Actin genes in *Arabidopsis thaliana* exist as an ancient gene family. *A. thaliana* actins comprise two ancient classes, reproductive and vegetative. *ACT1* and *ACT3*, the result of a duplication event 30-60 MYA, comprise one of the five ancient subclasses and both are expressed at high levels in mature pollen, embryo sac, and organ primordia. *ACT1* and *ACT3* share little homology with exception of a 55 base pair sequence within the 5' flanking region. A fluorescence assay measuring reporter gene expression was designed to quantify changes in expression in various promoter/GUS fusion constructs. Truncations of the 5' upstream flanking region, within an adjacent intron, and site-directed mutagenesis constructs demonstrate that all these regions are needed to promote high levels of mature pollen expression. The first intron in *ACT1* is required for high levels of pollen expression and organ primordia expression. This intron functions in a gene-specific manner as substitution of the first intron of *ACT1* with the first intron of a vegetatively expressed actin efficiently down-regulates pollen expression. Multiple sequence elements within the *ACT1* 5' flanking region direct high levels of mature pollen specific and organ primordia expression.

Ancient classes of actin could have been preserved because their expression was under strong selective constraint. *Cis*-elements and transcription factors may function in
distant plant species. 5' flanking regions of the ACT1 gene translationally fused to a reporter gene were transformed into distantly related *Nicotiana tabacum* and *Oryza sativa*. ACT1 was expressed almost identically in *N. tabacum* and *O. sativa* as in *A. thaliana*. The heterologous transformation results clearly show that the expression patterns of the reproductive classes of actin genes are highly conserved in heterologous species in reproductive organs, but not vegetative organs. Both traditional methods of promoter dissection and examination of gene expression within a heterologous system are useful in analyzing duplicated regulatory regions throughout the evolution of the actin gene family.

THE EVOLUTION OF *ARABIDOPSIS THALIANA* ACTIN GENE REPRODUCTIVE SUBCLASS EXPRESSION PATTERN

by

ANGELA VICTORIA VITALE

B. S., Berry College, 1993

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2001
© 2001

Angela Victoria Vitale

All Rights Reserved
THE EVOLUTION OF *ARABIDOPSIS THALIANA* ACTIN GENE REPRODUCTIVE SUBCLASS EXPRESSION PATTERN

by

ANGELA VICTORIA VITALE

Approved: Richard B. Meagher

Major Professor: Richard B. Meagher

Committee: Katherine Spindler
John McDonald
Michael Bender
Ron Nagao

Electronic Version Approved:

Gordhan L. Patel
Dean of the Graduate School
The University of Georgia
August 2001
DEDICATION

To Kurt for more than I can ever say. And to Sparky and Lilith for teaching me about the really important things: life and living it happily. Thanks to my family, especially Dad and Christopher, who in the end would not let me give up.
ACKNOWLEDGEMENTS

Thanks to everyone who both literally and figuratively stopped me from leaping from a very high place. Starting with my fourth grade teacher, Mrs. Guest and continuing with my brilliant and dedicated high school instructors, Carolyn Becker, Beth Caccioli, Billy Jones, Lori Ziecker and William Eudaly – I thank you for catching me when I often stumbled. My good friends, Lee Biola, Briana Barnhill, and Chandra Weils, I love you all very much. From my early college years, I owe an immense debt of gratitude to John Smith, theatre director extrordinare. Within the Biology Department, I thank Steve Schwartz, John Graham, and Robert Martin, the latter both a mentor and close friend. In graduate school, I thank everyone who knew I had my dark days but took care of me anyway. Thank you Adriana Kajon, Becky Balish, Laura Gilliland, Yolanda Lay, Anne Marie, Domagoi Vucic, Maria Sanchez, David Powell, Gina, Natalia Roques, Lucia Polowski, Hyla Guigino, Robert Shapiro and Kathy Spindler. You all held on to me and without your attention I would not have made it! Even as I transition into my new life here in Boston I have to express gratitude to Kathleen Riley and Patrice Tellier for their concern and friendship. Last but not least, thank you Ira Herman, for helping make this new life a reality.
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>........................................................................................................</td>
<td>v</td>
</tr>
<tr>
<td>CHAPTER 1</td>
<td>ANTECEDENTS AND LITERATURE REVIEW</td>
<td>1</td>
</tr>
<tr>
<td>CHAPTER 2</td>
<td>THE 5’ FLANKING REGION OF \textit{ACT1 ARABIDOPSIS} ACTIN GENE CONTAINS MULTIPLE ELEMENTS NECESSARY FOR HIGH-LEVEL, MATURE POLLEN EXPRESSION</td>
<td>31</td>
</tr>
<tr>
<td>CHAPTER 3</td>
<td>REPRODUCTIVE CLASSES OF \textit{ARABIDOPSIS THALIANA} ACTIN GENES GUIDE HOMOLOGOUS PATTERNS OF EXPRESSION IN EVOLUTIONARILY DISTANT PLANT SPECIES</td>
<td>65</td>
</tr>
<tr>
<td>CHAPTER 4</td>
<td>CONCLUSIONS AND FUTURE DIRECTIONS</td>
<td>92</td>
</tr>
</tbody>
</table>
CHAPTER 1

ANTECEDENTS AND LITERATURE REVIEW
I. Actin is a ubiquitous eukaryotic cytoskeletal protein.

Actin is a cytoskeletal protein that is found throughout the eukaryotic kingdom. Although my work focuses on the expression and evolution of reproductive plant actin genes, it is imperative to review the diverse roles actin plays in the cytoskeleton and the many proteins which interact with actin. Thus, this review will examine not only cytoplasmic plant and animal actins, but also actin’s roles and associated proteins in animal muscle cells.

Actin in a monomeric state (G-actin) is a 43kd protein and is found throughout the cytoplasm and within the nucleus (Nakayasu and Ueda, 1983). Monomeric actin is also defined by its ability to hydrolyze ATP to ADP and to bind divalent cations \( \text{Mg}^{2+} \) and \( \text{Ca}^{2+} \) within the cytoplasm (Pollard and Cooper, 1986). Though found only as a monomeric polypeptide in the nucleus, actin is known primarily for playing a dynamic role in the cytoskeleton through its ability to form microfilaments (F-actin) (Pollard, 1990). Though actin is a ubiquitous protein, it has both muscle and cytoplasmic isoforms that behave and are sequestered very differently within the cytoplasm (Herman, 1993; Shuster and Herman, 1998).

G- and F-actin have been examined using X-ray crystallographic techniques that have resolved their structures at the atomic level (Holmes et al., 1990; Kabsch et al., 1990). G-actin typically is 375-377 amino acids in length and forms a squarish dimeric structure. The amount of free monomer actin in cells is always much greater than that needed to spontaneously form filaments. Much is known about muscle actin isoforms
and their ability to generate concerted force (Geeves and Holmes, 1999; Littlefield and Fowler, 1998).

Actin filament elongation is generally considered a four step process (Shuster and Herman, 1998). First the monomer binds with a divalent cation. The divalent cation binding site is located between two halves of the protein in a hinge-like structure. Next, follows a rate-limiting step whereby these activated monomers nucleate in groups of three. After this nucleation elongation of the filament is rapid with monomers favorably added to the barbed end (+) over the pointed end (-) (so named because of the appearance of arrowhead structures formed upon decoration of the filament with myosin). Finally filaments are able to join together in a barbed end (+) to pointed end (-) manner, forming longer filaments. Once in a filamentous structure, actin will eventually hydrolyze bound ATP to ADP and the formation of ADP will destabilize the filament (Pollard et al., 2000).

The behavior of F-actin structures in vertebrate non-muscle cells was recently reviewed (Shuster and Herman, 1998). Although animal, protozoa and fungal cells behave differently from plant cells, much knowledge useful to the plant actin field has been accumulated. Animal non-muscle cells move by ruffling their membrane edges. Many models exist to explain the ability of actin to generate a motive force. Through the use of various actin binding, capping, and severing proteins discussed below, microfilaments cooperatively act to move animal cells. These microfilaments form
vast numbers, making three dimensional arrays in conjunction with many specific actin
binding proteins.

The Arp2/3 protein, touted as the most important recent discovery in the
regulation of actin filaments in nonmuscle cells, binds the F-actin pointed (-) end and
allows rapid extension of the barbed (+) end to occur (Ma et al., 1998; Welch et al., 1997;
Winter et al., 1997). Many external cell structures that depend on specific cell
morphology, such as psuedopods or membrane ruffles, are required to guide cell
locomotion. Eukaryotic phyla almost universally display common networks of filaments
below the cell’s leading edge. This highly organized cross-linking of F-actin is facilitated
by Arp2/3. Arp2/3 is a seven subunit complex, the products of two genes, and is
conserved across a broad spectrum of organisms. Arp2/3 creates a 70° branching pattern
from actin filaments by binding to the end of one filament and to the side of another
(Pollard et al., 2000). These branching networks facilitate motion by allowing nucleation
of growing filaments to occur at the edges of membranes. The Arp2/3 protein also works
in association with actin depolymerizing factor, which assists in the recycling of G actin
back onto growing filament ends.

The actin depolymerizing factor (ADF)/cofilin family is a ubiquitous actin
binding protein family that is possibly one of the best studied for its role in regulating the
actin cytoskeleton (Cooper and Schafer, 2000; Wear et al., 2000). Found in birds,
mammals, plants and insects, ADF/cofilin protein sizes vary from 118 to 168 amino acids
and 13 to 19 kD. ADF and coflin, which are different gene products, behave in a similar fashion. From in vitro studies both proteins bind actin subunits in a 1:1 ratio and will increase the pool of G-actin monomer at biological pH (Bamburg, 1999). All ADF/cofilins thus far are composed of four $\alpha$-helices and five $\beta$-sheets. In most eukaryotic phyla the penultimate residue is acetylated and/or Ser3 or the equivalent Ser4 is phosphorylated. The crystal structure determined for yeast coflin has laid the groundwork for comparison with other actin binding proteins (Bamburg, 1999). For instance, ADF/cofilin shares a domain that is homologous to repeated domains within both the gelsolin and villin families (Bamburg, 1999). Based on mutagenesis work, the actin binding domain of ADF/cofilin is predicted to be at the N-terminus of the protein (Bamburg, 1999).

Using electron cryomicroscopy and image reconstruction a crude structure of ADF/cofilin bound to F-actin has been deduced. Basically, ADF/cofilin lies on the top and bottom of the filament at the pointed end. ADF and coflin both bind monomeric actin as well and can sever actin filaments, though ADF seems to depolymerize actin much more quickly (Bamburg, 1999). ADF also binds F-actin cooperatively and more significantly ADF/cofilin will more readily bind F-actin-ADP than either F-actin-ATP or F-actin-ADP+Pi. When ADF binds F-actin-ADP a slight conformational change in the filament is induced. This twist may help ADF/cofilin to depolymerize actin, though when bound to ADF/cofilin nucleotide exchange of G-actin cannot occur. The ADF/cofilin-G-
actin monomer somehow dissociates and the high concentration of free cofilin will promote nucleotide exchange, thus allowing the monomers to be recycled back onto the barbed end of a filament (Bamburg, 1999; Pollard et al., 2000).

Specific isoforms of ADF/cofilin are seen in multicellular organisms (Bamburg, 1999). These isoforms arose in a variety of ways, from alternative splicing, gene duplication to post-translational modification schemes. Like actin in vertebrate cells, there exist both muscle and non-muscle isoforms. ADF/cofilin is regulated by phosphatidylinositol 4-phosphate (PIP) and 4,5-bis-phosphate (PIP2)(Wear et al., 2000).

