).

5.5 Summary

Interactions between MADS-domain proteins, with particular emphasis on those expressed in embryonic tissue are reported here. Interactions between AGL15 and SEPALLATA 3 (SEP3) and between SEP3 and LEAFY COTYLEDON 1 (LEC1) are of particular interest because they hint that AGL15 and LEC1, both promoters of somatic embryogenesis, may exist in the same complex together, via a shared association with SEP3. The C-terminal domain of SEP3, but not AGL15, is able activate reporter genes in yeast, suggesting that recruitment of SEP3 by a promoter bound AGL15 might account for the activation of a subset of AGL15 downstream target genes. Plants carrying the *35S:SEP3* transgene and those carry a transgene coding for an activated form of AGL15, *35S:AGL15-VP16*, exhibit similar phenotypes, and this could be due to activation of a shared set of genes. Ectopic expression of *SEP3* permits maintenance of development in the embryonic mode, most likely via its activation of *AGL18*.

Preface

The work reported herein formed a manuscript intended for publication as a short paper:

Authors: Weining Tang, Kristine Hill and Sharyn E. Perry, 2003

The author this dissertation was involved in the initial cloning of *GmAGL15* in collaboration with Dr. Weining Tang, who performed the expression analysis. The original manuscript has been elaborated on and discusses some additional studies reported in Weining Tang's dissertation, University of Kentucky, 2004.

6.1 Introduction

MADS-domain proteins are a family of transcriptional regulatory factors found in eukaryotic organisms. In plants, MADS domain proteins are the central players in many developmental processes, including control of flowering-time, homeotic regulation of floral organogenesis, fruit development, and seed pigmentation (Parenicová *et al*., 2003, and the references therein). There are over 100 MADS box genes in *Arabidopsis thaliana*. Based on the primary sequence, the MADS family can be divided into five subfamilies, namely MIKC, Mα, Mβ, Mγ and Mς, with MIKC being the best studied group. Most MADS proteins with known functions fall into this subfamily (Parenicová *et al*., 2003). Interestingly, among the 39 *A. thaliana* MIKC-type MADS box genes, only one gene, *AGL15* (for *Agamous-Like 15*) has been reported as preferentially expressed during *Arabidopsis* embryogenesis (Heck *et al*., 1995; Rounsley *et al*., 1995). AGL15 and putative orthologs accumulate in the nuclei of angiosperm tissues developing in an embryonic mode, regardless of the origin of the embryos, suggesting a strong correlation between AGL15 and embryogenesis (Perry *et al*., 1996, 1999). Furthermore, persistent accumulation of AGL15 in tissues constitutively expressing the gene promotes somatic embryogenesis (Harding *et al*., 2003).

Although immuno-reactive proteins could be detected using AGL15-specific antibodies in a variety of embryos or embryonic tissues from dicots as well as monocots (Perry *et al*., 1996, 1999, and unpublished observations), little is known about the molecular nature of *AGL15* orthologs in higher plants other than *Arabidopsis thaliana* and *Brassica napus*. In an effort to isolate *AGL15* orthologs from other plant species, the authors cloned a soybean (*Glycine max*) MADS box gene. This gene, designated *GmAGL15*, had highest sequence similarity to the previously published *AGL15*'s. Although numerous MADS box genes have been reported from dozens of plant species, including soybean, to date no full-length cDNA or protein has been identified from soybean (Genbank, http://www.ncbi.nlm.nih.gov; SMART, http://smart.embl-heidelberg.de), making the sequences reported here the first case for this agriculturally important crop.

6.2 Results

6.2.1 Isolation of GmAGL15

Isolation of sequences encoding GmAGL15 was initiated by searching the Genbank EST database for possible candidates. One entry (accession number: AW756465) was found annotated as “similar to… AGL15”. BLASTX program (http://www.ncbi.nlm.nih.gov/blast) confirmed this EST represented the N-terminal 37 amino acid residues of the conserved MADS domain. To obtain the full-length cDNA, RNA was isolated from a soybean somatic embryo culture (Reddy *et al*., 2001). Oligonucleotide primers were designed based on the EST sequence, and 3’-RACE PCR was performed as described (Ausubel *et al*., 1998). After sequencing this partial cDNA, additional primers were designed for 5’-RACE PCR to recover the full-length cDNA. All primer sequences are available upon request.

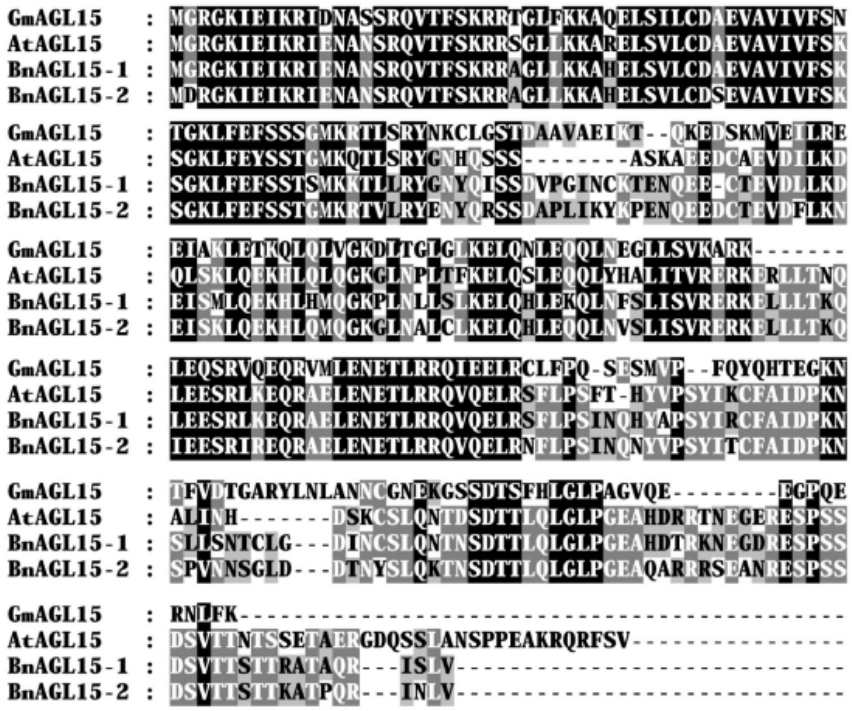
The longest cDNA clone (Genbank accession: AY370659) consisted of a ~270 bp 5’-UTR, ~260 bp 3’-UTR plus polyA tail, and an ORF of 708 bp, which encodes a protein of 235 aa. BLASTP search was performed using the protein sequence. The highest scoring matches were to *Arabidopsis* AGL15 (AtAGL15), *Brassica* type I and type II AGL15 (BnAGL15-1 and -2; accession numbers: Q38847, T07867 and T07869, respectively). BLASTN using the DNA coding sequence yielded the same result. Pairwise comparison of GmAGL15 protein with the other three AGL15s revealed an identity of approximately 50%. Multiple sequence alignment showed a moderate conservation among these proteins (Figure 5.1). This was not unexpected, because even between *A. thaliana* and *B. napus*, two closely related species, considerable divergence in AGL15 exists. Nevertheless, the soybean sequence displayed overall homology to the AGL15 proteins, including divergent domains outside the conserved MADS domain. In addition, the soybean protein contained several “signature” sequences that are rarely found in MADS domain proteins other than AGL15, such as the C-terminal LENETLRRQ and LGLP motifs. Therefore the authors propose that GmAGL15 is most likely the soybean ortholog of AGL15.

The 5.8 kb genomic region of *GmAGL15* (accession: AY370660) was amplified from *G. max* cv. Jack genomic DNA, using primers corresponding to the UTRs of the cDNA. An

alignment of the genomic and cDNA sequences revealed that *GmAGL15* contained eight exons and seven introns (Figure 5.2). The introns were longer than found in *Arabidopsis*. Nevertheless, the exon-intron boundary locations appeared to be identical between the two species, as often observed among evolutionarily conserved orthologs. This further suggested *GmAGL15* was the soybean counterpart of *AGL15*.

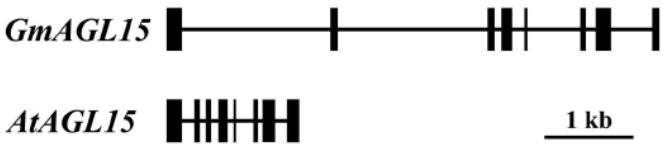
A phylogenetic tree was constructed using protein sequences of GmAGL15, BnAGL15s, and all 39 *Arabidopsis* MIKC-type MADS domain proteins (Parenicová *et al*., 2003). Indeed, GmAGL15 was grouped more closely to the AGL15s (Figure 5.3).

Figure 6.1 Sequence alignment between GmAGL15, AtAGL15, BnAGL15-1 and BnAGL15-2



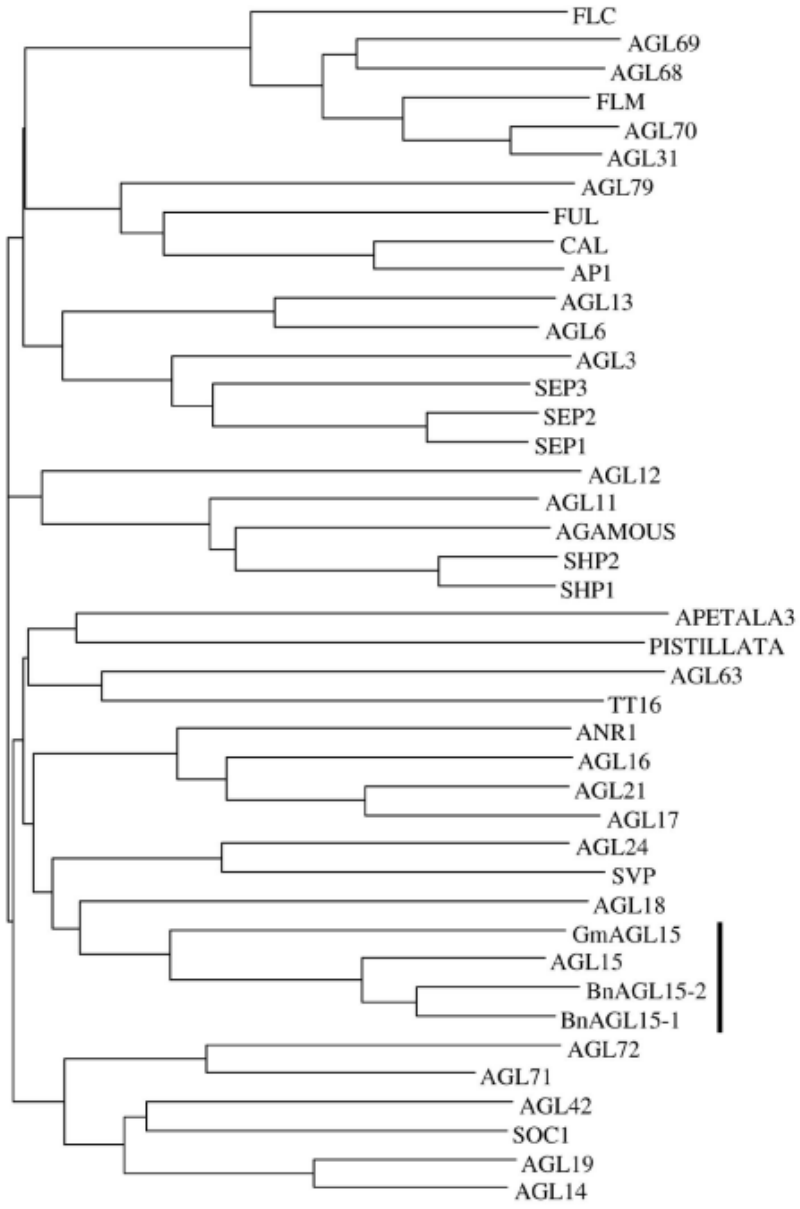
GeneDoc (www.pac.edu/biomed/genedoc) sequence alignment between GmAGL15, AtAGL15, BnAGL15-1 and BnAGL15-2. The shade levels represent conservation degrees.

Figure 6.2 Schematic representation of the gene structures of *GmAGL15* and *AtAGL15*



The numbers and positions of the introns/exons were conserved between these two species. Boxes, exons; lines, introns.

Figure 6.3 A phylogenetic tree generated from GmAGL15, BnAGL15s and all 39 MIKC-type *Arabidopsis* MADS domain proteins

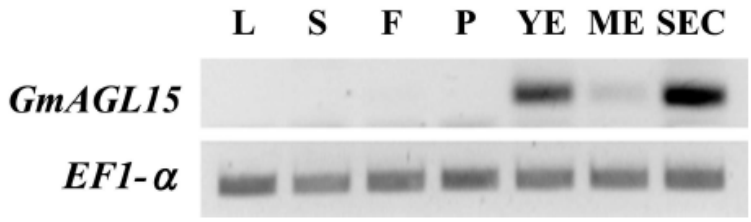


Neighbor joining method with a bootstrap number of 1000 was used (ClustalX 1.81, http://www-igbmc.u-strasbg.fr/BioInfo/ClustalX). AGL15 group is indicated.

6.2.2 Expression pattern of *GmAGL15*

To investigate the expression pattern of *GmAGL15*, semi-quantitative RT-PCR was performed on RNAs isolated from various tissues. As shown in Figure 6.4, *GmAGL15* transcript was not detected in the vegetative tissues (leaves, stems), or in the open flowers. *GmAGL15* mRNA was not detectable at very early stages of seed pod development, but was more abundant in the young developing embryos, and the level declined after maturation. This pattern was consistent with that previously reported for *AGL15*s (Heck *et al*., 1995; Rounsley *et al*., 1995). Notably, in the somatic embryo culture, the highest level of *GmAGL15* mRNA accumulation was detected. AGL15-specific antiserum (Perry *et al*., 1999) detected immuno-reactive protein in nuclear extracts prepared from soybean somatic embryos (data not shown). Additionally, reaction of the antiserum against the product of the cloned *GmAGL15* was tested by expressing the soybean gene in *Escherichia coli*. The AGL15-specific antiserum recognized the *E. coli* produced protein (data not shown).

Figure 6.4 Expression pattern of *GmAGL15*



Semi-quantitative RT-PCR was performed on RNA derived from various tissues of *G. max* cv. Jack. The coding region of *Elongation Factor 1-α* (*EF1-α*) was amplified for normalization. L, young leaves; S, stems; F, open flowers; P, seed pods containing very young embryos; YE, young embryos (average length 2 mm); ME, mature green embryos; SEC, somatic embryo culture.

6.3 Discussion

In conclusion, *GmAGL15*, an embryo MADS box gene was isolated from soybean. Sequence similarity, gene structure and expression pattern support *GmAGL15* as the likely soybean ortholog to previously reported AGL15's from *A. thaliana* and *B. napus*. This is the first time that a full-length MADS box gene has been reported from soybean. MADS domain proteins have fascinating developmental roles, and AGL15 may be involved in specification of embryo identity. Not only is there a correlation between presence of AGL15 and development in embryonic mode (Perry *et al*., 1999), constitutive expression of *Arabidopsis AGL15* is sufficient to promote somatic embryo development in some developmental contexts (Harding *et al*., 2003). Therefore, it will be especially interesting to ascertain if there is a link between expression of *GmAGL15* and promotion/maintenance of somatic embryos in soybean.