Knockout mutants of ADF/cofilin in *Saccharomyces cerevisiae* and *Caenohabditis elegans* are either lethal or produce paralysis in the case of *C. elegans*(Ono et al., 1999). Site-directed mutants in yeast resulted in lowered turnover of actin monomers (Carlier, 1998). *Drosophila melanogaster*, with a defective protein, and *Xenopus laevis*, when injected with an inhibitory antibody, both fail to undergo normal cytokinesis(Bamburg, 1999).

Gelsolin is another actin associated protein responsible for binding G-actin monomers as well as capping barbed filament ends and cleaving filaments (McLaughlin et al., 1993). Gelsolin is composed of six 16 kd subunits, and its ability to bind actin rests on Ca\(^{2+}\) levels and cellular pH. Also, like other actin-associated cytoplasmic proteins, phosphoinosotides can down-regulate the activity of gelsolin by removing the protein from filaments (Cooper and Schafer, 2000). Interestingly when a gelsolin knockout was
observed in cell culture, loss of F-actin organization was observed, including loss of stress fibers (Cooper and Schafer, 2000). However in live mice with the gelsolin gene knocked out, more stress fibers were seen (Cooper and Schafer, 2000). Recently, gelsolin has been demonstrated to control neural differentiation. In a recent study it was shown that overexpression of wildtype gelsolin in certain neural cells leads to failure of the cells to develop lamellepodia, ruffling edges and the actin cable network associated with these morphologies (Westberg et al., 1999). These findings, though in vitro studies, suggest a reason for the connection between a point mutation in a human gelsolin gene and the dominantly inherited disease familial amyloidosis of the Finnish type (FAF). All studies of gelsolin thus far suggest it regulates the actin cytoskeleton in a very complex manner (Cooper and Schafer, 2000).

Working in conjunction with ADF and gelsolin, profilin is a well characterized actin monomer binding protein that binds actin in a 1:1 ratio (Pollard and Cooper, 1986) and blocks filament formation by preventing nucleation (Pollard, 1984). Profilin has been intimately associated with signaling cascades since when it is unbound to actin it binds the membrane-bound PIP2. Phosphorylated phospholipase C hydrolyzes the phospholipid, releasing profilin into the cytoplasm where it then can sequester actin monomers (Goldschmidt-Clermont et al., 1991). Formation of actin filaments can thus be tied to certain cellular signaling pathways. Yeast cells deficient in profilin grow slowly, are multinucleate and devoid of actin cables, although delocalized thick actin
filaments can be detected (Haarer et al., 1990). Mutations in the *Drosophila* profilin gene disrupt many actin related processes and result in formation of bundled filaments, abnormal mitosis, multinucleate cells and cells that are unable to move to correct positions during organismal development (Verheyen and Cooley, 1994).

Myosin is the most studied actin motor protein. Myosin’s role in muscle contraction is well known. It also acts as a motor protein with cytoplasmic actin, aiding nutrient delivery to cell localities and assisting with the movement of certain organelles (Rodriguez and Cheney, 2000). At least ten forms of myosin are known. Myosin II forms filaments in all cell types. Because it is the conventional myosin in muscle cells, myosin II is perhaps the best studied of all the myosins, but other nonconventional myosins from classes such as I, IV, VI, VII and X also play important roles as actin-based motor proteins for intracellular transport (Sellers, 2000). These other myosins are similar to class II in the head, or motor domain, but differ widely in the tail domains. For instance class I myosin is known to interact with the Arp2/3 proteins and thus is likely involved in nucleation and subsequent elongation of actin at the cell periphery (Lechler et al., 2000). Recently, myosin X was cloned and its possible functions in the cytoplasm characterized (Berg et al., 2000). Myosin X has a tail domain containing three plecstrin homology domains thought to interact with the PIP signal transduction pathway. Class X has also been localized to cell membrane structures such as lamellapodia edges, membrane
ruffles, and actin bundle tips of filopodia. Thus, these unconventional myosins most likely interact with actin in an active, dynamic fashion.

II. Roles of plant actin

F-actin has a number of associations and vital roles in the various important general and specialized functions cytoskeletal proteins provide in plant cells. The more generalized functions include determination of the cell division plane since F-actin is found localized at the preprophase band of Allium cells (Ding et al., 1991). F-actin also partially makes up the phragmoplast, the structure that will develop into the cell wall (Staehelin and Hepler, 1996). If the F-actin localized within the phragmoplast is inhibited either through treatment with cytochalasin B or the application of the actin binding protein profilin, cell walls are slow to form, adopt incorrect division planes, or are significantly weakened (Gallagher and Smith, 1997). Nearly all plant cells rely on cytoplasmic streaming to aid in gas exchange and the distribution of organelles at the cell periphery (Kuroda, 1990). Actin is thought to be intimately connected with the phenomenon of cytoplasmic streaming through its association with many other cytoskeletal proteins.

In animal cells, the actin cytoskeleton has been associated with the localization of mRNA in the cytoplasm (Nasmyth and Jansen, 1997). Actin is also likely to play a role in mRNA localization in plant cells. During early embryogenesis of the algae Fucus, mRNA is preferentially distributed as cell division occurs (Bouget et al., 1996).
Treatment of Fucus with cytochalasin D disrupts this mRNA asymmetry, suggesting that actin plays a fundamental role in guidance of the mRNA distribution. In animal cells, a number of actin binding proteins have been suggested to play an important role in mRNA localization. For example, mutations in the profilin gene of Drosophila disrupt transportation of oskar RNA (Erdelyi et al., 1995). Different dynamic arrays of microfilaments were first reported in cultured cells of alfalfa (Seagull et al., 1987). Three arrays of filaments were obvious during interphase: randomly placed filaments as a network in the cortical cytoplasm, larger bundles near the central vacuole, and a “basket” of microfilaments that surrounded the nucleus. Actin filaments lose association with many structures as cell division occurs, instead becoming associated with the spindle and phragmoplast (Smith, 1999).

Associated with plant cell division, actin microfilaments are also correlated with cell elongation (Seagull et al., 1987) and have been shown to play a functional role in these processes (Thimann et al., 1992). Microfilaments are most obvious in stem and root cells where they form linear structures in a growing cell parallel to the direction of elongation. In stems, actin is involved in maize gravitropism signalling (Collings et al., 1998). As cells continue to elongate, the microfilaments reorient transverse to the axis of elongation, becoming aligned with microtubules. This helps to maintain the shape of the cell against turgor pressure. Root cells too depend on actin to help guide gravitropic responses. Using a newly optimized technique for visualizing actin filaments in maize
columella cells it was shown that actin filaments have distinct characteristics (Collings et al., 2001). Most importantly F-actin was found associated with amyloplasts, the gravity sensing organelle found in a polarized distribution pattern within the root columellar cells.

Organelle movement is related to cytoplasmic streaming, but often seen in response to specific external stimuli. Actin has been shown to be in close association with the movement of the endoplasmic reticulum in *Allium* (Quader et al., 1987) and with golgi bodies in tobacco (Boevink et al., 1998). Most striking is actin’s involvement with chloroplast movement, seen in a variety of plant species (Witzram and Pathasarthy 1985; Menzel and Schliwa 1986) (Kadota and Wada, 1992). It has been shown unambiguously that actin is associated with chloroplasts along major filament bundles which at times traverse the entire length of the cell (Kandasamy and Meagher, 1999). Minor actin filaments form baskets around the chloroplasts which are thought to be used as an anchoring aid to the transverse filament bundles. The major cytoskeletal protein tubulin is not found in alliance with chloroplasts (Kandasamy and Meagher, 1999).

F-actin plays a vital role in pollen germination and tip growth, since application of cytochalasin B will inhibit pollen germination (Mascarenhas and Lafountain, 1972). F-actin becomes organized as filamentous structures surrounding the vegetative nucleus upon hydration. These structures are then reorganized into large cable structures as the pollen undergoes germination and the pollen tube emerges. It has been shown that actin
filaments are associated with directing vital cell wall components to the pollen tip via associated motor proteins (Cai et al., 1997). Other proteins known to interact either directly or indirectly with actin have been shown to play an important role in pollen tip elongation. Various proteins have also been shown to interact with the pollen actin cytoskeleton. One of the most studied is a member of the Rho-GTPase family \textit{Rop1} (Li et al., 1999). Null mutants of \textit{Rop1} will germinate in vivo, but fail to fertilize plant egg cells. In these mutants the actin cytoskeleton is severely disrupted, and it is predicted that \textit{Rop1} acts on actin filaments via plant profilin proteins (Li et al., 1999). Actin is also associated with pollen tube growth, signaling a self incompatibility (SI) response. SI will prevent pollen tubes from growing if the stigma and the germinating pollen tube are too closely related. Pollen arrest was thought to be coordinated with the expression of certain \textit{Rnases} and influenced by hormonal control (Kandasamy et al., 1994). It was recently found that the actin cytoskeleton will function abnormally if pollen growth is arrested during SI, but not if the tube growth is stopped by other means (Geitmann et al., 2000).

Thus there is a connection between SI and hormonal control.

\textbf{III. Plant actin gene family structure and evolution}

Plant actin genes have a typical gene structure consisting of a putative TFIID binding site (TATA-box), multiple start sites of transcription, a leader mRNA region split by a large leader intron (150 - 500 base pairs) and three small introns (100 base pairs) between codon positions 20 and 21, at codon position 151, and between positions 355
and 356 (Hightower and Meagher, 1985). This gene structure is found in all known plant actins with the exception of *A. thaliana ACT2* which has a deletion of the first intron (An et al., 1996b) and some actin pseudo genes (Drouin and Dover, 1990),

All land plants have multiple copies of actin genes. *Nicotiana tabacum* has approximately 25 genes (Thangavelu et al., 1993) and *Oryza sativa* approximately eight (McElroy et al., 1990), while some plants such as *Petunia hybrida* have over 100 copies of actin in their genome (Baird and Meagher, 1987). An interesting question to pose is why are there so many actin genes? That is, is each of these gene copies responsible for either necessary protein functions and/or differential regulation of actin in plants? Plant actins have evolved from a single copy gene that originally diverged sometime before the evolution of land plants (Hightower and Meagher, 1986). Thus actin genes between monocots and dicots are often more closely related than intragenic copies, demonstrating that actin is truly an ancient gene family.

Recently, the actin gene family in *A. thaliana* was isolated and the evolutionary relationships of the genes elucidated (McDowell et al., 1996b). Within *A. thaliana* there are ten actin genes, eight of which are expressed actively. These eight genes can be divided into the vegetative and reproductive classes of genes, and then further subdivided into five subclasses. Each class has a corresponding expression pattern representative of all members of that class. For instance in the reproductive class, *ACT1, ACT3, ACT12, ACT4* and *ACT11* all are highly expressed in reproductive tissues (An et al., 1996a;
Huang et al., 1996a; Huang et al., 1997). Coincidentally, each subdivision of actin genes based on sequence homology also has its own specific expression pattern. In the reproductive actin subclass I, both ACT1 and ACT3 are highly expressed in mature pollen and organ primordia. All expressed A. thaliana actin genes are unlinked (McKinney and Meagher, 1998) and do not show any signs of gene conversion phenomena (McDowell et al., 1996b). Therefore the vegetative and reproductive classes of actin have evolved separately, suggesting that indeed the gene products and/or the regulation of those gene products are necessary for plant survival.