Many cultivars of soybean are recalcitrant to somatic embryogenesis. Preliminary data suggests that soybean transformation experiments using *35S:GmAGL15* may improve the recovery of viable transformed embryos (W. Tang *et al.,* unpublished data). Therefore, *GmAGL15* may be a valuable biotech tool as a co-transformation vector. Alternatively somatic embryo cultures expressing *GmAGL15* could be used as target tissue for transformation. However, *AtAGL15* appears not to enhance somatic embryogenesis in soybean. Likewise, preliminary experiments suggest that in *Arabidopsis* seedlings, the *35S:GmAGL15* transgene does not enhance somatic embryogenesis. However, *AtAGL15* overexpressors seedlings examined at the same time did show typical enhancement of somatic embryo production (W. Tang, University of Kentucky Graduate School Dissertation, 2004). This is intriguing. While the *Arabidopsis* and soybean *AGL15* genes appear to enhance somatic embryogenesis in their original species *AtAGL15* has no noticeable effect on soybean or *GmAGL15* on *Arabidopsis* somatic embryogenesis. One reason could be is that the differences between AtAGL15 and GmAGL15 render them unable to form complexes with interaction partners. For example, AGL15 interacts with SAP18 and the highly conserved LxLxL motif seems to be important (Figure 2.4). However, in the soybean sequence the first leucine of the LxLxL motif is replaced by a phenylalanine residue (Chapter 2, Figure 2.2). Therefore GmAGL15 might not be able to interact with AtSAP18 (or another interaction partner) or AtAGL15 with GmSAP18.

6.4 Materials and Methods

6.4.1 RACE-PCR

RACE-PCR (Rapid Amplification of cDNA Ends-PCR) was performed as described (Ausubel *et al*., 1998) with minor modifications. For 3'-RACE, 2 µg of total RNA was reverse transcribed into cDNA using an oligo dT-adaptor primer (CCGGATCCTCTAGAGCGGCCGCTTTTTTTTTTTTTTTT) and M-MLV Reverse Transcriptase (Promega, Madison, WI). The 3' part of the cDNA was amplified by PCR using the adaptor primer (CCGGATCCTCTAGAGCGGCCGCT) and two nested primers (AATGCCAACAGCAGACAAGT and GTGGAGATTCTAAGAGAGG). For 5'-RACE, 9 µg of total RNA was annealed with the internal primers (CCTCTCTTAGAATCTCCAC or TCGCAAAGTTTCATTCTCCAAC) and reverse transcription performed. The cDNA was tailed with multiple dA using Terminal Transferase (New England Biolabs, Beverly, MA), according to the manufacturers' protocol. Second-strand cDNA synthesis and PCR were performed as for 3'-RACE, except with different nested specific primers (ACTTGTCTGCTGTTGGCATT or TCGCAAAGTTTCATTCTCCAAC). The final PCR products representing the 5' and 3' portion of the cDNA were cloned into pGEM T vectors (Promega) for sequencing.

6.4.2 Sequence Analysis

Arabidopsis thaliana AGL15 and Brassica napus type I and II AGL15 sequences were obtained from the Genbank (accession numbers: Q38847, T07867 and T07869, respectively). GmAGL15 protein sequence was deduced and compared with the other AGL15 sequences using the GeneTool Lite and PepTool Lite software package (BioTools Inc., Edmonton, Alberta, Canada). Sequences were aligned using the GeneDoc program (www.pac.edu/biomed/genedoc). Sequences of the 39 Arabidopsis MIKC-type MADS domain proteins were extracted from the Genbank. The phylogenetic tree was generated using the ClustalX 1.81program (http://www-igbmc.u-strasbg.fr/BioInfo/ClustalX). Neighbor Joining method with a bootstrap number of 1,000 was used.

6.4.3 Semi-Quantitative RT-PCR

Total RNA was isolated from various soybean tissues and cDNA synthesized as described above. Semi-quantitative PCR was performed for GmAGL15, and for ELONGATION FACTOR 1-α (EF-1α) as a normalization control. The primers were as follows:

GmAGL15
Forward 5'GTGGAGATTCTAAGAGAGG'3
Reverse 5'TCGCAAAGTTTCATTCTCCAAC'3
EF-1α
Forward 5'ACGCTCTACTTGCTTTCACC'3
Reverse 5'GCACCGTTCCAATACC'3

The PCR program used included an initial denaturation at 95°C for 3 min, followed by 27-35 cycles of 94°C for 30 sec, 57°C for 30 sec and 72°C for 30 sec, and a final extension of 10 min at 72°C. The PCR products were visualized on 1.5% agarose gels and images acquired using a ChemImager (Alpha Innotech Corp., San Leandro, CA).

6.5 Summary

A full-length cDNA encoding a MADS domain protein was isolated from soybean somatic embryos. Subsequently the corresponding genomic region of the gene was obtained. This gene, designated *GmAGL15*, encodes a protein with highest similarity to AGL15 from *Arabidopsis thaliana* and *Brassica napus*. *GmAGL15* was preferentially expressed in developing embryos, the same pattern as observed for *Arabidopsis* and *Brassica AGL15's.* The expression pattern, combined with the intriguing roles MADS domain proteins in general, and AGL15 in particular, play in plant development suggests that GmAGL15 may be a central regulatory factor during soybean embryo development.

Preface

This project set out to identify proteins able to interact with AGL15 and to characterize one or two of these interactions. This objective has been fulfilled (Chapters 2 thru 5). However, in science, questions are often answered with more questions. The purpose of this chapter is to reassert the main findings presented in this dissertation and to discuss future directions ongoing research might take.

7.1 Main findings and suggestions for further directions

7.1.1 Co-regulation of AGL15 target genes

Yeast two-hybrid screens have indentified novel AGL15-interacting proteins (Table 4.2). The interactions likely to be of biological significance include transcription factors, co-regulators, putative chromatin remodeling factors and RNA-binding proteins. Investigation into the role these interactions might be playing in regulating gene expression has been initiated and findings are reported in Chapters 2 and 3. Research programs, lead by Dr. Sharyn Perry, have identified *in vivo* AGL15-binding sites, including those reported here. A reasonable hypothesis is that a subset of AGL15 bound target genes will also be bound by AGL15-interacting transcription factors or co-regulators. A current graduate student, Yumei Zheng, is using Affymetrix whole genome arrays to map *in vivo* AGL15 binding sites. Although beyond the scope of this study, further projects might involve mapping *in vivo* binding sites of AGL15-interacting partners to identify co-regulated genes.

7.1.2 AGL15 functions as a repressor protein by recruitment of a histone deacetylase complex

In Chapter 2 an interaction between AGL15 and members of the SWI-INDEPENDENT 3/HISTONE DEACETYLASE (SIN3/HDAC) complex is reported. Three previously unreported direct targets of *AGL15*: *LEA76*, *CBF2*, and *AGL18*, whose transcripts accumulation is decreased in response to AGL15, are described. Two of these AGL15-target genes are also responsive to a member of the SIN3/HDAC1 complex, SIN3 ASSOCIATED POLYPEPTIDE OF 18 KD (SAP18). AGL15 exhibits *in vivo* transcriptional repressor activity, and within a region necessary for the repressive function of AGL15 resides a conserved motif, similar to the previously reported LxLxL/EAR repression domains (Tiwari *et al.,* 2001, Ohta *et al.,* 2001). What is more, the aforementioned motif mediates the association of AGL15 with SAP18 in yeast two-hybrid assays, thus providing a possible mechanism for AGL15's role in repressing gene expression via recruitment of a histone deacetylase complex.

It is has been demonstrated that ectopic expression of either *AGL15* or *SAP18* significantly decreases transcript levels of direct AGL15 target genes, *LEA76* and *CBF2* (Figure 2.5). The proposed model (Figure 2.9) is that AGL15 acts as a platform, recruiting the SIN3/HDAC1 complex to a subset of target genes. Therefore it can be hypothesized AGL15 is a requirement for the SIN3/HDAC1 mediated repression of a subset of AGL15 regulated genes. While reduced *LEA76* transcript levels are observed when *SAP18* is ectopically expressed in a wild-type background, preliminary data suggests that the same *35S:SAP18* lines no longer exhibit this repression, when crossed into the *agl15-2* background. Plants have been recovered that are homozygous for *35S:SAP18* and *agl15-2* alleles, and are being compared to the homozygous parental line. At the time of writing these experiments are holding up in biological repeats, providing additional evidence of an *in vivo* interaction between SAP18 and AGL15.

Technical issues have prevented sufficient immunoprecipitation using anti-c-myc. An antibody raised against *Drosophila* SAP18 recognizes *Arabidopsis* SAP18 *in vitro*, but the high level cross-reactivity *in planta* may hinder effective immunoprecipitation *in vivo*. However, technical difficulties aside, it would be worthwhile to pursue chromatin immunoprecipitation and ultimately show that target genes such as *LEA76* and *CBF2* are bound *in vivo* by members of the SIN3/HDAC1 complex. One way to approach this might be to utilize commercially available antibodies against histone acetylation states. Antibodies which recognize acetylated and un-acetylated histones, H3 and H4, are available, and have been shown to work in *Arabidopsis* (Benhamed *et al.,* 2006). Benhamed *et al.,* 2006, performed chromatin immunoprecipitation experiments to compare the acetylation states of light regulated genes in wildtype and mutant seedling carrying null alleles of a histone acetyltransferase (*GCN5*) or the histone deacetylase, *HDA19* gene. A similar approach could be adapted to this study by comparing the acetylation states of AGL15 bound genes in wild type, *agl15* (or *agl15/agl18* double mutant), and *35S:AGL15* tissues. The hypothesis would be that increased levels of AGL15 cause a decrease in the acetylation state of a subset of target genes, such as *LEA76* or *CBF2*, due to recruitment of the SIN3/HDAC1 complex.

7.1.3 Co-regulation by AGL15 and GRP2, and cold tolerance

An interaction between AGL15 and the COLD SHOCK DOMAIN (CSD) –containing protein, GRP2 is reported in Chapter 3. Preliminary data showing enhanced tolerance to freezing stress conferred by seedlings carrying the *35S:AGL15* transgene is also presented. However, the freezing tolerance phenotype of *35S:AGL15* seedlings needs to be verified under using a series of controlled temperature conditions. AGL15 directly binds and regulates the expression of other CSD-containing proteins, which may act to enhance translation under cold conditions. AGL15 binds *GRP2b* and appears to repress its transcription. *GRP2b* transcript levels also decrease in response to cold, whereas *CSD4* increases (Karlson and Imai, 2003). *AGL15* binds *CSD4* in vivo, but the effect of AGL15 levels on *CSD4* remains to be tested. Heterologous expression of *CSD4* is able to complement the cold sensitivity of mutant *Escherichia coli* that lack four cold shock proteins (Kim *et al.,* 2007b). If AGL15 increased levels of CSD4 protein it might explain the apparent freezing tolerance of *35S:AGL15* seedlings.

Unlike other CSD-containing proteins, *35S:GRP2* does not appear to enhance the freezing tolerance of seedlings (K. Hill, unpublished data, Kim *et al.,* 2007a). *GRP2b* transcript levels are increased in *35S:GRP2* seedling, suggesting that *GRP2b* does not enhance cold tolerance either. This is in keeping with *GRP2b* being decreased in response to cold (Karlson and Imai, 2003), but contrary to what one would expect were AGL15 and GRP2 co-regulating gene expression. A possible explanation is that GRP2 and AGL15 have antagonizing effect at certain promoters. If GRP2 functioned solely as a transcriptional activator, increased levels might siphon AGL15 away from repression complexes. It would be worth testing if genes whose transcript levels are normally increased by AGL15 are also increased in response to GRP2.

Demonstration that GRP2 bound some of the *in vivo* sites identified as AGL15 –binding sites, would be evidence of AGL15 and GRP2 co-regulation. Preliminary experiments hinted that *GRP2b* is bound in *vivo* by c-myc:GRP2. However, technical issues prevented sufficient immunoprecipitation using anti-myc. Plans are underway to repeat the experiment using the newly available antibody raised to GRP2 (generously provide by Dr. Gilberto Sachetto—Martins, University of Rio de Janeiro, Brazil). Cis-element prediction programs such as Promomer (Toufighi *et al.,* 2005) can be employed to find

putative Y-box sequences adjacent to AGL15 bound regions, and identify putative candidates to test for binding GRP2.

7.1.4 Co-regulation by AGL15 and SEP3

Interactions between AGL15 and SEPALLATA 3 (SEP3) and between SEP3 and LEAFY COTYLEDON 1 (LEC1) are of particular interest because they hint that AGL15 and LEC1, both promoters of somatic embryogenesis, may exist in the same complex together, via a shared association with SEP3. What is more, ectopic expression of SEP3 permits maintenance of development in the embryonic mode, most likely via its activation of *AGL18*. The C-terminal domain of SEP3, but not AGL15, is able activate reporter genes in yeast, suggesting that recruitment of SEP3 by a promoter bound AGL15 might account for the activation of a subset of AGL15 downstream target genes. C-myc tagged versions of SEP3 theoretically enable chromatin immunoprecipitation experiments to be performed. This approach has been fraught with difficulties immunoprecipitating sufficient protein with c-myc antibodies. However, very tentative data suggests that some AGL15 target genes might also be bound by SEP3.

7.1.5 Other interesting AGL15 interacting proteins

Yeast two-hybrid screens were undertaken to identify novel proteins able to interact with AGAMOUS-Like 15 (AGL15). A number of interesting and potentially biologically important AGL15-interacting partners were discovered (Table 4.2). These include a member of the SIN3 histone deacetylase complex, SAP18, and a CSD-containing protein, GRP2, which are described Chapters 2 and 3 respectively. Other interesting AGL15-interacting proteins include a K-homology domain/CCCH type zinc finger protein, a bZIP transcription factor (bZIP1), a histone acetyl transferase (HAG5), a homeobox-leucine zipper protein (HDG8), a LOB-domain containing protein (LOB40), and an Agenet domain containing protein. Regions of AGL15 mediating these interactions were mapped to one of three regions: the K-domain, the C-domain, or the K-and C-domains together. Some, but not all of the AGL15-interacting proteins were also able to associate with other MADS-domain proteins in yeast two-hybrid assays.

On candidate worthy of further investigation is LOB40. *In vivo* LEC2, a promoter of somatic embryogenesis, induces the expression of both *AGL15* and *LOB40* and *in vitro* LEC2 can bind the RY motifs (CATGCA) present in both *AGL15* and *LOB40* promoter regions (Braybrook *et al*., 2006). Braybrook *et al*., 2006 make the statement "*LEC2 may also induce somatic embryo development by increasing tissue competency to undergo somatic embryogenesis through AGL15*" and it will be interesting to determine if, like AGL15 (Harding *et al.,* 2003), over-expression of *LOB40* also promotes somatic embryogenesis, or if over-expression of both *LOB40* and *AGL15* enhance this. T1 plants carrying the *35S: LOB40* transgene are currently growing and await further analysis.

Another intriguing candidate is the K-homology/zinc finger domain protein. Other K-homology proteins in plants have been shown to bind RNA in the first intron of *FLC* (Mockler *et al*., 2004) and to play a role in splicing of the first intron of *AG* (Cheng *et al*., 2003). Both *AG* and *FLC* contain putative CArG elements in their first intron. AGL15 binds to a CArG motif in its own promoter (Zhu and Perry, 2005) and higher levels of AGL15 accumulation are observed in transgenic plants carrying a genomic version of *AGL15* compared to a form lacking introns (Fernandez *et al.,* 2000). Therefore, it would be interesting to test if there was a connection between transcription and RNA-processing, mediated by the interaction of AGL15 with the KH-domain protein. T1 plants, transformed with a c-myc tagged version of the K-homology protein are currently growing and await further analysis.

Given the suggested antagonist role of the histone acetyl transferase, GCN5, which is similar to HAG5, a putative AGL15 interacting protein, and HDA19 (Benhamed *et al.,* 2006), further study regarding the association between AGL15 and HAG5 may be warranted. Other candidates worthy of further investigation are the Agenet domain-containing protein and HDG8. The interaction between AGL15 and the Agenet domain-containing protein is intriguing because of the inferred role of the Agenet domain in chromatin remodeling. HDG8 is a homeodomain leucine-zipper protein that contains a START domain which is predicted to function in lipid signaling. Therefore this protein offers a potential link between transcription and signaling.

7.1.6 Soybean AGL15

A full-length cDNA encoding a MADS domain protein was isolated from soybean somatic embryos. This gene, designated *GmAGL15*, encodes a protein with highest similarity to AGL15 from *Arabidopsis thaliana* and *Brassica napus*. *GmAGL15* was preferentially expressed in developing embryos, the same pattern as observed for *Arabidopsis* and *Brassica AGL15.*

Many cultivars of soybean are recalcitrant to somatic embryogenesis. Therefore, *GmAGL15* may be a valuable biotech tool as a co-transformation vector. Alternatively somatic embryo cultures expressing *GmAGL15* could be used as target tissue for transformation. *Arabidopsis* and soybean *AGL15* genes appear to enhance somatic embryogenesis in their original species, but *AtAGL15* has no noticeable effect on Soybean or *GmAGL15* on Arabidopsis somatic embryogenesis (W. Tang, and S. Perry, unpublished data). One reason could be is that the differences between AtAGL15 and GmAGL15 render them unable to form complexes with certain interaction partners. For example, AGL15 interacts with SAP18, and the conserved LxLxL motif seems to be important to this interaction (Figure 2.4). In *GmAGL15* the first leucine residue of the conserved LxLxL motif is replaced by a phenylalanine (Figure 2.2). Therefore GmAGL15 might not be able to interact with AtSAP18, or AtAGL15 with GmSAP18. This hypothesis could easily be tested via yeast two-hybrid assays. Of course SAP18 is one example, but many yeast two-hybrid constructs, containing AtAGL15-interaction partners, are already available and testing their ability to interact with GmAGL15 would be relatively non-labor intensive. Likewise generating and testing putative Soybean orthologs (sequence data permitting) of AGL15- interacting partners would be a relatively simple and inexpensive experiment.

7.2 Concluding remarks

This project set out to identify novel AGL15-interacting proteins and characterize interactions of biological relevance. This objective has been fulfilled. A number of interactions likely to be of biological significance have been discovered, namely those involving transcription factors, co-regulators, putative chromatin remodeling factors and RNA-binding proteins. Several of these interactions have been investigated further and are presented in this dissertation. Of particular note is the interaction between AGL15 and members of a histone deacetylase complex, because it suggests a possible mechanism to explain the observed transcriptional repressive capacity of AGL15.

Appendix A: GAL4-DBD- fusion “bait” constructs used in this study

DBD-	AAs coded for by cDNA	RE	Oligonucleotides/Intermediate Restriction Site
AGL15 (MIKC)	At5g13790 AA 1 - 268	Nco1 BamH1	Cloned from intermediate in pET vector
AGL15 (IKC)	At5g13790 AA 64 - 268	Nco1 BamH1	Cloned from intermediate in pET vector
AGL15 (IK)	At5g13790 AA 62 – 152	Nde1 EcoR1	GGAATTCCATATGAAGCAAACACTTTCCAG GGAATTCTCGTTGTTCCTTGAGGCG
AGL15 (MI½)	At5g13790 AA 1-105	Nco1 Pst1	Cloned from DBD-AGL15 (MIKC) Utilized internal Pst1
AGL15 (C)	At5g13790 AA 152-268	Nde1 BamH1	GGAATTCCATATGGCAGAGTTGGAAAACGAG GCGGATCCCTAAACAGAGAACCTTTGTC
AGL15 (KC)	At5g13790 AA 84-268	Nde1 EcoR1	GGAATTCCATATGTGTGCAGAGGTGGATATTTTAAAG GCGGATCCCTAAACAGAGAACCTTTGTC
AGL15 (IKC***)	At5g13790 AA 64 – 268	Nco1 BamH1	Mutated LxLxL to AxAxA
AGL15 (KC***)	At5g13790 AA 84-268	Nde1 EcoR1	Mutated LxLxL to AxAxA
AGL15 (½KC)	At5g13790 AA 118 – 268	EcoR1 BamH1	Cloned from AD-rec-AGL15 (½KC)
AGL15 (½MIKC)	At5g13790 AA 31 – 268	Nde1 EcoR1	GGAATTCGCTCGTGAGCTCTCTG GCGGATCCCTAAACAGAGAACCTTTGTC
SAP18	At2g45640 AA 1- 152	Nde1 BamH1	GGAATTCCATATGACTGAAGCAGCGAGAAG CGCGGATCCCTAGTAAATTGCCACATC
HDA6	At5g63110 AA 1-287	Nde1 EcoR1	GGAATTCCATATGGAGGCAGACGAAAGCGG CCGGAATTCGTTGAAGCAACCCAACCG
HDA19	At4g38130 AA 1-356	EcoR1 BamH1	GGAATTCGATACTGGCGGCAATTCGCTG CGCGGATCCACTTGGAGCAACGTGAAGTG
SVP (MIKC)	At2g22540 AA 1- 239	Nde1 BamH1	GGAATTCCATATGGCGAGAGAAAAGATTC CGCGGATCCACCACCATACGGTAAG
SVP (IKC)	At2g22540 AA 61- 239	Nde1 BamH1	GGAATTCCATATGAAGGAAGTCCTAGAGAG CGCGGATCCACCACCATACGGTAAG
SOC1 (MIKC)	AT2G45660 AA 1-214	Nde1 BamH1	GGAATTCCATATGGTGAGGGGCAAAACTCAG CGCGGATCCTCACTTTCTTGAAGAACAAG
SOC1 (IKC)	AT2G45660 AA 55-214	EcoR1 BamH1	Cloned from BD-SOC1 (MIKC) Utilized internal EcoR1 site
AGL18 (MIKC)	At3g57390 AA 1- 256	Nde1 BamH1	GGAATTCCATATGAGAGGAAGGATTGAG CGCGGATCCATCAGAAGCCACTTG
AGL18 (IKC)	At3g57390 AA 62- 256	EcoR1 BamH1	GGAATTCATGGAGCAAATTCTTTCTAG CGCGGATCCATCAGAAGCCACTTG
SHP1 (MIKC)	AT3G58780 AA 1-248	Nde1 BamH1	GGAATTCCATATGGAGGAAGGTGGGAGTAGTC CGCGGATCCTTACACAAGTTGAAGAGG
PI (MIKC)	AT5G20240 AA 1-208	Nde1 EcoR1	GGAATTCCATATGGGTAGAGGAAAGATCG CCGGAATTCTCAATCGATGACCAAAG
FLC (MIKC)	At5g10140 AA 1- 196	Nde1 EcoR1	GGAATTCCATATGCGCAACGGTCTCATCGAG CCGGAATTCCTAATTAAGTAGTGGGAGAG
SEP3 (MIKC)	AT1G24260 AA 1- 250	Nde1 EcoR1	GGAATTCCATATGGGAAGAGGGAGAGTAG CGCGGATCCAATAGAGTTGGTGTC
SEP3 (½KC½)	AT1G24260 AA 119-196	EcoR1 BamH1	GGAATTCGGGTATCAGATGCCAC CGCGGATCCACGACCGTAGTGATCAACCTC
SEP3 (MIK)	AT1G24260 AA 1- 176	Nde1 BamH1	GGAATTCCATATGGGAAGAGGGAGAGTAG CGCGGATCCATCAGCTAACCTTAGTC
SEP3 (MIKC½)	AT1G24260 AA 1- 196	Nde1 BamH1	GGAATTCCATATGGGAAGAGGGAGAGTAG CGCGGATCCACGACCGTAGTGATCAACCTC
SEP2 (MIKC)	At3g02310 AA 1 - 250	Nde1 BamH1	GGAATTCCATATGGGAAGAGGAAGAGTAGAGCTC CGGGATCCTCACAGCATCCAGCCAGG
LEC1	At1g21970 AA 1-208	Nde1 BamH1	GGAATTCCATATGACCAGCTCAGTCATA CGCGGATCCCTTATACTGACCATAATG

Appendix B: GAL4-AD- fusion “prey” constructs used in this study

AD-	AAs coded for by cDNA	RE	Oligonucleotides/Intermediate Restriction Site
AGL15 (MIKC)	At5g13790 AA 1 - 268	Nco1 BamH1	Cloned from intermediate in pET vector
AGL15 (IKC)	At5g13790 AA 64 - 268	Nco1 BamH1	Cloned from intermediate in pET vector
AGL15 (C)	At5g13790 AA 152 - 268	Nde1 BamH1	GGAATTCCATATGGCAGAGTTGGAAAACGAG GCGGATCCCTAAACAGAGAACCTTTGTC
AGL15 (KC)	At5g13790 AA 84-268	Nde1 EcoR1	GGAATTCCATATGTGTGCAGAGGTGGATATTTTAAAG GCGGATCCCTAAACAGAGAACCTTTGTC
AGL15 (IK)	At5g13790 AA 62 – 152	Nde1 EcoR1	GGAATTCCATATGAAGCAAACACTTTCCAG GGAATTCTCGTTGTTCCTTGAGGCG
AGL15 (½KC)	At5g13790 AA 118 – 268		Recovered clone
AGL15 (½MIKC½)	At5g13790 AA 32-180	EcoR1 BamH1	GGAATTCGCTCGTGAGCT GCGGATCCTCAGGATGGAACATAGTGGGAG
LEC1	At1g21970 AA 1-208	Nde1 BamH1	GGAATTCCATATGACCAGCTCAGTCATA CGCGGATCCCTTATACTGACCATAATG
SEP3 (MIKC)	AT1G24260 AA 1- 250	Nde1 BamH1	GGAATTCCATATGGGAAGAGGGAGAGTAG CGCGGATCCAATAGAGTTGGTGTC
SEP3 (MIK)	AT1G24260 AA 1- 176	Nde1 BamH1	GGAATTCCATATGGGAAGAGGGAGAGTAG CGCGGATCCATCAGCTAACCTTAGTC
SEP3 (½KC)	AT1G24260 AA 119-250	EcoR1 BamH1	GGAATTCGGGTATCAGATGCCAC CGCGGATCCAATAGAGTTGGTGTC
SEP3 (C)	AT1G24260 AA 177-250	EcoR1 BamH1	GGAATTCGGGTATCAGATGCCAC CGCGGATCCAATAGAGTTGGTGTC
SEP3 (MIKC½)	AT1G24260 AA 1- 196	Nde1 BamH1	GGAATTCCATATGGGAAGAGGGAGAGTAG CGCGGATCCACGACCGTAGTGATCAACCTC
SEP1 (KC)	AT5G15800 AA 80-250	BamH1 Xho1	CGGGATCCAAGTCAACAACAAACCTGCCAAAG CCGCTCGAGTCAGAGCATCCACCCCGGG
AGL18 (MIKC)	At3g57390 AA 1- 256	Nde1 BamH1	GGAATTCCATATGAGAGGAAGGATTGAG CGCGGATCCATCAGAAGCCACTTG
AGL18 (IKC)	At3g57390 AA 62- 256	EcoR1 BamH1	GGAATTCATGGAGCAAATTCTTTCTAG CGCGGATCCATCAGAAGCCACTTG
SVP (MIKC)	At2g22540) AA 1- 239	Nde1 BamH1	GGAATTCCATATGGCGAGAGAAAAGATTC CGCGGATCCACCACCATACGGTAAG
SVP (IKC)	At2g22540 AA 61- 239	Nde1 BamH1	GGAATTCCATATGAAGGAAGTCCTAGAGAG CGCGGATCCACCACCATACGGTAAG
SHP1 (MIKC)	AT3G58780 AA 1-248	Nde1 BamH1	GGAATTCCATATGGAGGAAGGTGGGAGTAGTC CGCGGATCCTTACACAAGTTGAAGAGG
PI (MIKC)	AT5G20240 AA 1-208	Nde1 EcoR1	GGAATTCCATATGGGTAGAGGAAAGATCG CCGGAATTCTCAATCGATGACCAAAG
SOC1 (MIKC)	AT2G45660 AA 1-214	Nde1 BamH1	GGAATTCCATATGGTGAGGGGCAAAACTCAG CGCGGATCCTCACTTTCTTGAAGAACAAG
SOC1 (IKC)	AT2G45660 AA 55-214	EcoR1 BamH1	Cloned from AD-SOC1 (MIKC) Utilized internal EcoR1 site
FLC (MIKC)	At5g10140 AA 1- 196	Nde1 EcoR1	GGAATTCCATATGCGCAACGGTCTCATCGAG CCGGAATTCCTAATTAAGTAGTGGGAGAG
AD-FLC (IKC)	At5g10140 AA 60 - 196	Nde1 EcoR1	GGAATTCCATATGGGCGATAACCTGGTCAAG CCGGAATTCCTAATTAAGTAGTGGGAGAG
GRP2 (CSD)	At4g38680 AA 1-132	Nde1 BamH1	GGAATTCCATATGAGCGGAGACAACGGC GGATCCGTAGCAGTCGCTGCCTC
GRP2 (-CSD)	At4g38680 AA 63-203	Nde1 BamH1	CATATGATCGACAACAACAACCG CGCGGATCCACGTCCAACGCTGGTGC
HDA6	At5g63110 AA 1-287	Nde1 EcoR1	GGAATTCCATATGGAGGCAGACGAAAGCGG CCGGAATTCGTTGAAGCAACCCAACCG
HDA19	At4g38130 AA 1-356	EcoR1 BamH1	GGAATTCGATACTGGCGGCAATTCGCTG CGCGGATCCACTTGGAGCAACGTGAAGTG

Appendix C: Recovered clones that autoactivate (false positives)

Gene	Description	Class	Screens
At1g51510	RNA binding protein	A	1
At3g02550	LOB41	A	5, 7, 11
At5g44210	ERF domain protein 9 (ERF9)	A	7, 11
At4g14930	Predicted acid phosphatase	B	2
At3g54350	Forkhead-associated domain-containing p	B	7, 11
At4g20270	Clavata1	B	8
AT5g42990	Ubiquitin-conjugating enzyme-like (UBC18)	C	7
At4g05320	UBQ10/SEN3	C	2^{III}
AT5G40480	Nuclear pore protein-like (EMB3012)	D	11, 12
At5g40770	Prohibitin	D	1
At5g40480	Nuclear pore like protein	D	6, 11
At4g31490	b-cop	D	2
At5g16280	GSG1 domain	E	2
AT5g17790	Zinc finger protein VAR3, chloroplast precursor	E	7
At1g04870	Protein arginine N-methyltransferase family	E	7
At1g11930	Putative Proline sythatase ass.	E	1
At1g42970	G3PDH	E	2^{IV}, 3
At1g13440	G3PDH-related	E	2^{IV}, 6
At4g34870	Cyclophillin – Proline isomerase	E	6

Screens, the number designated to the respective independent screen/s said clone was recovered from (see Table 4.1); Roman numerals above the screen number denote the number of unique clones coding for the same protein that were recovered from that screen. Based upon the known or inferred functions of conserved domains (Marchler-Bauer *et al.,* 2007), or predicted sub-cellular localization (Emanuelsson *et al*., 2000), sequenced clones were categorized into the following classes: A, Transcription factors, chromatin remodeling factors, RNA-binding proteins; B, Cytosolic proteins with a possible role in posttranslational modification or signaling; C, Chaperons, protein folding, degradation, proteases; D, Secreted and structural proteins; E, Ribosomal, chloroplastic, mitochondrial proteins, metabolic enzymes; X, unclassified.