Animal actin isoforms are divided into two classes, the non-muscle actins β and γ and smooth muscle γ, and the muscle actins cardiac α, skeletal α and vascular α (Herman, 1993). These isoforms are functionally distinct and necessary proteins through studies of developing muscle tissue. For example, during switching from proliferation to differentiation myoblast, cells will stop production of non-muscle actins and expression of skeletal α actin occurs (Herman, 1993). It has also been shown that iso-actin subcellular sorting occurs because vascular endothelial cells in culture localize β-actin only at the edges of the plasma membrane (Hoock et al., 1991). Isoforms in plants point to some regulation at the protein level (Meagher et al., 1999a; Meagher et al., 2000).
DNA sequence analysis demonstrated the existence of three ancient and evolutionarily divergent classes of soybean actin: κ, λ, and υ (Hightower and Meagher, 1985). These classes are separated by at least 350 million years, clearly predating the divergence of angiosperms and gymnosperms. To determine if these actin genes produce protein products that are necessary for appropriate regulation, polyclonal antibodies were raised against synthetic actin peptides and the antisera was examined in soybean tissues for differential patterns of expression (McLean et al., 1990). Antisera recognizing κ and υ were found in root cap. Because the soybean actin gene family is so large, further analysis of the plant actin superfamily has been conducted in A. thaliana. A. thaliana is an excellent plant model system as it contains a small genome estimated to be between 130 and 140 Mbp of DNA. A. thaliana has excellent genetic and physical maps available for analysis. It also has a short life cycle and takes very little space in a growth chamber, and perhaps most importantly is easily transformable (The Arabidopsis Genome Initiative, 2000).

To more directly address the functionality of the actin proteins in plants, a series of monoclonal antibodies was raised against a series of actin MAP peptides. Originally three class of antibodies were reported, MAbGEa, MAbGPa, and MAb45a (Kandasamy et al., 1999). MabGEa is a general antibody that reacts with all A. thaliana actins as well as vertebrate actins. MAbGPa reacts with all A. thaliana actins, additionally there are
other monoclonals that react with mature pollen of not only *A. thaliana*, but of all angiosperms and advanced gymnosperms. Because antibodies can detect a mature pollen-specific epitope over an ~400 million year separation, it strongly suggests that these pollen specific actins have been conserved because they are functionally necessary (Meagher et al., 1999a).

**IV. Plant actin expression patterns**

A 40 base pair sequence in the 3’ end of the chicken β-actin gene is responsible for down-regulation of expression during myogenesis (Deponti-Zilli et al., 1988). Smooth muscle actin has also been shown to have elements within the 5' intron important for high levels of expression (Mack and Owens, 1999). Like animal actins, plant actins have complex temporal and spatial patterns of expression, and it is hypothesized that these patterns are driven by key elements within the 5' flanking region of each gene.

All eight expressed *A. thaliana* actin genes have been examined for expression using Northerns, RT-PCR and constructs with the 5' flanking region fused to a β-glucuronidase reporter gene. The vegetative ACT2/8 subclass has very high expression in all vegetative tissues, such as root, stems, and leaves (An et al., 1996b). Very low amounts of *ACT2* RNA in pollen are detected through RT-PCR analysis, though no expression can be detected in reporter gene assays. No translational activity was found in any other reproductive tissue. *ACT7*, representative of a second subclass of reproductive genes, was found strongly expressed in rapidly dividing tissue (McDowell et al., 1996a).
It was also shown that the \textit{ACT7} 5’ flanking region contains elements responsive to several growth hormones (Kandasamy et al., 2001; McDowell et al., 1996a).

The reproductive class of actins includes related genes that are evolutionarily ancient, and as the name implies, expressed mainly in reproductive tissues. There are three subclasses of reproductive actin genes: \textit{ACT4/12}, \textit{ACT11}, and \textit{ACT1/3}. The \textit{ACT4/12} subclass is expressed primarily in mature pollen, though very low expression is seen in \textit{ACT12} root tissue, mostly in pericycle tissue surrounding lateral roots (Huang et al., 1996b). \textit{ACT11} is a single gene in a second subclass and is expressed throughout pollen development, though most strongly in mature pollen. \textit{ACT11} is distinct from the other reproductive actins in that it is the only actin gene expressed throughout ovule development as shown by reporter gene assays driven by the 5' flanking region (Huang et al., 1997). The \textit{ACT1/3} subclass is expressed predominantly in mature pollen and both reproductive and vegetative organ primordia (An et al., 1996a). Specifically these genes are expressed in the tips of both primary and lateral roots, lateral root organ primordia, leaf primordia, and floral organ primordia. Additionally, \textit{ACT1} and \textit{ACT3} are both expressed weakly in unfertilized ovules and immature pollen. Though \textit{ACT1} shares an identical expression pattern with \textit{ACT3}, it is expressed 1000 times more strongly than any other reproductive actin in mature pollen as determined by quantitative RT-PCR (An et al., 1996a).

V. Cytoskeletal gene expression patterns
Plant actin genes arose from a common ancestor and diverged around the time of land plant emergence (McDowell et al., 1996b). Sequence analysis, RNA expression data, and class specific antibody epitope studies (Kandasamy et al., 1999) indicate that plant actin genes exist as vegetative and reproductive classes. Because these classes were formed around the estimated time of land plant organ development, it is hypothesized that plant actin gene classes evolved cooperatively with vegetative structures such as leaves and reproductive structures such as pollen. A second hypothesis based on the evolution of the actin gene family is the proposed coevolution of actin with other cytoskeletal genes (Meagher et al., 1999b). This coevolution could have been selected for in the same fashion as the actin gene family itself, since these genes were needed to form vegetative and reproductive organs either in an expression-dependent manner and/or protein interaction dependant manner. For instance, profilin genes in both maize and A. thaliana are expressed divergently in vegetative and reproductive tissues (Kovar et al., 2000; Meagher et al., 1999b). Additionally, maize ADF and Arabidopsis α–tubulin and β-tubulin multigene families also display tissue-specific expression patterns whereby the expression can be divided into vegetative and reproductive types (Meagher et al., 1999b).

The expression pattern of ACT1 is a fascinating representative of the reproductive class of actin genes to further study because it is the major actin gene present in mature pollen. Thus ACT1 was the focus of the work in this dissertation. ACT1 and ACT3 differ in only one conservative amino acid change within their coding regions. However they
have no homologous sequences in noncoding sections such as introns, 3’ UTR, or the 5’ flanking region with the exception of a 55 base pair region immediately upstream of the TATA box (An et al., 1996a). Because an epitope of mature pollen has been conserved throughout advanced land plant evolution, we examined the evolution of actin pollen-specific expression in heterologous plant species (Chapter 3). We examined the fidelity with which the ACT1 promoter would guide expression in evolutionarily distant organisms, reasoning that if the expression pattern was conserved then transactivating factors were likely also conserved. An active promoter displaying correct expression would strongly suggest the necessity of the ACT1 gene throughout land plant evolution.

It was shown that the ACT1 5’ flanking region contains many elements which control high levels of expression in pollen and organ primordia. Transcriptional machinery guiding reproductive actin gene expression was shown to be conserved between distantly related species supporting the hypothesis that actin gene duplications could have preserved due to the necessary expression of the protein in specific spatial and temporal patterns. The ACT11 promoter was also examined in an additional species because it is representative of another actin gene subclass strongly expressed in both mature pollen and developing pollen and its expression varies slightly from ACT1.

To more thoroughly test elements of the ACT1 promoter important for extremely strong pollen expression, many expression constructs mutagenizing the 5’ flanking region translationally fused to a β-glucuronidase reporter gene were made (Chapter 2). Many
parts of the 5' flanking region were tested for their ability to guide ACT1 expression.

Most notably, the conserved region was tested for regions important for expression as well as the distinctive 430 bp leader intron. Transgenic plants containing various constructs were examined qualitatively in floral and seedling tissues. A novel quantitative assay was created to measure β-glucuronidase expression in mature pollen.
Literature Cited


CHAPTER 2

THE 5' FLANKING REGION OF ACT1 ARABIDOPSIS ACTIN GENE CONTAINS MULTIPLE ELEMENTS NECESSARY FOR HIGH-LEVEL, MATURE POLLEN EXPRESSION\textsuperscript{1}

\textsuperscript{1} A. V. Vitale and R. B. Meagher to be submitted to\textit{Planta}
Abstract

Actin genes in *Arabidopsis thaliana* exist as an ancient gene family. *A. thaliana* actins are comprised of two ancient classes that can be subdivided further into five subclasses. *ACT1* and *ACT3*, the result of a duplication event 30-60 MYA, comprise one of the five ancient subclasses and both are expressed at high levels in mature pollen, embryo sac, and organ primordia. *ACT1* and *ACT3* share little homology with the exception of a 55 base-pair conserved sequence within the 5' flanking region. Modifications of the *ACT1* 5' flanking region were examined in transgenic *A. thaliana* to determine qualitative differences of expression guided by various regions of the 5' flanking region. A sensitive fluorescence assay measuring reporter gene expression was designed to quantify changes in expression in various promoter/GUS fusion constructs. Truncations of the 5' upstream flanking region and within an adjacent intron, and site-directed mutagenesis constructs of the conserved region demonstrated that all these regions are needed to promote high levels of mature pollen expression. Even a subtle change in the conserved region in the form of a replacement construct caused a nine-fold decrease in expression in mature pollen. The first intron in *ACT1*, located in the 5' UTR, is required for high levels of pollen expression and organ primordia expression. The leader intron functions in a gene-specific manner since substitution of the leader intron of *ACT1* with the leader intron of a vegetatively expressed actin efficiently down-regulated pollen expression 10-20 fold. Thus, multiple sequence elements within the *ACT1* 5' flanking region direct high levels of mature pollen-specific and organ primordia expression.
Introduction

Plant actin cytoskeleton genes participate in many important plant processes such as control of the cell division plane, cytoplasmic streaming, cell elongation, pollen tip growth, and organelle transport (Boevink et al., 1998; Cai et al., 1997; Kandasamy et al., 2001; Kandasamy and Meagher, 1999; McCormick, 1993; Smith et al., 1998). In this work we dissect the cis-acting regulatory regions of \textit{ACT1}, a highly expressed reproductive \textit{A. thaliana} actin gene. Review of the expression and evolution of the plant actin gene family indicated the \textit{ACT1} subclass has an important and unique expression pattern. Investigation of similarly expressed cytoskeletal genes and other highly expressed reproductive genes may illuminate reproductive patterns of expression, pointing to global mechanisms of gene expression and the co-evolution of these genes with floral structures.