Appendix D: Recovered clones that failed activate reporter genes when retested

Gene	Descriptions	Class	Screens
At4g28830	RNA methylase	A	1, 2
At2g30800	RNA helicase	A	2
At1g07820	Histone H4	A	2
At4g35750	Sec14p-like lipid-binding domain	B	8
At5g19140	Auxin/aluminum-responsive protein	E	5
At2g41430	Dehydration-induced protein (ERD15)	X	5
At4g24973	Tapetum Determinant Protein (TD1)	X	5, 6, 11
At1g16210	DUF 1014 domain – unknown function	X	2
At5g26760	Unknown	X	7, 10

Screens, the number designated to the respective independent screen/s said clone was recovered from (see Table 4.1); Roman numerals above the screen number denote the number of unique clones coding for the same protein that were recovered from that screen. Based upon the known or inferred functions of conserved domains (Marchler-Bauer *et al.,* 2007), or predicted sub-cellular localization (Emanuelsson *et al*., 2000), sequenced clones were categorized into the following classes: A, Transcription factors, chromatin remodeling factors, RNA-binding proteins; B, Cytosolic proteins with a possible role in posttranslational modification or signaling; C, Chaperons, protein folding, degradation, proteases; D, Secreted and structural proteins; E, Ribosomal, chloroplastic, mitochondrial proteins, metabolic enzymes; X, unclassified.

Appendix E: Recovered clones that have not been re-tested

Gene	Descriptions	Class	Screens
At1g44900	MCM-like	A	2
At2g42310		X	1
At4g33780		X	10
At5g48160	DUF1423	X	2
At3g01650	RING-finger domain	B	11
At1g70810	Protein Kinase C	B	11
At1g73010	Putative Phosphatase	B	4
At4g01330	Kinase family protein	C	11
At2g30110	Ubiquitin activating enzyme (UBA1)	C	10
At1g05180	NEDD8-activating enzyme E1 (AXR1)	C	1
At4g24280	cpHsc70-1	C	2
At1g47540	Trypsin inhibitor 2	C	2
At3g14240	Subtilisin-like serine protease	C	2^{II}, 7
At1g11910	Aspartic proteinase	C	2^{II}
At1g62290	Aspartyl protease family protein	C	7
At1g62290	Pepsin A	C	11
At1G63120	ATRBL2 - serine-type endopeptidase	E	10
At2g16600	Cytosolic cyclophilin (ROC3)	C	11
At1g78830	Signal peptide selection derived (sps843)	D	11
At1g07920	Elongation factor 1-alpha	D	2
At1g21310	Proline Rich Extensin 5 family	D	2
At3g62360	PM5 collagenase	D	2
At5g58070	Outer mem. Lipo prot.-like	D	3
At5g26760	Myosin tail 1	D	10
At5G05520	Outer membrane OMP85 family protein	D	11
At3g01280	Putative porin	D	2
At2g17380	Clathrin assembly protein	D	2

Appendix E: Recovered clones that have not been re-tested (continued)

Gene	Descriptions	Class	Screens
At1g67700	Oligopeptidase	E	1
At3g21720	Isocitrate lyase	E	4
At2g17720	Prolyl 4-hydroxylase	E	1
At4g13430	Aconitase	E	2, 11
At2g01350	Nicotinate-nucleotide diphosphorylase	E	2
At2g42600	PEPC2	E	2
At3g58750	Citrate synthase -like	E	11
At1g72370	40 S ribosomal protein	E	6
At2g19450	TAG1 - diacylglycerol O-acyltransferase	E	11
At3g46830	Rab GTPase homolog A2c	E	11
At3g18490	Chloroplasid Nucleoid DNA binding prot.	E	2^{III}
At3g12800	Short-chain alcohol dehydrogenase	E	11
At1g29930	Putative Chlo a/b binding prot	E	2
At5g23940	Acyltransferase	E	12
At2g36530	Enolase (2-phospho-D-glycerate hydroylase)	E	12
At3g12120	FAD2/delta-12 desaturase	E	2
At4g03280	Rieske FeS prot	E	2
At3g48000	ALDH	E	2
At2g36530	Enolase	E	4
At5g43940	Alcohol dehydrogenase	E	5
At4g15530	Pyruvate phosphate dikinase family protein	E	5
At1g73270	40S ribosomal protein Lamin receptor like	E	6, 10
At1g68560	Alpha-xylosidase	E	7
At4g26860	Proline synthetase associated protein	E	7
At5g24380	Metal-nicotianamine transporter YSL2	E	7
At1g25350	Glutamine-tRNA ligase (OVA9)	E	10
At3g51670	SEC14 phosphoglyceride transfer	E	7
At1g72150	Phosphoglyceride transfer	E	7
At1g07240	UDP-glucoronosyl/UDP-glucosyl transferase	E	7, 9

Appendix E: Recovered clones that have not been re-tested (continued)

Gene	Descriptions	Class	Screens
At1g11260	Putative glucose transporter protein	E	7
At5g56710	60S ribosomal protein	E	10
At1G74020	SS2 (STRICTOSIDINE SYNTHASE 2)	E	10
At4g34030	MCCB (3-methylcrotonyl-CoA carboxylase)	E	10
At1G80600	Acetylornithine transaminase (ArgD)	E	10
At1g04270	40S ribosomal protein S15	E	10^{IV}
At2g20890	THYLAKOID FORMATION1 (PSB29)	E	10
At1G80600	Acetylornithine transaminase (ArgD)	E	10^{II}
At1G54130	RELA/SPOT HOMOLOG 3 (RSH3)	E	10
At5g09510	RPS15D	E	10
At1g74960	3-ketoacyl-ACP synthase (Kas4)	E	10
At4G30270	endo-xyloglucan transferase (SEN4)	E	11
At3g57610	Adenylosuccinate synthetase	E	11
At3g55430	β-1,3-gluconase	E	2
At5g13980	α-mannosidase	E	2
At4g03415	Protein phosphatase 2C family protein	E	10
At4g37990	Mannitol dehydrogenase 2	E	11
At1g07250	UDP-glucose glucosyltransferase	E	7

Screens, the number designated to the respective independent screen/s said clone was recovered from (see Table 4.1); Roman numerals above the screen number denote the number of unique clones coding for the same protein that were recovered from that screen. Based upon the known or inferred functions of conserved domains (Marchler-Bauer *et al.,* 2007), or predicted sub-cellular localization (Emanuelsson *et al*., 2000), sequenced clones were categorized into the following classes: A, Transcription factors, chromatin remodeling factors, RNA-binding proteins; B, Cytosolic proteins with a possible role in posttranslational modification or signaling; C, Chaperons, protein folding, degradation, proteases; D, Secreted and structural proteins; E, Ribosomal, chloroplastic, mitochondrial proteins, metabolic enzymes; X, unclassified.

References

Abe, M., Katsumata, H., Komeda, Y., and Takahash, T. (2003) Regulation of Shoot Epidermal Cell Differentiation by a Pair of Homeodomain Proteins in Arabidopsis. Development 130, 635-643

Abe, M., Takahashi, T., Komeda, Y. (2001) Identification of a Cis-Regulatory Element for L1 Layer-Specific Gene Expression, which is Targeted by an L1-Specific Homeodomain Protein. The Plant Journal 26, 487-494

Adamczyk, B, Lehti-Shiu, M., and Fernandez, D. (2007) The MADS-Domain Factors AGL15 and AGL18 Act Redundantly as Repressors of the Rloral Transition in Arabidopsis. The Plant Journal 50, 1007-1019

Alonso-Blanco, C., Gomez-Mena, C., Llorente, F., Koornneef, M., Salinas, J., and Martinez-Zapater, J. (2005) Genetic and Molecular Analyses of Natural Variation Indicate CBF2 as a Candidate Gene for Underlying a Freezing Tolerance Quantitative Trait Locus in Arabidopsis. Plant Physiology 139, 1304–1312

Arnaud, P., and Feil, R. (2006) MEDEA Takes Control of Its Own Imprinting. Cell 124, 468-470

Ausubel, F., Brent R., Kingston, R., Moore, D., Seidman, J., Smith, J., Struhl, K., eds. (1998). Current Protocols in Molecular Biology. New York, NY: John Wiley & Sons, Inc., 15.6.1–15.6.10.

Bastow, R., Mylne, J., Lister, C., Lippman, Z., Martienssen, R., and Dean, C. (2004) Vernalization Requires Epigenetic Silencing of FLC by Histone Methylation. Nature 427, 164–167

Baumbusch, L., Hughes, D., Galau, G., and Jakobsen, K. (2004) LEC1, FUS3, ABI3 and Em Expression Reveals No Correlation With Dormancy in Arabidopsis. Journal of Experimental Botany 55, 77-87

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

Bertos, N., Wang, A., and Yang, X. (2001) Class II Histone Deacetylases: Structure, Function, and Regulation. Biochemistry and Cell Biology 79, 243-252.

Benhamed, M., Bertrand, C., Servet, C and Zhou, D-X (2006) Arabidopsis GCN5, HD1, and TAF1/HAF2 Interact to Regulate Histone Acetylation Required for Light-Responsive Gene Expression. The Plant Cell 18, 2893-2903

Berger. F., and Gaudin. V. (2003) Chromatin Dynamics and Arabidopsis Development. Chromosome Research 11, 277-304

Bewley, J. (1997) Seed Germination and Dormancy. The Plant Cell 9, 1055-1066

Bewley, J., and Black, M. (1994) Seeds: Physiology of Development and Germination. Kluwer Academic Publisher, pg 1.

Blazquez, M., and Weigel, D. (2000) Integration of Floral Inductive Signals in Arabidopsis. Nature 404, 889-892.

Boavida, L., Becker, J., and Feijó, J. (2005a) The Making of Gametes in Higher Plants. The International Journal of Developmental Biology 49, 595-614

Boavida, L., Vieira, A., Becker, J., and Feijó, J. (2005b) Gametophyte Interaction and Sexual Reproduction: How Plants Make a Zygote. The International Journal of Developmental Biology 49, 615-632

Boss, P., Bastow, R., Mylne, J., and Dean. C. (2004) Multiple Pathways in the Decision to Flower: Enabling, Promoting, and Resetting. The Plant Cell 16, S18–S31

Boutilier, K., Offringa, R., Sharma, V., Kieft, H., Ouellet, T., Zhang, L., Hattori, J., Liu, C-M., van Lammeren, A., Miki, B., Custers, J., and van Lookeren-Campagne, M. (2002) Ectopic Expression of BABY BOOM Triggers a Conversion from Vegetative to Embryonic Growth. The Plant Cell 14, 1737–1749

Bouvet, P., Matsumoto, K., and Wolffe, A. (1995) Sequence-Specific RNA Recognition by the Xenopus Y-Box Proteins. An Essential Role for the Cold Shock Domain. Journal of Biological Chemistry 270, 28297-28303

Bouvet, P., and Wolffe, A. (1994) A Role for Transcription and FRGY2 in Masking Maternal mRNA Within Xenopus oocytes. Cell 77, 931-941

Bowler, C., Benvenuto, G., Laflamme, P., Molino, D., Probst, A., Tariq, M., and Paszkowski, J. (2004) Chromatin Techniques for Plant Cells. The Plant Journal 39, 776–789

Burley, S., and Kamada, K (2002) Transcription Factor Complexes. Current Opinion in Structural Biology 12, 225–230

Busch, M., Mayer, U., and Juergens, G. (1996) Molecular Analysis of the Arabidopsis Pattern Formation Gene GNOM: Gene Structure and Intragenic Complementation. Molecular Genomics and Genetics 250, 681-691

Braybrook, S., Stone, S., Park, S., Bui, A., Le, B., Fischer, R., Goldberg, R., and Harada, J. (2006) Genes Directly Regulated by LEAFY COTYLEDON2 Provide Insight into the Control of Embryo Maturation and Somatic Embryogenesis. Proceedings of the National Academy of Science 103, 3468–3473

Castillejo, C., Romera-Branchat, M., and Pelaz, S. (2005) A New Role of the Arabidopsis SEPALLATA3 Gene Revealed by its Constitutive Expression. The Plant Journal 43, 586–596

Causier, B., Cook, H., and Davies, B. (2003) An Antirrhinum Ternary Complex Factor Specifically Interacts Wwith C-Function and SEPALLATA-Like MADS-Box Factors. Plant Molecular Biology 52, 1051-62

Causier, B., and Davies, B. (2002) Analysing Protein-Protein Interactions with the Yeast Two-Hybrid System. Plant Molecular Biology 50, 855–870

Chalfun-Junior, A., Franken, J., Mes, J., Marsch-Martinez, N., Pereira, A., and Angenent, G. (2005) ASYMMETRIC LEAVES2-LIKE1 Gene, a Member of the AS2/LOB Family, Controls Proximal–Distal Patterning in Arabidopsis Petals. Plant Molecular Biology 57, 559–575

Cheng, Y., Kato, N., Wang, W., Li, J., and Chen, X. (2003) Two RNA Binding Proteins, HEN4 and HUA1, Act in the Processing of AGAMOUS Pre-mRNA in Arabidopsis thaliana. Developmental Cell 4, 53–66

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T., Higgins, D., and Thompson, J. (2003) Multiple Sequence Alignment with the Clustal Series of Programs. Nucleic Acids Research 31, 3497-3500

Cheng, S., and Bishop, J. (2002) Suppressor of Fused Represses Gli-Mediated Transcription by Recruiting the SAP18-mSin3 Corepressor Complex. Proceedings of the National Academy of Science 99, 5442-5447

Cho, S., Jang, S., Chae, S., Chung, K., Moon, Y., An, G. and Jang, S. (1999) Analysis of the C-Terminal Region of Arabidopsis thaliana APETALA1 as a Transcription Activation Domain. Plant Molecular Biology 40, 419–429

Choi, Y., Harada, J., Goldberg, R., and Fischer, L. (2004) An Invariant Aspartic Acid in the DNA glycosylase Domain of DEMETER is Necessary for Transcriptional Activation of the Imprinted MEDEA gene. Proceedings of the National Academy of Sciences 101, 7481-7486

Citovsky, V., Lee, L., Vyas, S., Glick, E., Chen, M., Vainstein, A., Gafni, Y., Gelvin, S., and Tzfira, T. (2006) Subcellular Localization of Interacting Proteins by Bimolecular Fluorescence Complementation in Planta. Journal of Molecular Biology 362, 51120-1131

Clough S., and Bent A. (1998) Floral Dip: a Simplified Method for Agrobacterium-Mediated Transformation of Arabidopsis thaliana. The Plant Journal 16, 735-43.