All higher plants contain a multigene family of actin genes which can be divided based on evolutionary sequence analysis into two major groups, reproductive and vegetative (McDowell et al., 1996b). In \textit{Arabidopsis}, the two major classes of actin genes are comprised of eight expressed genes whose patterns of gene expression can be superimposed onto the phylogenetic relationship of both the reproductive and vegetative gene classes. The actin genes diverged into these two classes sometime around the evolution of land plant vegetative and reproductive tissues. The Arabidopsis reproductive actin class contains five genes that can be organized into three subclasses based on phylogenetic relationships and expression patterns deduced from experiments involving reporter gene constructs and RNA steady-state levels (An et al., 1996a). Each of these subclasses has its own unique temporal and spatial patterns of expression. The \textit{ACT12} and \textit{ACT4} genes form
one subclass and are expressed in mature pollen. *ACT1* forms a separate second subclass and is the only actin expressed throughout ovule, seed, and fruit development. The third subclass, that of *ACT1* and *ACT3*, is the focus of this work. The *ACT1* and *ACT3* genes form a specific reproductive subclass expressed in mature pollen, unfertilized ovules and floral organ primordia. *ACT1* and *ACT3* also display some expression in vegetative structures such as young vascular tissues, root tip, lateral root primordia, cotyledons, and leaf primordia. Previous work has demonstrated that *ACT1* contains the late pollen-specific epitope that has been conserved among higher gymnosperms (Gnetales), monocots, and dicots (Kandasamy et al., 1999). Thus, the late pollen-specific gene subclasses arose at least 220 million years ago. It is hypothesized that this ancient class of actin genes has been preserved throughout land plant evolution because it encodes proteins with specialized actin functions (Meagher et al., 1999a) and/or because the specific spatio-temporal expression patterns are essential for reproductive tissue development (Meagher et al., 1999b). Almost nothing is known about the cis-sequences that control the reproductive expression pattern of the *ACT1* subclass. The expression patterns of the *ACT1* gene and the closely related *ACT3* gene were originally determined using RNA steady-state measurements and 5' flanking region translational fusions to a β-glucuronidase reporter gene (An et al., 1996a). This reporter gene included the leader exon, leader intron, and initial 19 amino acid codon positions. Using 5' flanking regions of 2.7 kb and 2.1 kb from the start of translation, respectively, it was shown that *ACT1* and *ACT3* directed expression most strongly in mature pollen with significant expression also in all organ primordia, root tip, root vascular tissue, and ovules. Quantitative RT-PCR studies determined that *ACT1* was expressed 1000 times
more strongly in mature pollen than in leaves (An et al., 1996a). For this reason we chose to study ACT1 elements important in guiding actin expression in reproductive tissue.

It was originally found that a gene, TN167, was transcribed in the opposite orientation approximately 900 base pairs upstream of the ACT3 gene (An et al., 1996a). It was therefore postulated that elements important for ACT1 gene expression were located within ~1 kb upstream of the major start site of transcription in an area corresponding to the region between the start site of transcriptions of TN167 and ACT3. Within the ACT1 5' flanking region there is no sequence similarity to ACT3 with the exception of an almost perfectly conserved region of 55 base pairs immediately upstream of the proposed TATA box. This homology is likely significant because ACT1 and ACT3 share identical expression patterns, the ACT1 and ACT3 amino acid sequence differs by a single conservative change and analysis of synonymous nucleotide substitution within codons suggests the two genes recently diverged approximately 30-60 MYA (McDowell et al., 1996a).

ACT1 is the most strongly expressed actin in mature Arabidopsis pollen and is categorized as a late pollen-expressed gene (An et al., 1996a; Mascarenhas, 1992; Meagher et al., 1999b). Other late pollen specific genes such as the Lycopersicon esculentum LAT59, LAT56 and LAT52 genes have been shown to contain conserved promoter elements that guide mature pollen expression (McCormick et al., 1991; Twell et al., 1990; Twell et al., 1991). Because of problems in examining pollen-specific expression in transient assays due to ectopic expression and the large amount of pollen needed for protein analysis of transgenic plants we developed a quantitative assay using whole transgenic anthers. Additionally, we examined the DNA sequence for elements within the 5' flanking region of
to identify those contributing to the overall expression pattern and those most important for high levels of pollen specific expression.

**Materials and Methods**

*Construction of 5' deletion plasmids:* 2.2 kb of *ACT1* 5' flanking region with respect to the major start site transcription from pBI1E1 (An et al., 1996a) was subcloned into pBluscriptSK- to make p2200bs which was used as the template for all PCR cloning. Sense oligonucleotides encompassing 5' *HindIII* restriction sites were designed to create 5' deletions via PCR cloning relative to *ACT1* major start site of transcription at positions -310bp, -238bp, -88bp, -41bp, and -32bp (Figure 2-1 and Table 2-1). All sense oligonucleotides were used with the antisense oligonucleotide *ACT1*-EX1N (EX1N) (Table 2-1), which was designed complementary to the first 19 codon positions and included a *Bgl*II site at the 5' end. PCR products from -310/EX1N, -238/EX1N and -88/EX1N primer pairs were subcloned into pSL1180 (Pharmacia) creating plasmids p310bs, p238bs, and p88bs respectively. PCR products from primer pairs -41/EX1N and -32/EX1N were subcloned into pCR2.1 using Invitrogen’s PCR Cloning Kit to create plasmids p41bs and p32bs respectively. The subcloned 5' flanking regions were then cloned into plant expression binary vector pBI101.1 (Clontech) using *HindIII* as the upstream cloning site and a *Bgl*II/ BamHI fusion downstream creating p310, p238, p88, p41, and p32.

*Construction of site-directed mutagenesis plasmids:* p310bs was used as the template for all the site-directed mutagenesis constructs. *ACT1* and *ACT3* conserved promoter regions were targeted for mutagenesis (Figure 2-1). *ACT1* promoter sequence was replaced with DNA sequences containing an *Xba*I site and random nucleotides using Stratagene’s
Quickchange Site-Directed Mutagenesis kit. Positions relative to the major start site of transcription were mutageneized at -87 to -78, -77 to -68, -67 to -58, -57 to -48, and -47 to -36 to make plasmids p164Mbs, p87Mbs, p77Mbs, p67Mbs, p57Mbs, and p47Mbs, respectively (Table 2-1). The mutagenized promoters were cloned into pBI101.1 as before to create p87M, p77M, p67M, p57M, and p47M.

ACT1 intron deletion and replacement constructs: An expression construct containing a deletion of the ACT1 intron L was made by first amplifying the 5’ flanking area upstream of the leader intron using the sense primer -310 and antisense primer ACT1LE-EX1N (Table 2-1). ACT1LE-EX1N created a PCR product with 3’ end complementary to EX1N. The product was purified using agarose gel extraction and reamplified with the same sense primer and EX1N as the antisense primer to add the first 19 codon positions to the product. The second PCR product was then subcloned into pSL1180 to create p310.ivsbs and the subcloned fragment was then moved to pBI101.1 to make p310ivs. The leader intron was cloned upstream of the ACT1 5’ flanking region sequences in p310ivs by first PCR amplifying the leader intron with the sense oligo ACT1LIS and the antisense oligo ACT1LIA. The intron fragment was then cloned into the HindIII site of p310ivs in the forward and reverse directions to produce pLIB-L and pLIB-R respectively. The orientations were determined by automated sequencing.

The ACT1 leader intron was replaced by the ACT2 leader intron via a PCR cloning strategy. Using the sense primer -310 and antisense primer ACT1-457A PCR was performed to produce a DNA fragment between -310 and the start of the ACT1 leader intron. A DNA fragment containing the ACT2 leader intron complementary to the ACT1
UTR at its 5’ end and complementary to the ACT1 UTR and first exon and its 3’ end was created by PCR with the sense primer ACT1-2LIS and antisense primer ACT1-2LIA using the template pAACT2S (An et al., 1996b). The two fragments were joined and the first 19 codon positions were added along with a BamHI site at the 3’ end using overlap extension PCR with the sense oligo -310 and the antisense oligo AC1EX1BAM (Table 2-1). The overlap extension PCR product was cloned into pCR2.1 using Invitrogen’s PCR Cloning Kit to construct pLIRcr. The intron replacement was then cloned from pLIRcr using HindIII and BamHI restriction sites into pBI101.1 to create pLIR.

Construction of gain-of-function plasmids: A minimal 35S promoter expression vector (pMIN) was constructed by restricting the -45 CaMV promoter fragment (Benfey et al., 1990) with SalI and BamHI from pBM259 (Marcotte et al., 1989) and cloning the fragment into pBI101.1. A fragment of DNA representing the conserved ACT1 promoter region was made by annealing two 5’ phosphorylated oligonucleotides (ACT1CON-S and ACT1CON-A) together. Equimolar amounts of ACT1CON-S and ACT1CON-A were heated at 94°C for 5 minutes and then cooled to 4°C. When annealed, the fragments formed a partial HindIII site at the 5’ end and a partial SalI site at the 3’ end. The 55 bp region of the ACT1 promoter was then cloned into pMIN using HindIII and SalI restriction sites, creating pCIN.

Plant transformation and analysis: Arabidopsis thaliana (RLD) was transformed by the method in (An et al., 1996a) or Green (http://www.bch.msu.edu/pamgreen/vac.htm)(Green, 2001). T1, T2, T3 or T4 plants were used for qualitative and quantitative GUS analysis. 7- to 10-day-old seedlings, 3-week-old seedlings for selected constructs, and mature plant
inflorescences were examined qualitatively for GUS activity according to the method described previously (An et al., 1996a). To determine 5’ flanking regions important for the high levels of pollen expression displayed by ACT1, quantitative analyses of GUS expression in mature pollen were performed using 1 millimolar methylumbelliferone. (MUG) Dehiscent anthers of a single plant from flowers at stage 13 - 14 (Smyth et al., 1990) were placed two at a time into 8 wells of a 96-well microtiter plate (Falcon) each containing fifty µl of 7 percent sucrose pH 7.4. 50 µl of GUS extraction buffer (Jefferson et al., 1987) was added to each well and over a time course from 5, 10, 15, and 25 minutes 100 µl of 2M NaCO₃ was added to 2 wells at a time to quench the reaction. The fluorescence from each well was read at 360 nm/480 nm on a Biolumen 960 microtitre plate reader (Molecular Dynamics). The fluorescence values from the same time point were averaged together and normalized to plant lines transformed with p310T. Every run of samples contained an anther time course from an untransformed plant as a zero expression standard. To detect variations due to equipment and to serve as an additional measurement for sample comparison, InSpeck Blue beads (Molecular Probes) were used at dilutions of 100, 30, 3, 1, 0.3 and 0 percent and shown to have similar values with each set of assays.

Results

5' deletion of the ACT1 5' flanking region qualitative and quantitative results

In order to define a 5’ region of ACT1 that contained all the elements needed for normal expression, an initial truncation of the ACT1 5' flanking region was constructed -310 bps
from the major start site of transcription (Figure 2-1A). This construct, p310, directed
expression exactly as a previously described ACT1 reporter gene construct containing 2.2
kb of upstream sequence from the major start site of transcription (An et al., 1996a). This
construct included strong pollen ovule and organ primordia expression (Figure 2-2A-E).
p310 was used as the base construct in a series of additional 5' flanking region dissections
(Figure 2-2A and Table 2-2A and Table 2-2B).

To identify important regions upstream of the conserved sequence box, two other 5'
deletion constructs were examined in transgenic A. thaliana. The first one corresponds to
the region -238 bp from the ACT1 major start of transcription (p238). The other
corresponds to the region -88 bp (p88) from the start site of transcription and contains a 5'
deletion of all but the conserved promoter area (Figure 2-1A). The p238 construct
demonstrated qualitative expression equivalent to that of the base p310 construct (Figure 2-
2A-E and Table 2-2A and Table 2-2B)). However, deletion of the region upstream of the
conserved promoter area in p88 reduced ovule expression and completely removed
expression in vegetative organ primordial tissue (Figure 2-2F and Figure 2-2H and Table 2-
2A and Table 2-2B). Floral organ primordia expression was absent from 5 out of 9 lines
examined and in the remaining lines expression was visibly lowered.