Colcombet, J., Boisson-Dernier A., Ros-Palau, R., Vera, C., and Schroeder, J. (2005) Arabidopsis SOMATIC EMBRYOGENESIS RECEPTOR KINASES1 and 2 are Essential for Tapetum Development and Microspore Maturation. The Plant Cell 17, 3350-3361

Cook, D., Fowler, S., Fiehn, O., and Thomashow M. (2004) A Prominent Role for the CBF Cold Response in Configuring the Low-Temperature of Arabidopsis. Proceedings of the National Academy of Sciences 101, 15243–15248

Corbesier, L., Vincent, C., Jang, S., Fornara, F., Fan, Q., Searle, I., Giakountis, A., Farrona, S., Gissot, L., Turnbull, C., and Coupland, G. (2007) FT Protein Movement Contributes to Long-Distance Signaling in Floral Induction of Arabidopsis. Science 316, 1030-1033

Courey, A., and Jia, S. (2001) Transcriptional Repression: The Long and the Short of it. Genes and Development 15, 2786–2796

Cseke L., Ravinder N., Pandey A., and Podila G. (2007) Identification of PTM5 Protein Interaction Partners, a MADS-Box Gene Involved in Aspen Tree Vegetative Development. Gene 39, 209-222.

Curaba, J., Moritz, T., Blervaque, R., Parcy, F., Raz, V., Herzog, M., and Vachon, G. (2004) AtGA3ox2, a Key Gene Responsible for Bioactive Gibberellin Biosynthesis, Is Regulated during Embryogenesis by LEAFY COTYLEDON2 and FUSCA3 in Arabidopsis. Plant Physiology 136, 3660–3669

Dalrymple M., McGeoch D., Davison A., and Preston C. (1985) DNA Sequence of the Herpes Simplex Virus Type 1 Gene Whose Product is Responsible for Transcriptional Activation of Immediate Early Promoters. Nucleic Acids Research 13, 7865-7879.

Ditta, G., Pinyopich A., Robles, P., Pelaz, S., Yanofsky, M. (2004) The SEP4 Gene of Arabidopsis thaliana Functions in Floral Organ and Meristem Identity. Current Biology 14, 1935–1940

de Folter, S., and Angenent, G. (2006) Trans Meets Cis in MADS Science. Trends in Plant Science 11, 224-31.

de Folter S., Immink R., Kieffer M., Parenicova L., Henz S., Weigel D., Busscher M., Kooiker M., Colombo L., Kater M., Davies B., Angenent G. (2005) Comprehensive Interaction Map of the Arabidopsis MADS Box Transcription Factors. The Plant Cell 17, 1424-33

Deng, W., Liu, C., Pei, Y., Deng, X., Niu, L., and Cao, X. (2007) Involvement of the Histone Acetyltransferase AtHAC1 in the Regulation of Flowering Time Via Repression of FLOWERING LOCUS C in Arabidopsis. Plant Physiology 143, 1660-1668

de Ruijter, A., Van Gennip, A., Caron, H., Kemp, D., and Kuilenburg, A. (2003) Histone Deacetylases (HDACs): Characterization of the Classical HDAC Family. The Biochemical Journal 370, 737-749

Dubnicoff, T., Valentine, S., Chen, G., Shi, T., Lengyel, J., Paroush, Z., and Courey, A. (1997) Conversion of Dorsal from an Activator to a Repressor by the Global Corepressor Groucho. Genes and Development 11, 2952-2957.

Egea-Cortines, M., Saedler, H. and Sommer, H. (1999) Ternary Complex Formation Between the MADS-Box Proteins SQUAMOSA, DEFICIENS, and GLOBOSA is Involved in the Control of Floral Architecture in Antirrhinum majus. The EMBO Journal 18, 5370–5379.

Ehlert, A., Weltmeier, F., Wang, X., Mayer, C., Smeekens, S., Vicente-Carbajosa, J., and Droge-Laser, W. (2006) Two-Hybrid Protein–Protein Interaction Analysis in Arabidopsis Protoplasts: Establishment of a Heterodimerization Map of Group C and Group S bZIP Transcription Factors. The Plant Journal 46, 890–900

Emanuelsson, O., Nielsen, H., Brunak, S., and von Heijne, G. (2000) Predicting Subcellular Localization of Proteins Based on Their N-terminal Amino Acid Sequence. Journal of Molecular Biology 300, 1005-1016

Espinas, M., Canudas, S., Fanti, L., Pimpinelli, S., Casanova, J., Azorin, F. (2000) The GAGA Factor of Drosophila Interacts with SAP18, a Sin3-Associated Polypeptide. EMBO Reports 1, 253-9

Fan, H., Hu, Y., Tudor, M., and Ma, H. (1997) Specific Interactions Between the K Domains of AG and AGLs, Members of the MADS Family of DNA Binding Proteins. The Plant Journal 12, 999-1010

Fang S., and Fernandez D. (2002) Effect of Regulated Overexpression of the MADS Domain Factor AGL15 on Flower Senescence and Fruit Maturation. Plant Physiology 130, 78-89

Favaro R, Immink R., Ferioli V, Bernasconi B, Byzova M, Angenent G., Kater M, and Colombo L. (2002) Ovule-Specific MADS-Box Proteins Have Conserved Protein-Protein Interactions in Monocot and Dicot Plants. Molecular Genetics and Genomics 268, 152-159.

Favaro, R., Pinyopich, A., Battaglia, R., Kooiker, M., Borghi, L., Ditta, G., Yanofsky, M., Kater, M., and Colombo, L. (2003) MADS-Box Protein Complexes Control Carpel and Ovule Development in Arabidopsis. The Plant Cell 15, 2603-11

Fernandez, D., Heck, G., Perry, S., Patterson, S., Bleecker, A., and Fang, S-C. (2000) The Embryo MADS Domain Factor AGL15 Acts Postembryonically: Inhibition of Perianth Senescence and Abscission via Constitutive Expression. The Plant Cell 12, 183–197

Ferrándiz, C., Liljegren, S., and Yanofsky, M. (2000) Negative Regulation of the SHATTERPROOF Genes by FRUITFULL During Arabidopsis Fruit Development. Science 289, 436-438.

Fernando, S., and Gamage, C. (2000) Abscisic Acid Induced Somatic Embryogenesis in Immature Embryo Explants of Coconut (Cocos nucifera L.). Plant Science 151, 193-198

Fields, S. and Song, O (1989) A Novel Genetic System to Detect Protein-Protein Interactions. Nature 340, 245–246.

Finch-Savage, W., and Leubner-Metzger, G. (2006) Seed Dormancy and the Control of Germination. The New Phytologist 171, 501-523.

Fowler, S., Cook, D., and Thomashow, M. (2005) Low Temperature Induction of Arabidopsis CBF1, 2, and 3 Is Gated by the Circadian Clock. Plant Physiology 137, 961–968

Franks, R., Wang, C., Levin, J., and Liu, Z. (2002) SEUSS, a Member of a Novel Family of Plant Regulatory Proteins, Represses Floral Homeotic Gene Expression with LEUNIG. Development 129, 253-263.

Fujita H, Takemura M, Tani E, Nemoto K, Yokota A, Kohchi T. (2003) An Arabidopsis MADS-Box Protein, AGL24, is Specifically Bound to and Phosphorylated by Meristematic Receptor-Like Kinase (MRLK). Plant Cell Physiology 44, 735–742

Furger, A., O'Sullivan, J., Binnie, A., Lee, B., and Proudfoot, N. (2002) Promoter proximal splice sites enhance transcription. Genes and Development 16, 2792–2799.

Fusaro, A., Bocca, S., Ramos, R., Barrôco, R., Magioli, C, Coutinho, T., Rangel-Lima, C., De Rycke, R., Inzé, D., Engler, G., and Sachetto—Martins, G. (2007) AtGRP2, a Cold-Induced Nucleo-Cytoplasmic RNA Binding Protein, has a Role in Flower and Seed Development. Planta 225, 1339-1351

Gaj , M., Zhang, S., Harada, J., and Lemaux, P. (2005) Leafy Cotyledon Genes are Essential for Induction of Somatic Embryogenesis of Arabidopsis. Planta 222, 977–988

Gamboa A, Paez-Valencia J, Acevedo G., Vazquez-Moreno L, Alvarez-Buylla R., (2001) Floral Transcription Factor AGAMOUS Interacts In Vitro With a Leucine-Rich Repeat and an Acid Phosphatase Protein Complex. Biochemical and Biophysical Research Communications 288, 1018-26

Gao, M-J, Parkin, I., Lydiate, D., and Hannoufa, A. (2004) An Auxin-Responsive SCARECROW-Like Transcriptional Activator Interacts with Histone Deacetylase. Plant Molecular Biology 55,417–431

Gao, M-J., Schafer, U., Parkin, I., Hegedus, D., Lydiate D., and Hannoufa, A. (2003) A Novel Protein from Brassica napus has a Putative KID Domain and Responds to Low Temperature. The Plant Journal 33, 1073-1086

Gilmour, S., Fowler S., and Thomashow, M. (2004) Arabidopsis Transcriptional Activators CBF1, CBF2, and CBF3 Have Matching Functional Activities. Plant Molecular Biology 54, 767–781

Girke, T., Todd, J., Ruuska, S., White, J., Benning, C., and Ohlrogge, J. (2000) Microarray Analysis of Developing Arabidopsis Seeds. Plant Physiology 124, 1570-1581

Giuliodori, A., Brandi, A., Gualerzi, C., and Pon, C. (2004) Preferential Translation of Cold-Shock mRNAs During Cold Adaptation. RNA 10, 265-276.

Goderis, I., De Bolle, M., François, I., Wouters, P., Broekaert, W., and Cammue, B. (2002) A Set of Modular Plant Transformation Vectors Allowing Flexible Insertion of up to Six Expression Units. Plant Molecular Biology 50, 17-27.

Golden, T., Schauer, S., Lang, J., Pien, S., Mushegian, A., Grossniklaus, U., Meinke, D., and Ray, A. (2002) SHORT INTEGUMENTS1/SUSPENSOR1/CARPEL FACTORY, a Dicer Homolog, is a Maternal Effect Gene Required for Embryo Development in Arabidopsis. Plant Physiology 130, 808-822

Golovkin, M., and Reddy, A. (2003). A Calmodulin-Binding Protein from Arabidopsis has an Essential Role in Pollen Germination. Proceedings of the National Academy of Science 100, 10558–10563

Goldberg, R.B., de Paiva, G., and Yadegari, R. (1994) Plant Embryogenesis: Zygote to Seed. Science 266: 605-614

Gómez-Mena, C., de Folter, S., Costa, M., Angenent, G., and Sablowski, R. (2005) Transcriptional Program Controlled by the Floral Homeotic Gene AGAMOUS During Early Organogenesis. Development 132, 429-438

Goodin, M., Dietzgen, R., Schichnes, D., Ruzin, S., and Jackson, A. (2002) pGD Vectors: Versatile Tools for the Expression of Green and Red Fluorescent Protein Fusions in Agro Infiltrated Plant Leaves. The Plant Journal 31, 375-383

Grasser, K. (2005) Emerging Role for Transcript Elongation in Plant Development. Trends in Plant Science 10, 484-490

Graumann, P., and Marahiel, M. (1998) A Superfamily of Proteins that Contain the Cold-Shock Domain. Trends in Biochemical Sciences 23, 286-290.

Green, M., Maskos, K., and Landry, S. (1998) Role of the J-domain in the Cooperation of HSP40 with HSP70. Proceedings of the National Academy of Science 95, 6108–6113

Grewal, S., and Elgin, S. (2002) Heterochromatin: New Possibilities for the Inheritance of Structure. Current Opinion in Genetics and Development 12, 178–87

Gu, Q., Ferrándiz, C., Yanofsky, M., and Martienssen, R. (1998) The FRUITFULL MADS-box Gene Mediates Cell Differentiation During Arabidopsis Fruit Development. Development 125, 1509-1517

Gualerzi, C., and Pon, C. (1990) Initiation of mRNA translation in prokaryotes. Biochemistry 29, 5881-5889

Gutierrez-Cortines, M., and Davies, B. (2000) Beyond the ABCs: Ternary Complex Formation in the Control of Floral Organ Identity. Trends Plant Science 5, 471-476

Hall, T. (2005) Multiple Modes of RNA Recognition by Zinc Finger Proteins. Current Opinion in Structural Biology 15, 367–373

Hall, D., Zhu, H., Zhu, X., Royce, T., Gerstein, M., and Snyder, M. (2004) Regulation of Gene Expression by a Metabolic Enzyme. Science 306, 482-484

Harding, E., Tang, W., Nichols, K., Fernandez, D. and Perry, S. (2003) Expression and Maintenance of Embryogenic Potential is Enhanced Through Constitutive Expression of AGAMOUS-Like 15. Plant Physiology 133, 653-663.

Hartl, F.U., and Hayer-Hartl, M. (2007) Molecular Chaperones in the Cytosol: from Nascent Chain to Folded Protein. Science 295, 1852-1858

Hartmann, U, Höhmann, S., Nettesheim, K., Wisman, E., Saedler, H., and Huijser, P. (2000) Molecular Cloning of SVP: a Negative Regulator of the Floral Transition in Arabidopsis. The Plant Journal 21, 351-360

Hatfield, P., Gosink, M., Carpenter, T., and Viestra, R. (1997). The Ubiquitin Activating Enzyme (E1) Gene Family in Arabidopsis thaliana. The Plant Journal 11, 213-226

He, Y., Michaels, S., and Amasino R. (2003). Regulation of Flowering Time by Histone Acetylation in Arabidopsis. Science 302, 1751-1754.

Heck, G., Perry, S., Nichols, K. and Fernandez, D. (1995) AGL15, a MADS Domain Protein Expressed in Developing Embryos. The Plant Cell 7, 1271–1282

Henderson, I., and Dean, C. (2004) Control of Arabidopsis Flowering: the Chill before the Bloom. Development 131, 3829-3838

Hecht, V., Vielle-Calzada, J., Hartog, M., Schmidt, E., Boutilier, K., Grossniklaus, U., de Vries, S. (2001) The Arabidopsis SOMATIC EMBRYOGENESIS RECEPTOR KINASE 1 Gene is Expressed in Developing Ovules and Embryos and Enhances Embryogenic Competence in Culture. Plant Physiology 127, 803-816

Hollung, K., Espelund, M., Schou, K., and Kjetill S. Jakobsen, K. (1997) Developmental, Stress and ABA Modulation of mRNA Levels for bZip Transcription Factors and Vp1 in Barley Embryos and Embryo-Derived Suspension Cultures. Plant Molecular Biology 35, 561–571

Honma, T and Goto, K. (2000) The Arabidopsis Floral Homeotic Gene PISTILLATA is Regulated by Discrete Cis-Elements Responsive to Induction and Maintenance Signals. Development, 127, 2021-30

Honma, T. and Goto, K. (2001) Complexes of MADS-box Proteins are Sufficient to Convert Leaves into Floral Organs. Nature 409, 525–529.

Hsieh, T., and Fischer, R. (2005) Biology of Chromatin Dynamics. Annual Review of Plant Biology 56, 327–251

Hu, H., Xiong, L., and Yang, Y. (2005) Rice SERK1 Gene Positively Regulates Somatic Embryogenesis of Cultured Cell and Host Defense Response Against Fungal Infection. Planta 222, 107–117

Ito, T., Wellmer, F., Yu, H., Das, P., Ito, N., Alves-Ferreira, M., Riechmann, J., and Meyerowitz, E. (2004) The Homeotic Protein AGAMOUS Controls Microsporogenesis by Regulation of SPOROCYTELESS. Nature 430, 356-360

Ikeda, Y., Kobayashi, Y., Yamaguchi, A., Abe, M., and Araki, T. (2007) Molecular Basis of Late-Flowering Phenotype Caused by Dominant Epi-Alleles of the FWA Locus in Arabidopsis. Plant Cell Physiology 48, 205–220

Ikeda-Iwai, M., Satoh, S, and Kamada, H. (2002) Establishment of a Reproducible Tissue Culture System for the Induction of Arabidopsis Somatic Embryos. Journal of Experimental Botany 53, 1575-1580

Ikeda, M., Umehara, M., and Kamada, H. (2006) Embryogenesis-Related Genes; Its Expression and Roles During Somatic and Zygotic Embryogenesis in Carrot and Arabidopsis. Plant Biotechnology 23, 153–161

Immink, R., Gadella, T., Ferrario, S., Busscher, M., Angenent, G. (2002) Analysis of MADS Box Protein–Protein Interactions in Living Plant Cells. Proceedings of the National Academy of Sciences 99, 2416–2421

Jack, T. (2004) Molecular and Genetic Mechanisms of Floral Control. The Plant Cell 16, S1–S17,

Jackoby, M., Drodge-Laser, W., Vincente-Carbajosa, J., Thedemann, J., Kroj, T., and Paryy, F. (2002) bZIP Transcription Factors in Arabidopsis. Trends in Plant Science 7, 106-111

Jang, S., An, K., Lee, S., An, G. (2002) Characterization of Tobacco MADS-Box Genes Involved in Floral Initiation. Plant Cell Physiology 43, 230-238

Jenik, P., Gillmor, C., and Lukowitz, W (2007) Embryonic Patterning in Arabidopsis thaliana. Annual Review of Cell and Developmental Biology 23, 207–736

Jenuwein, T., and Allis, C. (2001) Translating the Histone Code. Science 293, 1074-1080

Jiang, W., Hou, Y., and Inouye, M. (1997) CspA, the Major Cold-Shock Protein of Escherichia coli, is an RNA Chaperone. Journal of Biological Chemistry 272, 196-202.

Jones, P., Krah, R., Tafuri, S., and Wolffe, A. (1992). DNA Gyrase, CS7.4, and the Cold Shock Response in Escherichia coli. Journal of Bacteriology 174, 5798-5802.

Jones, P., VanBogelen, R., and Neidhardt, F. (1987) Induction of Proteins in Response to Low Temperature in Escherichia coli. Journal of Bacteriology 169, 2092-2095

Jurgens, G. (2001) Apical-Basal Pattern Formation in Arabidopsis Embryogenesis. The EMBO Journal 20, 3609-3616

Kariola, T., Brader, G., Helenius, E., Li, J., Heino, P., and E. Palva, E. (2006) EARLY RESPONSIVE TO DEHYDRATION 15, a Negative Regulator of Abscisic Acid Responses in Arabidopsis. Plant Physiology 142, 1559–1573

Karlova R, Boeren S, Russinova E, Aker J, Vervoort J, de Vries S. (2006) The Arabidopsis SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE1 Protein Complex Includes BRASSINOSTEROID-INSENSITIVE1. The Plant Cell 18, 626-38.

Karlson, D., and Imai, R. (2003) Conservation of the Cold Shock Domain Protein Family in Plants. Plant Physiology 131, 12-15

Karlson, D., Nakaminami, K., Toyomasu, T., and Imai, R. (2002) A Cold-Regulated Nucleic Acid-Binding Protein of Winter Wheat Shares a Domain with Bacterial Cold Shock Proteins. Journal of Biological Chemistry 277, 35248-35256

Kato, M., Hata, N., Banerjee, N., Futcher, M., and Zhang, M. (2004) Identifying Combinatorial Regulation of Transcription Factors and Binding Motifs. Genome Biology 5, R56

Kaufmann, K., Melzer, R., and Theissen, G. (2005) MIKC-Type MADS-Domain Proteins: Structural Modularity, Protein Interactions and Network Evolution in Land Plants. Gene 347, 183–198

Keene, J. (2007) RNA Regulons: Coordination of Post-Transcriptional Events. Nature Reviews Genetics 8, 533-543

Kim, Y-O., Kim, J., and Kang, H. (2005) Cold-Inducible Zinc Finger-Containing Glycine-Rich RNA-Binding Protein Contributes to the Enhancement of Freezing Tolerance in Arabidopsis thaliana. The Plant Journal 42, 890–900

Kim, J., Park, S., Jang, B., Jung C-H., Ahn, S., Goh, C-H., Cho, K., Han, O., and Kang, H. (2007a) Functional Characterization of a Glycine-Rich RNA-Binding Protein 2 in Arabidopsis thaliana Under Abiotic Stress Conditions. The Plant Journal 50, 439–451

Kim, J., Park, S., Kwak, K., Kim, Y., Kim, J., Song, J., Jang, B., Jung., C., and Kang, H. (2007b) Cold Shock Domain Proteins and Glycine-Rich RNA-Binding Proteins from Arabidopsis thaliana Can Promote the Cold Adaptation Process in Escherichia coli. Nucleic Acids Research 35, 506–516.

Kingsley, P., and Palis, J. (1994) GRP2 Proteins Contain Both CCHC Zinc Fingers and a Cold Shock Domain. The Plant Cell 6,1522-1523

Kinoshita, T., Yadegari, R., Harada, J., Goldberg R., and Fischer, R. (1999) Imprinting of the MEDEA Polycomb Gene in the Arabidopsis Endosperm. The Plant Cell 11, 1945-1952

Ko, T-S., Nelson, R., and Korban, S. (2004) Screening Multiple Soybean Cultivars (MG 00 to MG VIII) for Somatic Embryogenesis Following Agrobacterium-Mediated Transformation of Immature Cotyledons. Crop Science 44, 1825-1831

Köhler, C., Hennig, L., Spillane, C., Pien, S., Gruissem, W., Grossniklaus, U. (2003) The Polycomb-Group Protein MEDEA Regulates Seed Development by Controlling Expression of the MADS-Box Gene PHERES1. Genes and Development 17, 1540-1553

Köhler, C., Page, D., 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

Kohno, K., Izumi, H., Uchiumi, T., Ashizuka, M., and Kuwano, M. (2003) The Pleiotropic Functions of the Y-Box Binding Protein, YB-1. BioEssays 25, 691–698

Kornblihtt, A., de la Mata, M., Fededa, J., and Munoz, M., and Nogues, G. (2004) Multiple Links Between Transcription and Splicing. RNA 1489-1498

Kotake, T., Takada, S., Nakahigashi, K., Ohto, M., and Goto, K. (2003) Arabidopsis TERMINAL FLOWER 2 Gene Encodes a Heterochromatin Protein 1 Homolog and Represses both FLOWERING LOCUS T to Regulate Flowering Time and Several Floral Homeotic Genes. Plant Cell Physiology 44, 555–564

Koornneef, M., Bentsink, L., and Hilhorst, H. (2002) Seed Dormancy and Germination. Current Opinions in Plant Biology 5, 33-36

La Teana, A., Brandi, A., Falconi, M., Spurio, R., Pon, C., Gualerzi, C. (1991) Identification of a Cold Shock Transcriptional Enhancer of the Escherichia coli Gene Encoding Nucleoid Protein H-NS. Proceedings of the National Academy of Sciences 88, 10907-10911

Laux, T., Wurschum, T., and Breuninger, H. (2004) Genetic Regulation of Embryonic Pattern Formation. The Plant Cell, 16: S190–S202

Lee, J., Yoo, S., Park, S., Hwang, I., Lee, J., and Ahn, J. (2007) Role of SVP in the Control of Flowering Time by Ambient Temperature in Arabidopsis. Genes and Development 21, 397-402

Lehti-Shiu, M., Adamczyk, B., and Fernandez, D. (2005) Expression of MADS-Box Genes During the Embryonic Phase in Arabidopsis. Plant Molecular Biology 58, 89-107

Liljegren, S., Ditta, G., Eshed, Y., Savidge, B., Bowman, J., and Yanofsky, M. (2000) SHATTERPROOF MADS-Box Genes Control Seed Dispersal in Arabidopsis. Nature 404, 766-770

Lim, J., Moon, Y., An, G., Jang, S. (2000) Two Rice MADS Domain Proteins Interact With OsMADS1. Plant Molecular Biology 44, 513-27

Liu, Y-Z, Dawson, S., Gerster, T., Friedl, E., Pengue, G., Matthias, P., Lania, L., and Latchman, D. (1996) The Ability of the Inhibitory Domain of the POU Family Transcription Factor Oct-2 to Interfere with Promoter Activation by Different Classes of Activation Domains Is Dependent upon the Nature of the Basal Promoter Elements. Journal of Biological Chemistry 271, 20853-20860

Liu, Z., and Little, J. (1998) The Spacing Between Binding Sites Controls the Mode of Cooperative DNA-Protein Interactions: Implications for Evolution of Regulatory Circuitry. Journal of Molecular Biology 278, 331-338

Liu, Z., and Meyerowitz, E. (1995) LEUNIG Rregulates AGAMOUS Expression in Arabidopsis Flowers. Development 121, 975-91.

Liu, H., Wang, S., Yu, X., Yu, J., He, X., Zhang, S., Shou., H and Wu, P. (2005) ARL1, a LOB-Ddomain Protein Required for Adventitious Root Formation in Rice. The Plant Journal 43, 47–56

Liu, L., White, M., and MacRae T. (1999) Transcription Factors and Their Genes in Higher Plants Functional Domains, Evolution and Regulation. European Journal of Biochemistry 262, 247-257

Loidl, P. (2004) A Plant Dialect of the Histone Language. Trends in Plant Science 9, 84-90.

Loo, S., and Rine, J. (1994) Silencers and Domains of Generalized Repression. Science 264, 1768-1771

Lorkovic, Z., and Barta, A. (2002) Genome Analysis: RNA Recognition Motif (RRM) and K Homology (KH) Domain RNA-Binding Proteins from the Flowering Plant Arabidopsis thaliana. Nucleic Acids Research 30, 623-635

Lotan, T., Ohto, M., Yee, K., West, M., Lo, R., Kwong, R., Yamagishi, K., Fischer, R. L., Goldberg, R., and Harada, J. (1998) Arabidopsis LEAFY COTYLEDON1 is Sufficient to Induce Embryo Development in Vegetative Cells. Cell 93, 1195-1205

Ma, H. (1999) Seed Development: With or Without Sex? Current Biology 9, R636–R639

Mahajan, S., and Tuteja, N. (2005) Cold, Salinity and Drought stresses: An Overview. Archives of Biochemistry and Biophysics 444, 139-158

Malcomber, S., and Kellogg, E. (2005) SEPALLATA Gene Diversification: Brave New Whorls. Trends in Plant Science 10, 427-435

Marchler-Bauer, A., Anderson, J., Derbyshire, M., DeWeese-Scott, C., Gonzales, N., Gwadz, M., Hao, L., He, S., Hurwitz, D., Jackson, J., Ke, Z., Krylov, D., Lanczycki, C., Liebert, C., Liu, C., Lu, F., Lu, S., Marchle, G., Mullokandov, M., Song, S., Thanki, N., Yamashita, R., Yin, J., Zhang, D., and Bryant, S. (2007) "CDD: a Conserved Domain Database for Interactive Domain Family Analysis.", Nucleic Acids Research 35, D237-40

Masiero. S., Imbriano. C, Ravasio. F., Favaro. R., Pelucchi. N., Gorla. M., Mantovani. R., Colombo. L., Kater, M. (2002) Ternary Complex Formation Between MADS-Box Transcription Factors and the Histone Fold Protein NF-YB. Journal of Biological Chemistry 277, 26429-26435.

Mantovani R. (1998) A Survey of 178 NF-Y Binding CCAAT Boxes. Nucleic Acids Research 26, 1135–1143

Matsumoto, T., Pardo, J., Takeda, S., Bressan, R., and Hasegawa, P. (2001) Tobacco and Arabidiopsis SLT1 Mediate Salt Tolerance of Yeast. Plant Molecular Biology 45, 489–500

Maurer-Stroh, S., Dickens, D., Hughes-Davies, L Tony Kouzarides, L., Eisenhaber, F., and Ponting, C. (2003) The Tudor Domain 'Royal Family': Tudor, Plant Agenet, Chromo, PWWP and MBT domains. Trends in Biochemical Sciences 28, 69-74

McCallum, S., Bazan, J., Merchant, M., Yin, J., Pan, B., de Sauvage, F., and Fairbrother, W. (2006) Structure of SAP18: A Ubiquitin Fold in Histone Deacetylase Complex Assembly. Biochemistry 45, 11974-11982

McCormick, S. (2004) Control of Male Gametophyte Development. The Plant Cell 16, S142-S153

McElver, J., Tzafrir, I., Aux, G., Rogers, R., Ashby, C., Smith, K., Thomas, C., Schetter, A., Zhou, Q., Cushmann, M., Tossberg, J., Nickle, T., Levin, J., Law, M., Meinke, D., and Patton, D (2001) Insertional mutagenesis of genes required for seed development in Arabidopsis thaliana. Genetics 159, 1751-1763

McGonigle, B., Bouhidel, K., and Irish V. (1996) Nuclear Localization of the Arabidopsis APETALA3 and PISTILLATA Homeotic Gene Products Depends on their Simultaneous Expression. Genes and Development 10, 1812-1821

Meinke, D., Franzmann, L., Nickle, T., and Yeung, E. (1994) Leafy Cotyledon Mutants of Arabidopsis. The Plant Cell 6, 1049-1064

Messenguy, F., and Dubois, E. (2003) Role of MADS Box Proteins and Their Cofactors in Combinatorial Control of Gene Expression and Cell Development. Gene 316, 1-21.