ACT1 is by far the most highly expressed of all four late pollen actins thus visible
confirmation of expression level is highly subjective and difficult. To quantify the role of
various elements within the ACT1 5' flanking region in pollen expression a MUG-based
fluorometric expression assay was created, as shown in Figure 2-3A. The GUS expression
levels in four pairs of dehiscent anthers from a transgenic line were analyzed after
incubation with MUG substrate for varying amounts of time. The relative rates of activity per anther was determined from the slope of the line created from this data (Figure 2-3A and Figure 2-3B). The -310 deletion construct was defined as having 100 percent activity and used to normalize the expression levels from other constructs. Using the p238 deletion construct it was found that the ACT1 promoter sequence from -310 and -238 does not contribute significantly to high levels of pollen expression because the expression decreases very little upon deletion of this sequence (Table 2-2 and Figure 2-3C). Data from the p88 deletion construct suggests that sequences from -238 to -88 contribute strongly to expression because MUG levels drop to 20 percent of the p310 control (Figure 2-3C). A fourth deletion construct, p41, which leaves only the putative TFIID-binding site TATATTTT, showed still lower expression in pollen, about 10 percent compared with the control. This result concurs with the visibly reduced expression in β-glucuronidase assays (Figure 2-2G and Figure 2-3C). When the TATA region of the ACT1 promoter is truncated, (p32) expression is essentially zero (Figure 2-2I and Figure 2-3C).

Site-directed mutagenesis of the ACT1 promoter region significantly affects quantitative pollen expression

To determine if the ACT1 conserved region contained elements specific for tissue expression and/or levels of expression in the context of the full promoter, five site-directed mutagenic constructs were designed to replace the conserved region in ten base pair increments using the base –310 bp deletion construct (Figure 2-1B). These constructs were transformed into A. thaliana and visible differences in GUS staining amount were present in some tissues. Most notably p87M, which replaces the promoter region from -87 to -78
bp from the major start site of transcription, had visibly decreased pollen expression in GUS qualitative assays (Table 2-2A and Table 2-2B) and was more than 85 percent lowered in quantitative assays (Figure 2-3D). Additionally p47MBI, which replaces -47 to -38 bp from the major start site of transcription, had no floral organ primordia expression in three out of four lines examined and visibly lower expression in mature pollen (Table 2-2A). The quantifiable differences in pollen-specific expression are shown for replacements for -67 to -57 and -47 to -37 (Figure 2-3D). Each construct had nine-fold lower expression when compared with the -310 construct.

**ACT1 conserved region alone is not sufficient to guide pollen expression**

Deletion of the conserved region drastically reduced pollen expression. In order to determine if the ACT1 conserved region was necessary and sufficient to guide strong pollen and ovule expression, the 55 bp conserved region was cloned in front of a minimal CaMV35S promoter and GUS reporter gene (pCIN). This same base promoter construct responds to auxin response elements (Benfey and Chua, 1990). The conserved sequence alone was not sufficient to act as an enhancer and direct ACT1 expression pattern (data not shown) because no pollen expression was directed from this construct.

**ACT1 first intron is an important regulator of the ACT1 expression pattern**

An interesting feature that is evolutionarily conserved in all plant actin 5' flanking regions characterized to date is a large first intron that separates most of the 5' UTR from the body of the coding sequence. To examine the contribution this intron makes to the regulation of actin expression, a construct was created that perfectly deleted the 423 bp leader intron of the ACT1 5' flanking region from the -310 (p310.ivs) (Figure 2-1). The
deletion of this intron from the construct prevented GUS expression in the organ primordia in most seedlings and all floral organ primordia expression (Figure 2-2J). Quantitative β-glucuronidase assays showed that levels of mature-pollen specific expression ten percent of the –310 construct (Figure 2-3E). Thus, the first intron is essential to all aspects of ACT1 gene expression.

Because both qualitative and quantitative expression were lost when the first intron was deleted, reporter gene constructs were designed to test the ability of the intron to function as a typical enhancer element. To this end, the 423 bp ACT1 first intron was cloned upstream in both forward and reverse orientations of the -310 bp promoter in modified p310.ivs constructs. The new constructs pLIBBIF and pLIBBIR (Figure 2-1C) displayed expression patterns identical to the original intron deletion construct. Expression in seedlings, organ primordia and unfertilized ovules was largely absent, but expression in mature pollen appeared strong according to β-glucuronidase staining assays (Table 2-2A and Table 2-2B). Though the expression pattern was not rescued, the reporter gene construct had GUS rates comparable the p310 control lines in mature pollen (Figure 2-3E). These data suggest that the first intron possesses strong enhancer-like activity for pollen expression,

Perhaps just the presence of any intron specifically upstream of the ACT1 coding region is essential for some properties of ACT1 expression. To test this the ACT1 leader intron (intron L) was precisely replaced with the ACT2 intron L to make the reporter construct pLIR. This construct, pLIR, replaced the ACT1 5' intron with an intron of similar size (423 bp in ACT1 vs. 480 bp in ACT2). The ACT2 intron L was chosen because it is
from another *A. thaliana* actin gene, and is expressed contrapuntally from ACT1; its expression is very high in vegetative tissues but very low in reproductive tissues such as ovules, embryos, and pollen. The resulting expression pattern from construct pLIR had recover expression in young seedlings and organ primordial tissue, but unexpectedly had virtually no residual expression in mature pollen. This loss is surprising considering that this construct still contains all 310 bp of the upstream regulatory information, the TATA domain, and an intact 5' UTR.

**Discussion**

*Expression elements found in the ACT1 promoter*

The ACT1 5' region was examined for areas contributing to the overall expression pattern of ACT1 through a series of 5' deletions translationally fused to the GUS reporter gene. Truncations furthest from the start site of transcription had no qualitative effect on expression. Truncations that included only the conserved region of the ACT1 promoter and the TATA region of the promoter retained pollen and ovule expression. The construct p88 containing the conserved region is sufficient to guide mature pollen expression (Figure 2-2F and Table 2-2A and Table 2-2B), though levels of GUS expression were reduced by half according to quantitative assays (Figure 2-3C). In contrast, a deletion construct (p32) that removed all sequence upstream of the TATA promoter element demonstrated reduced expression in all tissues examined including mature pollen (Figure 2-2I, Table 2-2A and
Table 2-2B, Figure 2-3C). The conserved region may thus contain elements important for qualitative and/or quantitative reproductive actin gene expression.

The area between –310 bp and –88 bp from the major start of transcription is needed to confer highest levels of pollen specific expression as well as expression in young seedlings, root, leaf, and floral organ primordia. As expected, excluding the TATA box in the -32 truncation construct completely abolished mature pollen expression in most transgenic lines. These expression results from the promoter deletion constructs agree with results from other pollen promoter dissection studies: most promoters contain multiple areas that direct high levels of expression (Twell et al., 1991). Unlike previous studies involving homologous regions of LAT52 and LAT59 pollen-specific genes, the conserved area, shown here to be important for pollen expression, was not sufficient to guide expression in a gain-of-function assay (Twell et al., 1991). This finding suggests that certain aspects of the totality of actin promoter structure, such as the interaction of factors bound to these different tissue-specific domains, is required for high expression levels.

**ACT1 conserved box contains multiple redundant areas that guide high levels of mature pollen expression**

Analysis of the constructs that had site-directed changes within the conserved box did not uncover any qualitative discrepancies in expression patterns compared to the full-length 310 bp construct (Figure 2-2A-E). Because this homologous 55 bp box is the only sequence conserved between *ACT1* and *ACT3*, we examined the region for any quantitative expression changes in mature pollen using a tissue-specific quantitative expression assay. Transgenic analysis of expression was preferable to transient transfection because the
uncertainties associated with bombarding cells (i.e., ectopic expression, gene copy number) would make the data hard to apply to an in vivo expression situation. The conserved region, rather than having separate elements that function to isolate ACT1 expression in a temporal manner, has multiple elements that guide high levels of expression in mature pollen. This is unlike the constitutive CaMV promoter, which contains short sequence elements specific for both temporal and spatial expression (Benfey and Chua, 1990). Rather, the expression guided by the ACT1 conserved region is like other pollen promoters studied, in which multiple elements contribute to high levels of pollen expression (Twell et al., 1990; Twell et al., 1991).

ACT1 first intron may specifically upregulate high levels of pollen expression

The ACT1 first intron qualitatively contributes to expression in ovules and organ primordia. Perfect deletion of the ACT1 first intron in a reporter gene construct resulted in loss of expression from most lines in seven-day-old seedlings, leaf primordia, floral organ primordia, and immature ovules. This finding agrees with the effect that 5' introns exert on plant reporter gene construct expression in monocot systems. It is generally observed that an intron at the 5' end of the gene is necessary for enhanced levels of gene expression (Donath et al., 1995; Mascarenhas et al., 1990; Snowden et al., 1996). Deletion of the first intron can lower expression in a transcription-dependent manner, and increased levels of steady-state RNA are associated with normal splicing of the 5' most introns (Tanaka et al., 1990). The intron-mediated enhancement (IME) of ACT1 expression by intron L is the first example of this enhancement occurring in a dicot system.
When the intron was ectopically present upstream of the promoter (pLIBBIF, pLIBBIR) region, it directed the high levels of gene expression. Quantitative assays indicated that expression in mature pollen was partially recovered in comparison to the intron deletion construct. It seems that the intron does function as a tissue-specific enhancer element, because these rearrangements to an upstream location rescued quantitative expression in pollen. Thus, the ACT1 intron is unique in conferring both typical enhancer and intron-mediated upregulation effects on the expression guided by the ACT1 promoter.

Substituting the ACT2 intron L for the ACT1 intron L severely regulated lower expression in mature pollen, but rescues expression in organ primordia and seven-day-old seedlings. IME is a nuclear process and unlinked to translation in monocots and most importantly is a gene-dependent process (Rethmeier et al., 1997). This is in contrast with vertebrates; IME in Xenopus oocytes is a product of efficient splicing in the nucleus that leads to highly efficient translation of the mRNA (Matsumoto et al., 1998). Introns can act post-transcriptionally to upregulate gene expression in the Arabidopsis tryptophan pathway gene PAT1 (Rose and Last, 1997). We propose that the IME demonstrated for ACT1 intron L is due to a gene-specific enhancement effect. However, an alternative explanation is that the ACT2 intron L inhibits pollen expression, while allowing or encouraging expression in organ primordia and seedlings. This finding also agrees with the data of An et al. (1996) that suggests ACT2 mRNA expression is uniquely repressed in pollen.
Literature Cited


Table 2-1 This table depicts all oligonucleotides referred to in Materials and Methods as well as the oligo sequence and region(s) of homology with ACT1, ACT2, mutagenized sequence or sequences necessary for vector construction such as restriction site additions.

<table>
<thead>
<tr>
<th>Oligo</th>
<th>Sequence</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACT1-32TS</td>
<td>5'TTGACAAGCTTTTTACTCCTCCCA TTCCCTTCT 3'</td>
<td>-32 to -10 from the major start site of transcription. <em>Hind</em>III site for cloning purposes underlined.</td>
</tr>
<tr>
<td>ACT1-41TS</td>
<td>5'TTGACAAGCTTTTTTTATATTTAC TCCTCCATT 3'</td>
<td>-41 to -19 from the major start site of transcription. <em>Hind</em>III site for cloning purposes underlined.</td>
</tr>
<tr>
<td>ACT1-47MA</td>
<td>5'GAGGAGTAAATATAAGTTGCTTAG ATAAAGGCTTCGACTTC 3'</td>
<td>Mutagenesis of the promoter region from -47 to -37 in italics from the major start site of transcription (antisense orientation)</td>
</tr>
<tr>
<td>ACT1-67MA</td>
<td>5'GGTCCAAAGGCTTCGGTACTCTA GAACCTTGTTTTCCGGT 3'</td>
<td>Mutagenesis of the promoter region from -67 to -58 in italics from the major start site of transcription (antisense orientation)</td>
</tr>
<tr>
<td>ACT1-77MA</td>
<td>5'TTCGACTTCGGCTCAGTCTAGACA CCGTTGCTATATTTC 3'</td>
<td>Mutagenesis of the promoter region from -77 to -68 in italics from the major start site of transcription (antisense orientation)</td>
</tr>
<tr>
<td>ACT1-57MS</td>
<td>5'AAGGTGAGCCGAAGTTCTCTAGA ACGGACCCCGTTTTATA 3'</td>
<td>Mutagenesis of the promoter region from -57 to -48 in italics from the major start site of transcription (antisense orientation)</td>
</tr>
<tr>
<td>Sample Name</td>
<td>Sequence</td>
<td>Description</td>
</tr>
<tr>
<td>-------------</td>
<td>----------</td>
<td>-------------</td>
</tr>
<tr>
<td>ACT1-87MA</td>
<td>5'GGCTCACCTTGTTTTTTCTAGAGC AATTTCATTTCAAAATC 3'</td>
<td>Mutagenesis of -87 to -78 in italics from the major start site of transcription (antisense orientation)</td>
</tr>
<tr>
<td>ACT1-77MS</td>
<td>5'GAAATATTGAACCGGTGCTCTAGA CCGAGCAGAAGTCGAA 3'</td>
<td>Mutagenesis of -77 to -68 in italics from the major start site of transcription (sense orientation)</td>
</tr>
<tr>
<td>ACT1-57MA</td>
<td>5'TATAAAAACGGGTCCGTCTAGAGAACTTCGGCTCACCTT 3'</td>
<td>Mutagenesis of the promoter region from -57 to -48 in italics from the major start site of transcription (sense orientation)</td>
</tr>
<tr>
<td>ACT1-87MS</td>
<td>5'GATTTGAAATGAAATTGCTCTAGAGTAAACAAGGTGAGCC 3'</td>
<td>Mutagenesis of the promoter region from -87 to -78 in italics from the major start site of transcription (sense orientation)</td>
</tr>
<tr>
<td>ACT1-47MS</td>
<td>5'GAAGTCGAAGCCTTTATCTAGACCTATATTTACTCCTC 3'</td>
<td>Mutagenesis of the promoter region from -47 to -38 in italics from the major start site of transcription (sense orientation)</td>
</tr>
<tr>
<td>ACT1-67MS</td>
<td>5'ACCGGAAAACAAGGTTCTAGAGTACCGAAGCCTTTGGACC 3'</td>
<td>Mutagenesis of the promoter region from -67 to -58 in italics from the major start site of transcription (sense orientation)</td>
</tr>
<tr>
<td>ACT1-310TS</td>
<td>5'AGCGTCAAGCTTGAATTCTAGACAOTACTACCACCA 3'</td>
<td>-310 to -282 from the major start site of transcription, HindIII and EcoRI site added on 5' end (underlined)</td>
</tr>
<tr>
<td>ACT1-88TS</td>
<td>5'TTGACAAAGCTTTATAGAACCGGA AAAACAAGGTG 3'</td>
<td>-88 to -66 from the major start site of transcription, HindIII restriction site at the 5' end (underlined)</td>
</tr>
<tr>
<td>Oligo Name</td>
<td>Sequence</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>ACT1-238TS</td>
<td>5'TTGACAAGCTTTTTATTTACTCCG ATTTGATGG3'</td>
<td>-238 to -216 from the major start site of transcription, HindIII site added at the 5' end (underlined)</td>
</tr>
<tr>
<td>ACT1LIUP</td>
<td>5'TTGACAAGCTTTGTACTTTTTGAG ACCCTTTGTT 3'</td>
<td>Act1 leader intron sense oligo. HindIII site added (underlined)</td>
</tr>
<tr>
<td>ACT1LIDOW</td>
<td>5'TTGACAAGCTTTGTACTTTTTGAG ACCCTTTGTT 3'</td>
<td>Act1 antisense leader intron oligo. HindIII site added (underlined)</td>
</tr>
<tr>
<td>ACT1CON-A</td>
<td>5'TCGACAAACGCGGTCCAAAGGCTT CGACTTCGGCTCACCCTTTTCCG GTTCTATA 3'</td>
<td>-88 to -38 from the major start site of transcription, partial SalI restriction at the 5' end, partial HindIII site at the 3' end (underlined)</td>
</tr>
<tr>
<td>ACT1CON-S</td>
<td>5'AGCTTTATAGAACCGGAAAACAA GGTGAGCCGAAGTCGAGGCGTT GACCCGTGTTG 3'</td>
<td>-88 to -38 from the major start site of transcription, partial HindIII restriction at the 5' end, partial SalI site at the 3' end (underlined)</td>
</tr>
<tr>
<td>ACT1-EX1N</td>
<td>5'CGCACAGATCTAACCATTCCAGT TCCATTTCGCAAAACAAGGTTG AATGTCTACCACCACGCTTTTC TTCTAC 3'</td>
<td>-10 to +57 from the Act1 translation start site, BglII site for cloning purposes underlined.</td>
</tr>
<tr>
<td>ACT1EX1BAM</td>
<td>5'ATCACGGTACCAACCATTCCAGT TCCATTTCGCAAAACAAGGTTG AATGTCTACCACCACGCTTTTC TTCTAC 3'</td>
<td>-10 to +57 from the Act1 translational start site. BamHI site for cloning purposes underlined.</td>
</tr>
<tr>
<td>1REP2LIA</td>
<td>5'TCAGCCATTTTTATGAGCTTTATTGCAAATCCAAACA3'</td>
<td>ACT2 exon 1 +647 to +654 bp from ACT2 start site of transcription, and ACT1 first intron +561 to +580 from the ACT1 major start site of transcription.</td>
</tr>
<tr>
<td>----------</td>
<td>------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1REP2LIS</td>
<td>5'TTCCGCTCTTTCTTTCCAAGGTACTTTTGAGACCCTT 3'</td>
<td>ACT2 leader exon +97 to +116 bp from ACT2 start site of transcription, and ACT1 first intron</td>
</tr>
</tbody>
</table>
Table 2-2A  
Qualitative transgenic GUS phenotypes recorded in inflouresces and mature flowers

<table>
<thead>
<tr>
<th>CON-MIN</th>
<th>LI deletion</th>
<th>LI forward</th>
<th>LI reverse</th>
<th>LI ACT2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>Low/ n.o.</td>
<td>Low/ n.o.</td>
<td>Low/ n.o.</td>
<td>Medium</td>
</tr>
<tr>
<td>Low/ n.o.</td>
<td>Low/ n.o.</td>
<td>Low/ n.o.</td>
<td>Low/ n.o.</td>
<td>Medium</td>
</tr>
<tr>
<td>Low/ n.o.</td>
<td>Low/ n.o.</td>
<td>Low/ n.o.</td>
<td>Low/ n.o.</td>
<td>Medium</td>
</tr>
<tr>
<td>Low/ n.o.</td>
<td>Low/ n.o.</td>
<td>Low/ n.o.</td>
<td>Low/ n.o.</td>
<td>Medium</td>
</tr>
<tr>
<td>Low/ n.o.</td>
<td>Low/ n.o.</td>
<td>Low/ n.o.</td>
<td>Low/ n.o.</td>
<td>Medium</td>
</tr>
<tr>
<td>Low/ n.o.</td>
<td>Low/ n.o.</td>
<td>Low/ n.o.</td>
<td>Low/ n.o.</td>
<td>Medium</td>
</tr>
<tr>
<td>Low/ n.o.</td>
<td>Low/ n.o.</td>
<td>Low/ n.o.</td>
<td>Low/ n.o.</td>
<td>Medium</td>
</tr>
<tr>
<td>Low/ n.o.</td>
<td>Low/ n.o.</td>
<td>Low/ n.o.</td>
<td>Low/ n.o.</td>
<td>Medium</td>
</tr>
</tbody>
</table>

The following values are used: High, Medium, Low and n.o. for expression not observed. A dash through a box denotes that no value was taken.
Table 2-2B Qualitative transgenic GUS phenotypes recorded in 7-10 day old seedlings.

<table>
<thead>
<tr>
<th>-310bp</th>
<th>Root Tip</th>
<th>Lateral Root Primordia</th>
<th>Vascular Tissue</th>
<th>Cotyledons</th>
<th>Leaf Primordia</th>
<th>Stipules</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>High</td>
<td>High/ Medium</td>
<td>Medium</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>-238bp</td>
<td></td>
<td>High</td>
<td>High</td>
<td>Medium</td>
<td>High</td>
<td>high</td>
</tr>
<tr>
<td>-88bp</td>
<td></td>
<td>None</td>
<td>None</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>-41bp</td>
<td></td>
<td>n.o.</td>
<td>n.o.</td>
<td>Low</td>
<td>n.o.</td>
<td>Low</td>
</tr>
<tr>
<td>-32bp</td>
<td></td>
<td>n.o.</td>
<td>n.o.</td>
<td>Low</td>
<td>n.o.</td>
<td>n.o.</td>
</tr>
<tr>
<td>-87–78M</td>
<td></td>
<td>High</td>
<td>High</td>
<td>High/ Medium</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>-77–68M</td>
<td></td>
<td>High</td>
<td>Medium</td>
<td>Low</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>-67–58M</td>
<td></td>
<td>High</td>
<td>Medium</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>-57–48M</td>
<td></td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>-47–38M</td>
<td></td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>LI deletion</td>
<td>n.o.</td>
<td>n.o.</td>
<td>n.o.</td>
<td>n.o.</td>
<td>n.o.</td>
<td>n.o.</td>
</tr>
<tr>
<td>LI forward</td>
<td>n.o.</td>
<td>n.o.</td>
<td>n.o.</td>
<td>n.o.</td>
<td>n.o.</td>
<td>n.o.</td>
</tr>
<tr>
<td>LI reverse</td>
<td>n.o.</td>
<td>n.o.</td>
<td>Medium</td>
<td>n.o.</td>
<td>n.o.</td>
<td>n.o.</td>
</tr>
<tr>
<td>LI ACT2</td>
<td></td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
</tbody>
</table>

The following values were used: High, Medium, Low, n.o. for not observed.
Figure 2-1. Schematic Diagrams of Chimeric ACTI-GUS Constructs

A. ACTI promoter truncations fused to GUS reporter gene. Construct names are at the left. The number of bases truncated from the major start site of transcription are shown at the immediate left of the construct diagram. The 55 base pair region of homology is hatched. Features shown are the major start site of transcription (arrow), the ACTI intron L that intervenes in the 5’UTR, the ACTI translation start site (first ATG) and the bacterial GUS ATG codon that is in frame with the ACTI translational start site. Construct names indicate the number of base in of flanking sequence DNA prior to the major start of transcription (i.e., p310 has 310 bp of DNA in front of the start of transcription).

B. Linker-scanner mutant constructs altering the 55 bp conserved region. All construct names are shown to the left and the areas of mutagenesis are written in an expanded form. Ten bp regions were substituted with foreign DNA containing an XbaI site. Unchanged base pairs are shown in capital letters.