Meyer, P. (2001) Chromatin Remodelling. Current Opinion in Plant Biology 4, 457–462

Meyerowitz, E. (2002) Plants Compared to Animals: The Broadest Comparative Study of Development. Science 295, 1482-1485

Mierny, J. (2001) The J-domain Proteins of Arabidopsis thaliana: An Unexpectedly Large and Diverse Family of Chaperones. Cell Stress and Chaperones 6, 209–218

Michaels, S., Ditta, G., Gustafson-Brown, C., Pelaz, S., Yanofsky, M., Amasino, R. (2003) AGL24 Acts as a Promoter of Flowering in Arabidopsis and is Positively Regulated by Vernalization. The Plant Journal 33, 867-874

Minter, A., Brennan, B., and Mapp, A. (2004) A Small Molecule Transcriptional Activation Domain. Journal of the American Chemical Association 126, 10504–10505.

Mitsuda, N., Hisabori, T., Takeyasu, K and Sato, M. (2004) VOZ; Isolation and Characterization of Novel Vascular Plant Transcription Factors with a One-Zinc Finger from Arabidopsis thaliana. Plant Cell Physiology 45, 845–854

Mizukami, Y., Huang, H., Tudor, M., Hu, Y., and Ma, H. (1996) Functional Domains of the Floral Regulator AGAMOUS: Characterization of the DNA Binding Domain and Analysis of Dominant Negative Mutations. The Plant Cell 8, 831-845

Mockler, T., Yu, X., Shalitin, D., Parikh, D., Michael, T., Liou , J., Huang, J., Smith, Z., Alonso, J., Ecker, J., Chory, J., and Lin, C. (2004) Regulation of Flowering Time in Arabidopsis by K Homology Domain Proteins. Proceedings of the National Academy of Science 101, 12759-12764

Mongelard, F., and Corces, V. (2001) Two Insulators Are Not Better Than One. Nature Structural Biology 8, 192-194

Moon, Y., Chen, L., Pan, R., Chang, H., Zhu, T., Maffeo, D., and Sung, Z. (2003a) EMF Genes Maintain Vegetative Development by Repressing the Flower Program in Arabidopsis. The Plant Cell 15, 681-693

Moon, J., Suh, S., Lee, H., Choi, K., Hong, C., Paek, N., Kim, S., and Lee, I. (2003b) The SOC1 MADS-Box Gene Integrates Vernalization and Gibberellin Signals for Flowering in Arabidopsis. The Plant Journal 35, 613-623

Moon, Y., Kang, H., Jung, J., Jeon, J., Sung, S., and An, G. (1999) Determination of the Motif Responsible for Interaction Between the Rice APETALA1/AGAMOUS-LIKE9 Family Proteins Using a Yeast Two-Hybrid System. Plant Physiology 120, 1193-204

Mordhorst, A., Voerman, K., Hartog, M., Meijer, E., van Went, J., Koornneef, M., and de Vries, S. (1998) Somatic Embryogenesis in Arabidopsis thaliana is Facilitated by Mutations in Genes Repressing Meristematic Cell Divisions. Genetics 149, 549–563

Moss, E., and Tang, L. (2003) Conservation of the Heterochronic Regulator LIN-28, its Developmental Expression and microRNA Complementary Sites. Developmental Biology 258, 432-442.

Murashige, T., and Skoog, F. (1962) A Revised Medium for Rapid Growth and Bioassays with Tobacco Tissue Cultures. Physiology Plant. 15, 473-497

Nakaminami, K., Karlson, D. T., and Imai, R (2006) Functional Conservation of Cold Shock Domains in Bacteria and Higher Plants. Proceedings of the National Academy of Sciences 103, 10122–10127

Nakaminami, K., Sasaki, K., Kajita, S., Takeda, H., Karlson, D., Ohgi, K., and Imai, R. (2005) Heat Stable ssDNA/RNA-binding Activity of a Wheat Cold Shock Domain Protein. FEBS Letters 579, 4887–4891

Nakamura, M., Katsumata, H., Abe, M., Yabe, N., Komeda, Y., Yamamoto, K., and Takahashi, T. (2006). Characterization of the Class IV Homeodomain-Leucine Zipper Gene Family in Arabidopsis. Plant Physiology 141, 1363–1375

Mansfield, G., and Briarty, L. (1996) The Dynamics of Seedling and Cotyledon Cell Development in Arabidopsis thaliana During Reserve Mobilization. International Journal of Plant Sciences 157, 280-295

Ng, M. and Yanofsky, M. (2001). Activation of the Arabidopsis B Class Homeotic Genes by APETALA1. The Plant Cell 13, 739–753.

Ohta, M., Matsui, K., Hiratsu, K., Shinshi, H., and Ohme-Takagi. M. (2001) Repression Domains of Class II ERF Transcriptional Repressors Share an Essential Motif for Active Repression. The Plant Cell 13, 1959–1968

Novillo, F., Alonso, J., Ecker, J., and Salinas J. (2004) CBF2/DREB1C is a Negative Regulator of CBF1/DREB1B and CBF3/DREB1A Expression and Plays a Central Role in Stress Tolerance in Arabidopsis. Proceedings of the National Academy of Sciences 101, 3985-3990

Ogas, J., Kaufmann, S., Henderson, J and Somerville, K. (1999) PICKLE is a CHD3 Chromatin Remodelling Factor that Regulates the Transition from Embryonic to Vegetative Development in Arabidopsis. Proceedings of the National Academy of Sciences, 96 13839–13844

Ogata, Y., Iizuka, M., Nakayama, D., Ikeda, M., Kamada, H., and Koshiba, T. (2005) Possible Involvement of Abscisic Acid in the Induction of Secondary Somatic Embryogenesis on Seed-Coat-Derived Carrot Somatic Embryos. Planta 221, 417-423

Ohashi, Y., Oka, A., Rodrigues-Pousada, R., Possenti, M., Ruberti, I., Morelli, G., and Aoyama T. (2003) Modulation of Phospholipid Signaling by GLABRA2 in Root-Hair Pattern Formation. Science 300, 1427-1430

Ooms, J., Leon-Kloosterziel, K., Bartels, D., Koornneef, M., and Karssen, C. (1993) Acquisition of Desiccation Tolerance and Longevity in Seeds of Arabidopsis thaliana: A Comparative Study Using Abscisic Acid-Insensitive abi3 Mutants. Plant Physiology 102, 1185-1 191

Pandey, R., Muller, A., Napoli, C., Selinger, D., Pikaard, C., Richards, E., Bender, J., Mount, D., and Jorgensen, R. (2002) Analysis of Histone Acetyltransferase and Histone Deacetylase Families of Arabidopsis thaliana Suggests Functional Diversification of Chromatin Modification Among Multicellular Eukaryotes. Nucleic Acids Research 30, 5036-5055

Parcy F, Valon C, Raynal M, Gaubier-Comella P, Delseny M, Giraudat J. (1994) Regulation of Gene Expression Programs During Arabidopsis Seed Development: Roles of the ABI3 locus and of Endogenous Abscisic Acid. The Plant Cell 6, 1567–1582

Parenicová, L., de Folter, S., Kieffer, M., Horner, D., Favalli, C., Busscher, J., Cook, H., Ingram, R., Kater, M., Davies, B., Angenent, 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. The Plant Cell 15 1538–1551

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

Pelaz, S., Ditta, G., Baumann, E., Wisman, E., and Yanofsky, M. (2000) B and C Floral Organ Identity Functions Require SEPALLATA MADS-Box Genes. Nature 405, 200-203

Perry, S., Nichols, K. and Fernandez, D. (1996) The MADS Domain Protein AGL15 Localizes to the Nucleus During Early Stages of Seed Development. The Plant Cell 8, 1977–1989.

Perry, S., Lehti, M. and Fernandez, D. (1999) The MADS-Domain Protein AGAMOUS-Like 15 Accumulates in Embryonic Tissues with Diverse Origins. Plant Physiology 120, 121–130.

Phelps-Durr, T., Thomas, J., Vahab, P., and Timmermans, M. (2005) Maize rough sheath2 and its Arabidopsis Orthologue ASYMMETRIC LEAVES1 Interact with HIRA, a Predicted Histone Chaperone, to Maintain Knox Gene Silencing and Determinacy During Organogenesis. The Plant Cell 17, 2886-2898

Pickett, F. B., and Meeks-Wagner, D. R. (1995) Seeing Double: Appreciating Genetic Redundancy. The Plant Cell 7, 1347-1356

Poethig, R. (1990) Phase Change and the Regulation of Shoot Morphogenesis in Plants. Science 250, 923-930

Poethig, R. (2003) Phase Change and the Regulation of Developmental Timing in Plants. Science 301, 334-336

Polissi, A., De Laurentis, W., Zangrossi, S., Briani, F., Longhi, V., Pesole, G., and Dehò, G. (2003) Changes in Escherichia coli Transcriptome During Acclimatization at Low Temperature. Research in Microbiology 154, 573-580

Portereiko, M., Lloyd, A., Steffen, J., Punwani, J., Otsuga, D., and Drews, G. (2006) AGL80 Is Required for Central Cell and Endosperm Development in Arabidopsis. The Plant Cell 18, 1862–1872.

Preuss, D., Lemieux, B., Yen, G., and Davis, R. (1993) A Conditional Sterile Mutation Eliminates Surface Components from Arabidopsis Pollen and Disrupts Cell Signalling during Fertilization. Genes and Development 1993 7, 974-985

Prochasson, P., Florens, L., Swanson, S., Washburn, M., and Workman J. (2005) The HIR Corepressor Complex binds to Nucleosomes Generating a Distinct Protein/DNA Complex Resistant to Remodelling by SWI/SNF. Genes and Development 19, 2534-539.

Proudfoot, N., Furger, A., and Dye, M. (2002) Integrating mRNA Processing with Transcription. Cell 108, 501–512

Rambaldi, I., Kovàcs, E., and Featherstone, M. (1994) A Proline-Rich Transcriptional Activation Domain in Murine HOXD-4 (HOX-4.2). Nucleic Acids Research 11, 376-382.

Ranjan, M., Tafuri, S., and Wolffe A (1993) Masking mRNA From Translation in Somatic Cells. Genes and Development 7, 1725-1736

Raz, V., Bergervoet, J., and Koornneef, M. (2001) Sequential Steps for Developmental Arrest in Arabidopsis Seeds. Development 128, 243-252

Reddy, M., Ghabrial, S., Redmond, C., Dinkins, R., and Collins, G. (2001) Resistance to Bean Pod Mottle Virus in Transgenic Soybean Lines Expressing the Capsid Polyprotein. Phytopathology 91, 831-838.

Reyes, J. (2006) Chromatin Modifiers that Control Plant Development. Current Opinion in Plant Biology 9, 21-27

Reyes, J., Hennig, L., and Gruissem, W. (2002). Chromatin-Remodelling and Memory Factors. New Regulators of Plant Development. Plant Physiology 130, 1090–1101

Phizicky, E., and Fields, S. (1995) Protein-Protein Interactions: Methods for Detection and Analysis. Microbiological Reviews 59, 49-123

Reichmann, J., Krizek, B., and Meyerowitz, E. (1996) Dimerization Specificity of Arabidopsis MADS Domain Homeotic Proteins APETALA1, APETALA3, PISTILLATA, and AGAMOUS. Proceedings of the National Academy of Sciences 93, 4793–4798.

Reichmann, J. and Meyerowitz, E. (1997) MADS Domain Proteins in Plant Development. Biological Chemistry 378, 1079–1101

Riezman, H. (2004) Why Do Cells Require Heat Shock Proteins to Survive Heat Stress?. Cell Cycle 3, 61-63.

Rijpkema, A. S., Gerats, T., and Vandenbussche, M. (2007) Evolutionary Complexity of MADS Complexes. Current Opinion in Plant Biology 10, 32–38

Pina, C., Pinto, F., Feijó, J., and Becker, J. (2005) Gene Family Analysis of the Arabidopsis Pollen Transcriptome Reveals Biological Implications for Cell Growth, Division Control, and Gene Expression Regulation. Plant Physiology 138, 744–756

Rounsley, S., Ditta, G. and Yanofsky, M. (1995) Diverse Roles for MADS Box Genes in Arabidopsis Development. The Plant Cell 7,1259–1269

Sakhanokho, H., Ozias-Akins, P., May, O., and Chee, P. (2004) Induction of Somatic Embryogenesis and Plant Regeneration in Select Georgia and Pee Dee Cotton Lines. Crop Science 44, 2199–2205

Satoh, R., Fujita, Y., Nakashima, K., Shinozaki, K., and Yamaguchi- Shinozaki, K. (2004) A Novel Subgroup of bZIP Proteins Functions as Transcriptional Activators in Hypoosmolarity-Responsive Expression of the ProDH Gene in Arabidopsis. Plant Cell Physiology 45, 309–317

Savidge B, Rounsley S., and Yanofsky M. (1995) Temporal Relationship Between the Transcription of two Arabidopsis MADS Box Genes and the Floral Organ Identity Genes. The Plant Cell 7, 721-733.

Schmidt, E., Guzzo, F., Toonen, M., and de Vries S. (1997) A Leucine-Rich Repeat Containing Receptor-Like Kinase Marks Somatic Plant Cells Competent to form Embryos. Development 124, 2049-2062

Schrick, K., Nguyen, D., Karlowski, W., and Mayer, K. (2004) START Lipid/Sterol-Binding Domains are Amplified in Plants and are Predominantly Associated with Homeodomain Transcription Factors. Genome Biology, 5

Schubert, D., Primavesi, L., Bishopp, A., Roberts, G., Doonan, J., Jenuwein, T., and Goodrich, J. (2006). Silencing by Plant Polycomb-Group Genes Requires Dispersed Trimethylation of Histone H3 at Lysine 27. The EMBO Journal 25, 4638–4649

Shiota, H., Satoh, R., Watabe, K., Harada, H., and Kamada, H. (1998) C-ABI3, the Carrot Homologue of the Arabidopsis ABI3, is Expressed During both Zygotic and Somatic Embryogenesis and Functions in the Regulation of Embryo-Specific ABA-Inducible Genes. Plant and Cell Physiology 39, 1184-1193

Semiarti, E., Ueno, Y., Tsukaya, H., Iwakawa, H., Machida, C., and Machida, Y. (2001) The ASYMMETRIC LEAVES2 Gene of Arabidopsis thaliana Regulates Formation of a Symmetric Lamina, Establishment of Venation and Repression of Meristem-Related Homeobox Genes in Leaves. Development 128, 1771-1783

Serebriiskii, I., Estojak, J., Berman, B. and Golemis. E. (2000) Approaches to Detecting False Positives in Yeast Two-Hybrid Systems. BioTechniques 28, 328-336

Sharma, V., Carles, C., and Fletcher, J. (2003) Maintenance of Stem Cell Populations in Plants. Proceedings of the National Academy of Sciences 100, 11823–11829

Sharma, P., Sharma, N., and Deswal, R. (2005) The Molecular Biology of the Llow-Temperature Response in Plants. BioEssays 27, 1048–1059

Shchennikova, A., Shulga, O., Immink, R., Skryabin, K., Angenent, G. (2004) Identification and Characterization of Four Chrysanthemum MADS-box Genes, Belonging to the APETALA1/FRUITFULL and SEPALLATA3 Subfamilies. Plant Physiology 134, 1632-41

Shen, B., Sinkevicius, K., Selinger, D., and Tarczynsk, M. (2006) The Homeobox Gene GLABRA2 Affects Seed Oil Content in Arabidopsis. Plant Molecular Biology 60, 377–387

Shuai, B., Reynaga-Pena, C., and Springer, P. (2002) The LATERAL ORGAN BOUNDARIES Gene Defines a Novel, Plant-Specific Gene Family. Plant Physiology, 129, 747–761

Silverstein, R., and Ekwall, K. (2005) Sin3: a Flexible Regulator of Global Gene Expression and Genome Stability. Current Genetics 47, 1-17.

Singh, N., Zhu, W., and Hanes, S. (2005) Sap18 is Required for the Maternal Gene Bicoid to Direct Anterior Patterning in Drosophila melanogaster. Developmental Biology 278, 242–254

Skoog, F., and Miller, C. (1957) Chemical Regulation of Growth and Organ Formation in Plant tissues Cultured In Vitro. Symposia of the Society for Experimental Biology 54, 118-130

Sommerville, J (1999) Activities of Cold-Shock Domain Proteins in Translation Control. Bioessays 21, 319-325

Song, C, Agarwal, M., Ohta, M., Guo, Y., Halfter, U., Wang P., and Zhu J. (2005) Role of an Arabidopsis AP2/EREBP-Type Transcriptional Repressor in Abscisic Acid and Drought Stress Responses. The Plant Cell 17, 2384–2396

Song C., and Galbraith, D.W. (2006) AtSAP18, An Orthologue of Human SAP18, is Involved in the Regulation of Salt Stress and Mediates Transcriptional Repression in Arabidopsis. Plant Molecular Biology 60, 241–257

Struhl, K. (1998) Histone Acetylation and Transcriptional Regulatory Mechanisms. Genes and Development 12, 599-606

Souter, M., and Lindsay, K. (2000) Polarity and Signaling in Plant Embryogenesis. Journal of Experimental Botany 51, 971-983

Sridhar, V.V., Surendrarao, A., and Liu, Z. (2006) APETALA1 and SEPALLATA3 Interact With SEUSS to Mediate Transcription Repression During Flower Development. Development 133, 3159-3166

Stone, S., Kwong, L., Yee, K., Pelletier, J., Lepiniec, L., Fischer, R., Goldberg, R., and Harada, J. (2001) LEAFY COTYLEDON2 Encodes a B3 Domain Transcription Factor that Induces Embryo Development. Proceedings of the National Academy of Science 98, 11806-11811

Suzuki, M., Wang, H., and Mcarty, D. (2007) Repression of the LEAFY COTYLEDON 1/B3 Regulatory Network in Plant Embryo Development by VP1/ABSCISIC ACID INSENSITIVE 3-LIKE B3 genes. Plant Physiology 143, 902-911

Sunilkumar, G., and Rathore. K. (2001) Transgenic Cotton: Factors Influencing Agrobacterium-Mediated Transformation and Regeneration. Molecular Breeding 8, 37–52.

Tai, H., Tai, G., and Beardmore, T. (2005) Dynamic Histone Acetylation of Late Embryonic Genes During Seed Germination. Plant Molecular Biology 59, 909–925

Takada, S., and Jürgens, G. (2007) Transcriptional Regulation of Epidermal Cell Fate in the Arabidopsis Embryo. Development 134, 1141-1150

Tang, W. and Perry, S. (2003) Binding Site Selection for the Plant MADS Domain Protein AGL15: An In Vitro and In Vivo Study. Journal of Biological Chemistry 278, 28154–28159.

Tange, T., Shibuya, T., Jurica, M., and Moore, M. (2006) Biochemical Analysis of the EJC Reveals Two New Factors and a Stable Tetrameric Protein Core. RNA 11, 1869-1883

Traven, A., Jelicic, B., and Sopta, M. (2006) Yeast Gal4: a Transcriptional Paradigm Revisited. EMBO Reports 7, 496-499

The Arabidopsis Initiative (2000) Nature 408, 796-815

Tian, L., Wang, J., Fong, M., Chen, M., Cao, H., Gelvin, S., and Chen, Z. (2003) Genetic Control of Developmental Changes Induced by Disruption of Arabidopsis Histone Deacetylase 1 (AtHD1) Expression. Genetics 165, 399-409

Tian, L., and Chen, Z. (2001) Blocking Histone Deacetylation in Arabidopsis Induces Pleiotropic Effects on Plant Gene Regulation and Development. Proceedings in the National Academy of Sciences 98, 200-205

Titz, B., Thomas, S., Rajagopala S., Chiba, T., Ito, T., and Uetz, P. (2006) Transcriptional Activators in Yeast. Nucleic Acids Research 34, 955-967

Tiwari, .S.B, Hagen, G., and Guilfoyle, T.J. (2001) Aux/IAA Proteins Contain a Potent Transcriptional Repression Domain. The Plant Cell 16, 533–543.

Toby, G. ,and Golemis, E. (2001) Using the Yeast Interaction Trap and Other Two-Hybrid- Based Approaches to Study Protein–Protein Interactions. Methods 24, 201–217

Trolinder, N., and Xhixian, C. (1989) Genotype Specificity of the Somatic Embryogenesis Response in Cotton. Plant Cell Reports 8,133–136

Toufighi, K., Brady, S., Austin, R., Ly, E., and Provart, N. (2005) The Botany Array Resource: e-Northerns, Expression Angling, and promoter analyses. The plant Journal 43, 153-63

Tzafrir, I., Dickerman, A., Brazhnik, O., Nguyen, Q., McElver, J., Frye, C., Patton, D., and Meinke, D. (2003) The Arabidopsis SeedGenes Project. Nucleic Acids Research 31, 90-93

Tzafrir, I., Pena-Muralla, R., Dickerman, A., Berg, M., Rogers, R., Hutchens, S., Sweeney, T., McElver, J., Aux, G., Patton, D., and Meinke, D. (2004) Identification of Genes Required for Embryo Development in Arabidopsis. Plant Physiology 135, 1206–1220

Uhrig, J. (2006) Protein Interaction Networks in Plants. Planta 224, 771–781

Verelst, W., Saedler, H., and Munster, T. (2007) MIKC* MADS-Protein Complexes Bind Motifs Enriched in the Proximal Region of Late Pollen-Specific Arabidopsis Promoters. Plant Physiology 143, 447–460.

Vasil, I. (1988) Progress in the Regeneration and Genetic Manipulation of Cereal Crops. Nature Biotechnology 6, 397-402

Vernon, D., and Meinke, D. (1994) Embryogenic Transformation of the Suspensor in twin, a Polyembryonic Mutant of Arabidopsis. Developmental Biology 165, 566-573

Vernon, D., Hannon, M., Le, M., and Forsthoefel, N. (2001) An Expanded Role for the TWN1 Gene in Embryogenesis: Defects in Cotyledon Pattern and Morphology in the TWN1 Mutant of Arabidopsis (Brassicaceae). The American Journal of Botany 88, 570-582

Vierstra, R. (2003) The Ubiquitin/26S Proteasome Pathway, the Complex Last Chapter in the Life of Many Plant Proteins. Trends in Plant Sciences 8, 135-142

Vogel, G. (2005) How Does a Single Somatic Cell Become a Whole Plant? Science 309, 86

Vogel, J., Zarka, D., Van Buskirk, H., Fowler, S., and Thomashow, M. (2005) Roles of the CBF2 and ZAT12 Transcription Factors in Configuring the Low Temperature Transcriptome of Arabidopsis. The Plant Journal 41, 195–211

Wang, X., and Grumet, R. (2004) Identification and Characterization of Proteins that Interact with the Carboxy Terminus of Poly(A)-Binding Protein and Inhibit Translation In Vitro. Plant Molecular Biology 54, 85–98

Wang, H., Hill, K., and Perry, S. (2004) An Arabidopsis RNA Lariat Debranching Enzyme is Essential for Embryogenesis. Journal of Biological Chemistry 279, 1468-1473

Wang, H., Tang, W., Zhu, C. and Perry, S. (2002) A Chromatin Immunoprecipitation (ChIP) Approach to Isolate Genes Regulated by AGL15, a MADS Domain Protein that Preferentially Accumulates in Embryos. The Plant Journal 32, 831–843.

Wang, N., Yamanaka, K., and Inouye, M. (2000) Acquisition of Double-Stranded DNA-Binding Ability in a Hybrid Protein Between Escherichia coli CSPA and the Cold Shock Domain of Human YB-1. Molecular Microbiology 38, 526-534

Wang, H., Caruso, L., Downie, A., Perry, S. (2004) The Embryo MADS Domain Protein AGAMOUS-Like 15 Directly Regulates Expression of a Gene Encoding an Enzyme Involved in Gibberellin Metabolism. The Plant Cell 16, 1206-1219

West, M., and Harada, J. (1993) Embryogenesis in Higher Plants: an 0verview. Plant Cell 5, 1361-1369

West, M. , Yee, K., Danao, J., Zimmerman, J., Fischer, R., Goldberg, R., and Harada, J. (1994) LEAFY COTYLEDON1 is an Essential Regulator of Late Embryogenesis and Cotyledon Identity in Arabidopsis. The Plant Cell 6, 1731-1745

Wilkins, T., Rajasekaran, K., and Anderson, D. (2000) Cotton Biotechnology. Critical Reviews in Plant Sciences 19, 511–550

Willemsen, V. and Scheres, B. (2004) Mechanisms of Pattern Formation in Plant Embryogenesis. Annual Review of Genetics 38, 587–614

Wilson, Z., and Yang, C. (2004) Plant Gametogenesis: Conservation and Contrasts in Development. Development 128, 483-492

Wu, K., Malik, K., Tian, L., Brown, B., and Miki, B. (2000) Functional Analysis of a RPD3 Histone Deacetylase Homologue in Arabidopsis thaliana. Plant Molecular Biology 44, 167–176

Xu, J., Zhang, H-Y., Xie, C-H., Xue, H-W., Dijkhuis, P., and Liu. C-H. (2005) EMBRYONIC FACTOR 1 Encodes an AMP Deaminase and is Essential for the Zygote to Embryo Transition in Arabidopsis. The Plant Journal 42, 743–756

Yalovsky, S., Rodriguez-Concepcion, M., Bracha, K., Toledo-Ortiz, G., and Gruissem, W. (2000) Prenylation of the Floral Transcription Factor APETALA1 Modulates its Function. The Plant Cell, 12, 1257-66

Yamanaka, K., Fang, L. and Inouye, M. (1998) The CspA Family in Escherichia coli: Multiple Gene Duplication for Stress Adaptation. Molecular Microbiology 27, 247-255.

Yang, Y., Fanning, L., and Jack, T. (2003a) The K domain Mediates Heterodimerization of the Arabidopsis Floral Organ Identity Proteins, APETALA3 and PISTILLATA. The Plant Journal 33, 47–59

Yang Y, Xiang H, and Jack T (2003b) pistillata-5, An Arabidopsis B Class Mutant With Strong Defects in Petal But Not in Stamen Development. The Plant Journal 33, 177-188

Yang, Y., and Jack, T. (2004) Defining Subdomains of the K Domain Important for Protein-Protein Interactions of Plant MADS Proteins. Plant Molecluar Biology 55, 45-59

Yang, W-C., Ye, D., Xu, J and Sundaresan, V. (1999) The SPOROCYTELESS Gene of Arabidopsis is Required for Initiation of Sporogenesis and Encodes a Novel Nuclear Protein. Genes and Development 13, 2108-2117

Yang, W., Inouye, C., Zeng, Y., Bearss, D., and Seto, E. (1996) Transcriptional Repression by YY1 is Mediated by Interaction with a Mammalian Homolog of the Yeast Global Regulator RPD3. Proceedings of the National Academy of Science 93, 12845-12850

Ylstra, B., and McCormick, S. (1999) Analysis of mRNA Stabilities During Pollen Development and in BY2 Cells. The Plant Journal 20, 101-108

Yu, H., Ito, Y., Zhao, Y., Peng, J., Kumar, P., and Meyerowitz, M. (2004) Floral Homeotic Genes are Targets of Gibberellin Signaling in Flower Development. Proceedings of the National Academy of Sciences 101, 7827–7832

Zhang, Y., Iratni, R., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (1997) Histone Deacetylases and SAP18, a Novel Polypeptide, are Components of a Human Sin3 Complex. Cell 89, 357-64

Zheng, L., Roeder, R., and Luo, Y (2003) S Phase Activation of the Histone H2B Promoter by OCA-S, a Coactivator Complex that Contains GAPDH as a Key Component. Cell 114, 255–266

Zhou, C., Zhang, L.., Duan, B., Brian, M., and Wu, K. (2005) HISTONE DEACETYLASE19 Is Involved in Jasmonic Acid and Ethylene Signaling of Pathogen Response in Arabidopsis. The Plant Cell 17, 1196-1204

Zhu, W., Foehr, M., Jaynes, J., and Hanes, S. (2001) Drosophila SAP18, a Member of the SIN3/RPD3 Histone Deacetylase Complex, Interacts With Bicoid and Inhibits its Activity. Development Genes and Evolution 211, 109-17

Zhu, W., and Hanes, S. (2000) Identification of Drosophila Bicoid-Interacting Proteins Using a Custom Two-Hybrid Selection. Gene 245, 329-39

Zhu, C. and Perry, S. (2005) Control of Expression and Autoregulation of AGL15, a Member of the MADS-Box Family. The Plant Journal 41, 583–594

Zimmermann P, Hirsch-Hoffmann M, Hennig L, Gruissem W, (2004) GENEVESTIGATOR. Arabidopsis Microarray Database and Analysis Toolbox. Plant Physiology 136, 2621-2632.

Zuo, J., Niu, Q., Frugis, G., Chua, N. (2002) The WUSCHEL Gene Promotes Vegetative-to-Embryonic Transition in Arabidopsis. The Plant Journal 30, 349-359

Vita
Kristine Hill

Year of birth: 1973

Place of birth: Bromsgrove, United Kingdom

Educational institutions:

- B.Sc Biological sciences (Honours), 2000
- University of Plymouth, Plymouth, U.K. (1997-2000)
- University of Kentucky, Kentucky, U.S (exchange student, 1998-1999)

Professional positions held:

- 2001-2007 Research Assistant, University of Kentucky, U.S
- 1998-1999 Laboratory Assistant, University of Kentucky, U.S

Scholastic and professional honors:

- 2007, American Society of Plant Biologists (ASPB) Travel Grant Award
- 2006, 17th International Conference on Arabidopsis Research, Graduate Student Recognition Award
- 1999, Howard Hughes Medical Institute Undergraduate Summer Research Scholarship
- 1998 Selected for participation in the International Student Exchange Programme (ISEP)

Professional publications:

K. Hill, H. Wang, & S.E. Perry (2007) A Transcriptional Repression Motif in the MADS-factor AGL15 is Involved in Recruitment of Histone Deactylase Complex Components. The Plant Journal (In press)

H. Wang, **K. Hill**, & S.E. Perry (2004)
An *Arabidopsis* RNA Lariat Debranching Enzyme Is Essential for Embryogenesis
The Journal of Biological Chemistry **279** 1468–1473

D. Thakare, W. Tang, W., **K. Hill**, & S.E. Perry (2007) Isolation and characterization of a full-length soybean (*Glycine max*) embryo MADS box gene. (Manuscript in preparation)

Signature_____(Kristine Hill)