C. ACTI L intron rearrangement constructs and the ACT2 intron replacement construct. p310-ivs contains a precise deletion of the ACTI L intron. pLIB-F and pLIB-R add back the ACTI L intron in both forward and reverse directions, respectively, but upstream of the promoter. pLIR-BI contains a precise replacement of the ACTI L intron with the ACT2 L intron.
Figure 2-2. Constructs containing the *ACT1* promoter mutations fused to the GUS reporter gene and analyzed in transgenic plants

A. - E. GUS expression patterns of p310 construct in several transgenic *Arabidopsis* lines. This pattern also reflects that observed for p238 and all linker-scanner constructs.

A. Transgenic seedling at 7 –10 days old

B. Root tip from a 7 –10 day old seedling

C. Mature inflorescence

D. Dehiscent anther and mature pollen

E. Immature ovule

F. Mature inflorescence expressing p88 construct

G. Mature inflorescence expressing p41 construct

H. Root tip expressing p88 construct

I. Mature inflorescence expressing p32 construct

J. Mature inflorescence expressing intron deletion construct, p310-ivs

K. Root tip expressing *ACT2* intron replacement construct, pLIR.

L. Mature inflorescence expressing *ACT2* replacement construct, pLIR.
**Figure 2-3. GUS Quantification Assay**

A. Schematic of the quantitative GUS assay. Two anthers from transgenic or control plants were placed into each microtitre plate wells containing a buffered assay mix to which GUS substrate (MUG) is added. After various time intervals, the reaction was stopped and the fluorescence assayed. Anthers from p310 construct expressing line #1 and/or Inspeck Blue Beads as well as a control wild-type RLD line are included in each set of assays.

B. Results from the fluorescence assays were plotted and the resulting slopes for each line are determined. These rate values, a measure of β-glucuronidase activity, are then normalized to the level of activity in p310 line #1.

C.-E. Quantification of expression from various constructs. The average rate value for the full-length construct (p310) was assigned a 100% value that all other lines are standardized against. The rates for each line were then averaged. The standard errors depict the variation among lines.

C. Truncation constructs (see Figure 2-1A).

D. Linker scanner constructs (see Figure 2-1B).

E. Intron rearrangement constructs (see Figure 2-1C).
CHAPTER 3

REPRODUCTIVE CLASSES OF *ARABIDOPSIS THALIANA* ACTIN GENES GUIDE HOMOLOGOUS PATTERNS OF EXPRESSION IN EVOLUTIONARILY DISTANT PLANT SPECIES

---

1 A.V. Vitale, R. B. Meagher, R. Cheng, R. Wu, to be submitted to the *Journal of Molecular Evolution*
Abstract

Two actin gene classes, reproductive and vegetative, predate the separation of monocots and dicots. The conservation of the potentially ancient regulatory systems of two Arabidopsis thaliana reproductive actins was examined. The 5' flanking regions of ACT1 were fused in translational frame to a GUS reporter gene and transformed into the distantly related dicot Nicotiana tabacum and monocot Oryza sativa. The ACT11 5' flanking region translationally fused to the GUS reporter gene was transformed into N. tabacum. The ACT1 and ACT11 driven reporters were expressed in both pollen and ovules, but not organ primordia, of N. tabacum. ACT1 was strongly expressed in O. sativa pollen. These studies demonstrate that the cis-DNA regulatory elements and transacting factors controlling the expression patterns of the reproductive classes of actin genes are highly conserved in reproductive, but not vegetative organs.
Introduction

Many genes are expressed in a similar spatial and temporal fashion across families and even phyla within the plant and animal kingdoms. Within the animal kingdom homeobox genes such as the hox gene clusters, duplicated throughout metazoan phyla, are characterized examples of conservation of gene function and expression during development. For instance, *cis*-elements in mouse homeobox gene *Hoxb-1* direct early neural expression in the distant vertebrate species such as chicken and pufferfish (Marshall et al., 1994). This demonstrates that transcriptional machinery and regulatory pathways (in this instance retinoic acid sensing pathways) necessary to guide gene expression are conserved across taxonomic groups in vertebrate development (Marshall et al., 1994). Hox genes defining neural segmentation in *Drosophila melanogaster* are expressed similarly in the spider *Cupiennius salei*, demonstrating that head development in arthropods is of monophyletic origin (Damen et al., 2000).

A family of developmental transcription factors in plants is the MADS box genes. Within the plant kingdom the MADS box genes are transcription factors that act combinatorially to regulate the development of the reproductive structures of plants and, in the case of angiosperms, floral architecture. MADS transcription factors are a very ancient class of genes whose origin predates the divergence of plants and animals (Doyle, 1994). The separation of higher plant MADS box genes occurred over 300 million years
ago, well before the separation of angiosperm and gymnosperm lineages (Becker et al., 2000). Additionally, MADS box orthologs that are found in distantly related
gymnosperms have similar expression patterns as their angiosperm counterparts, i.e.
some MADS genes can be considered pollen specific (Munster et al., 1997). MADS box
genes, because they guide major aspects of plant development, possibly aided in the
division of vegetative and reproductive tissues and the divergence of gene expression in
these tissues (Theissen and Saedler, 1995). The conservation of MADS box genes and
their expression patterns in distantly related species suggest a floral “Bauplan” because
MADS genes control floral development. Floral homeotic genes diverged at distant time
points yet remain functionally homologous as well as equivalently expressed. It is
postulated that cytoskeletal genes might also reflect patterns of expression necessary for
land plant evolution (Meagher et al., 1999). *A. thaliana* contains eight expressed actin
genes that can be classified into two major classes, reproductive and vegetative, based on
expression patterns and sequence divergence. The reproductive class of actin genes can
be further divided into three subclasses containing *ACT1* and *ACT3, ACT4* and *ACT12*,
and *ACT11*, respectively (McDowell et al., 1996). Each subclass of the reproductive
actins displays a unique expression pattern. *ACT1* and *ACT3*, differing by only one
conservative amino acid change and more than a hundred silent nucleotide changes, are
highly expressed in mature pollen, immature ovules, floral and vegetative organ
primordia, and root tips (An et al., 1996). *ACT4* and *ACT12* are estimated to have
diverged from $ACT1$ and $ACT3$ more than 150 million years ago (McDowell et al., 1996). They also differ by only a single amino acid and more than a hundred silent nucleotide changes, but are expressed much more weakly in mature pollen and organ primordia (Huang et al., 1996a). These four actins, $ACT1$, 3, 4, and 12 have been termed the late pollen actins. Based on mapped immuno-epitopes they have been a separate and conserved group from the rest of plant actins for at least 220 million years (Kandasamy et al., 1999). In contrast $ACT11$, the representative of the third class, is the only actin gene to be expressed significantly in mature ovules and developing endosperm and carpel wall (Huang et al., 1997). All five reproductive actin genes are highly expressed in mature pollen, though the steady state level of $ACT1$ mRNA is ~100-fold greater than any other reproductive actin gene in this cell type (An et al., 1996). Because $ACT1$ is so abundantly expressed in mature pollen and $ACT11$ has a singular and significant expression pattern from embryogenesis through fruit development, these two genes were chosen to search for conserved regulatory mechanisms in reproductive tissue throughout monocot and dicot evolution.

Macroevolution of plant reproductive and vegetative structures appears to correlate with the separation of vegetative and reproductive classes of cytoskeletal genes (Meagher et al., 1999). The expression patterns of various cytoskeleton genes have been well-documented (Meagher et al., 1999). Profilins are easily divided into reproductively
and vegetatively expressed gene classes within both *Arabidopsis* (Huang et al., 1996b) and maize (Staiger et al., 1993). Other cytoskeletal genes that function in tandem with actin (such as actin depolymerizing factor, ADF) can be clearly divided into vegetative and reproductive classes based on correlating sequence and expression data in studies of *Arabidopsis* and maize (Jiang et al., 1997; Lopez et al., 1996). The cytoskeletal proteins critical for cell growth and development, α-tubulin and β-tubulin, also display divergence in expression patterns in monocot and dicot systems similar to actin gene class expression (Carpenter et al., 1993; Carpenter et al., 1992; Kopczak et al., 1992; Snustad et al., 1992). Reproductive expression patterns of cytoskeleton genes may be based on a broad underlying scheme common to different plant evolutionary lineages.

**Materials and Methods**

*Strains and Constructs:* Plasmids used to examine *ACT1* heterologous expression are described in Chapter 2 and shown in Figure 3-1. All of the translational fusions of the *ACT1* and *ACT11* genes included the native AUG start codon, the first 19 codon positions and the GUS gene containing its own AUG start codon. *ACT11* expression was determined using the 1.5 kb 5' flanking region fused to a GUS reporter gene in a plant binary vector (pAAc11) (Huang et al., 1997). The following plasmids and their neomycin phosphotransferase resistance markers were used: p2200 (pAAc1E1) (An et al., 1996), p310, p238, p88, p310-ivs (Chapter 2) and p1500 (pAAc11) (Huang et al., 1997). The
construct p2200 contains 2.2 kb of ACT1 5’ flanking region upstream of the major start site of transcription. The constructs p310, p238 and p88 contain 5’ deletions of p2200, retaining 310 bps, 238 bps, and 88 bps upstream of the major start site of transcription, respectively. The construct p1500 contains 1.5 kb of ACT1 5’ flanking region translationally fused to the GUS gene.

Plant transformation: N. tabacum was transformed with all of the vectors described above using a modification of the Agrobacterium mediated leaf disk methods of Horsch et al. (Goring et al., 1991; Horsch et al., 1988). O. sativa var. TN67 was transformed with a construct containing the full-length ACT1 GUS fusion from p2200 cloned into pCAMBIA (Hajdukiewicz, 1994), a hygromycin resistance vector. The vector was bombarded into rice and the selection and regeneration steps were carried out on hygromycin containing media (Cao et al., 1992; Hiei et al., 1994).

Qualitative expression assays: β-glucuronidase staining was carried out in both tobacco and rice essentially according to established methods (Jefferson et al., 1987). All tissue was first dissected by hand and stained in a 1 mM X-Gluc solution for 24 hours. The tissues were then stored in 70% ethanol. Sectioned ovule tissue, anthers and pollen grains were further bleached in a saturated solution of calcium hypochlorite for five to ten minutes until the tissue was lightened enough to observe any blue staining from the GUS cleavage product.
Results

Constructs containing the full length ACT1 and ACT11 promoters fused to the GUS reporter were examined in both tobacco and rice. The ACT1 reporter fusions in tobacco were examined in the greatest detail and are presented first.

*ACT1 5' flanking region shows conserved expression in tobacco.*

The construct p2200 contains 2.7 kb of 5' flanking region upstream of the ACT1 translation start site and 19 codons fused in frame to the GUS reporter and is referred to as the full length construct. In *A. thaliana* p2200 is strongly expressed in mature pollen and organ primordia (An et al., 1996), and weaker expression is seen in immature ovules. Analysis of transgenic GUS expression of p2200 in tobacco demonstrated full conservation of the native expression pattern in reproductive tissues. In all seven lines examined the ACT1 flanking region in p2200 strong reporter expression in mature pollen was observed (Figure 3-2A), but weak or no reporter expression in immature pollen. Strong expression was also seen in stylar tissue (Figure 3-2B). Interestingly, reporter expression in ovules was not confined to pre-fertilization stages as it was in *Arabidopsis*, but was maintained strongly throughout development (Figure 3-2C). Only one line examined demonstrated strong expression in tobacco seedlings and the expression in organ primordia observed for p2200 in *Arabidopsis* was not seen in tobacco (data not shown).
To determine sequences within the \textit{ACT1} 5' flanking region that may have evolutionary importance in guiding actin expression in reproductive tissues, other constructs containing deletions of the conserved region within the \textit{ACT1} promoter were examined in transgenic tobacco (Figure 3-1A). A deletion construct, p310, containing only -310 bp upstream of the major start site of transcription, was analyzed. To delineate the smallest 5' flanking region that would guide reporter gene expression, three other promoter truncations were examined in transgenic tobacco. The first construct, p238, contained sequences 238 bp upstream from the major start site of transcription and thus was 72 base pairs shorter than the base construct of p310 (Figure 3-1A). The second construct, p88, contained 88 base pairs from the major start site of transcription and included only the 55 base pair conserved region directly upstream of the TATA region and the rest of the 5' flanking region downstream of the TATA (Figure 3-1A). The third construct, p310-ivs, contained the full length -310 bp like p310, but with a precise internal deletion of the intervening sequence (intron) normally found within the 5'UTR (Figure 3-1B). p310--ivs was examined to determine any effect the intron may have on \textit{ACT1} expression.

The p310 construct demonstrated identical expression as that of the full-length p2200 construction in two lines examined (Figure 3-2D and Figure 3-2E). The p238 construct was expressed highly in mature pollen and more weakly in immature pollen.
Low levels of expression were seen in transmittal tissue, but no expression was seen in ovules (Figure 3-2F). The p88 construct exhibited little expression in mature pollen, no expression in non-dehiscent pollen, weak expression in the transmittal tissue and weak expression in ovules (data not shown). The intron deletion construct p310-ivs demonstrated high levels of GUS expression in mature dehiscent pollen and lower levels in immature pollen. Expression in the stylar transmittal track was very low and ovule expression was non-existent, in sharp contrast to the expression observed for p310 in these tissues.

*ACT11-GUS fusion with the full length 5’ flanking region is expressed similarly in N. tabacum and A. thaliana*

A second gene construct p1500, that of reproductively expressed *A. thaliana* *ACT11*, was assayed in transgenic tobacco. All seven tobacco lines examined demonstrated that the *ACT11* 5’ flanking region was able to guide expression in tobacco in a similar fashion as in *A. thaliana*. High levels of expression were seen in mature pollen, transmittal tissue and developing and mature ovules (Figure 3-2). Thus the expression of this reproductive actin subclass could have been conserved between tobacco and *Arabidopsis*, two dicots separated by approximately 100 million years.

*ACT1 full length construct p2200 guides high levels of expression in O. sativa*
The *ACT1* construct p2200 containing the full-length 5' flanking region directed expression in *O. sativa* that was very similar to that seen in *A. thaliana* (Fig. 2). Mature pollen stained strongly in ten independent transgenic rice lines that were examined. Ovule staining was not detected or was very weak, and in those lines exhibiting low levels of staining the GUS was diffuse throughout all female reproductive tissues and exhibiting no tissue specific staining pattern. Unlike *N. tabacum*, p2200-driven GUS staining was observed in root and young vegetative tissues in three out of four *O. sativa* lines examined. These lines showed high levels of staining in cells closer to the ground meristem of roots and also lateral root organ primordia, similar to that observed in *Arabidopsis*. Lateral roots demonstrated the same staining pattern as seen in the tap root. No staining was observed within the root cap or at the very tips of lateral roots. Thus the *A. thaliana ACT1* promoter may interact fully with conserved transcriptional machinery in a distantly related monocot.

**Discussion**

It has been postulated that cytoskeletal genes and in particular the plant actin gene family underwent ancient divergence in co-evolving with the separation of vegetative and reproductive structures (Meagher et al., 1999). Orthologs of actin genes throughout the higher plants share documented common expression patterns, though they are separated by hundreds of million years (McDowell et al., 1996). Other cytoskeletal genes such as
α-tubulins, β-tubulins, profilins, and ADFs that exist as diverse gene families also show
divergent gene expression in reproductive and vegetative structures (Meagher et al.,
1999). These diverse cytoskeletal genes may have co-evolved as a group, thus allowing
specialized interactions in reproductive versus vegetative tissues (Meagher et al., 1999).

*Arabidopsis* has a comparable number of gene families as other eukaryotes, but often has
a higher number of gene copies within these families than animals (The Arabidopsis
Genome Initiative, 2000). *A. thaliana* is likely descended from a tetraploid ancestor that
arose 112 MYA and after the divergence of the *Arabidopsis* and tomato plant lineages
(The Arabidopsis Genome Initiative, 2000). Though functional redundancy has been
shown for gross phenotypes of many of these duplicated genes, these studies may not
carefully address the necessity for a gene product in a specific spatial and temporal
location. For instance, separate knockouts of ACT2 and ACT7 produce no discernable
above ground phenotype in soil-grown plants (Gilliland et al., 1998), but population
analyses of both mutant alleles demonstrates that they are rapidly eliminated from the
population (Asmussen et al., 1998). Thus, selection pressure on actins may be due to
function or the regulation of actin protein products.

Other genes not directly participating in cytoskeletal phenomena also display
separate expression patterns in vegetative and reproductive tissues. Prominant are the
sugar modifying genes of the invertase and sucrose synthase gene families. Maize
invertase 1 (Ivr1) genes, invertase 2 (Ivr2) genes and maize sucrose synthase gene families of shrunken 1 (sh1) and sucrose synthase 1 (sus1) are differentially regulated. Ivr1 and sh1 are expressed largely in reproductive tissues and root tips, where sugar depletion is common. Ivr2 and sus1 are expressed over vegetative regions of the plant where sugar depletion is less likely to occur (Zeng et al., 1998). *A. thaliana* also has a large invertase gene family displaying a range of expression patterns (Tymowska-Lalanne and Kreis, 1998). Two of the four invertase genes studied were expressed in mainly vegetative tissues, while another was expressed in rapidly dividing tissue. Yet a fourth was found predominantly in floral tissues, reflecting the vegetative and reproductive split more commonly seen in many plant gene families, not only cytoskeletal genes.

Our work directly supports the hypothesis that families of cytoskeletal genes such as actin were selected over evolutionary time because their expression was necessary for the separate formation and development of reproductive and vegetative organs. Not only was it shown that the transcriptional machinery for cytoskeletal expression has been conserved over evolutionary time, but heterologous expression can be used as a test to determine what parts of a promoter may have been truly important in guiding gene expression during land plant evolution.
In this work we tested two *A. thaliana* actin gene promoters known to drive high levels of expression in reproductive organs: *ACT1* in mature pollen and *ACT11* in pollen and developing and mature ovules. *ACT11* full-length promoter fused to a GUS reporter gene, p1500, drove expression in a nearly identical fashion in tobacco as in *A. thaliana*. The full-length *ACT1* promoter p2200 and truncated promoter derivatives p310, p238, and p88 also drove expression in tobacco as they did in *A. thaliana*. p2200 and p310 truncation both directed GUS expression in mature pollen and throughout ovule development. Further truncations such as p88 failed to express in tobacco, suggesting that upstream elements of the promoter are essential to stabilize expression. Interestingly, the p310-ivs construct, containing -310 bps upstream from the major start site of transcription and deletion of the leader intron, guided high levels of mature pollen, but failed to guide expression in developing or mature ovules in tobacco, which was exactly as seen in *A. thaliana*. Thus, it is clear that regulatory elements within the first intron of *ACT1* are needed to direct ovule expression in these two distant organisms. The p2200 construct of *ACT1* directed high levels of expression in mature pollen of *O. sativa*, but failed to display high levels of GUS staining in developing or mature ovules, unlike the expression phenotypes of tobacco and *A. thaliana*.

These results illustrate the likely conservation of transactivating factors across major plant classes, with respect to actin gene promoters guiding reproductive
expression. *ACT1* and *ACT11*, though expressed in reproductive tissues, had specific expression patterns in their native *A. thaliana* background. In tobacco their expression was within reproductive organs, but the specific expression pattern seen appeared more generalized as both the full-length *ACT1* and *ACT11 5' flanking regions* were highly expressed in both mature pollen and developing ovules. The p2200 expression pattern in tobacco was slightly expanded in reproductive tissues as compared to expression in *A. thaliana*. High levels of expression throughout ovule development were seen in tobacco, while in *A. thaliana* expression was only seen during early ovule development. The presence of the *ACT1* intron seems essential for ovule expression in both dicots. Other *ACT1* deletion constructs showed little to no expression in tobacco, but high visible expression in *A. thaliana*. This was most likely because the *ACT1 5' flanking region* contains multiple tissue specific regulatory elements. When the promoter was expressed in a heterologous system the absence of one or more of these elements may have significantly lowered the amount of visible GUS expression. For instance in *A. thaliana* the -88 construct displayed visible GUS expression, but had a 20-fold reduction in expression levels. Therefore it is feasible that any expression guided by some of the *ACT1* constructs was not above the threshold of detection. *ACT1* (p2200) expression in *O. sativa* was remarkable in that high levels of pollen specific expression were preserved, though ovule expression was not. This suggests that the presence of the *ACT1* gene product in female gametophytic tissue may not have been selected during land plant
evolution. Overall, using heterologous organisms prediction of important regulatory regions. The strongest area of expression in *ACT1* (pollen) and *ACT11* (pollen and developing ovules) were conserved among all species examined. Assay sensitivity could be a problem when reporter gene expression levels in a particular assay may have been below the threshold of detection.

In evolutionary biology it has been long supposed that multi-gene families become stable over time because duplication of a gene permits one copy to be affected by mutations in the amino acid sequence, thus allowing for novel gene function leading to greater organismal fitness (Nadeau and Sankoff, 1997; Ohta and Kimura, 1970; Sidow, 1996). This hypothesis is challenged with an alternative explanation for the complexity of eukaryotic gene expression known as the duplication-degeneration-complementation (DDC) model (Force et al., 1999). The DDC model asserts that for genes under the control of multiple regulatory elements, mutations in regulatory regions of duplicated genes can help rather than hinder their fixation. The model stipulates that rather than duplication leading to novel gene function, regulatory mutation usually leads to partitioning of existing gene function, thus selecting for the preservation of duplicated genes. The DDC model also allows for occasional acquisition of novel gene function. As actin genes have been demonstrated to be monophyletic in origin (Hightower and Meagher, 1986; McDowell et al., 1996), an ancestral actin gene undergoing duplication
could both partition essential regulatory elements plus create novel elements necessary for reproductive tissue expression. As more gene family members duplicated, partitioning and/or adding of novel regulatory regions would lead to the preservation of individual members of a gene family, where every expressed member was required to maintain the fitness of the organism. With the examples above of reproductive actin gene expression in distant plant species it is likely that both the strong similarities and slight discrepancies shown for the specific constructs reflect this partitioning of duplicated regulatory regions throughout the evolution of the actin gene family.


Figure 3-1. Schematic Diagrams of Chimeric *ACTI*-GUS and *ACTII*-GUS

**Constructs**

**A.** *ACTI* truncation constructs fused to a GUS reporter gene. The number of bases remaining in front of the major start site of transcription are shown at the immediate left of the construct diagram along with the construct name. The 55 base pair region of homology with *ACT3* is crosshatched. Features showcased are the major start site of transcription (arrow), the *ACTI* intron L within the 5’UTR, the *ACTI* translation start site (first ATG) and the bacterial GUS ATG start codon in the same frame. The full-length 2200 bp promoter *ACTI* construct p2200 was examined in both *N. tabacum* and *O. sativa*, while the remaining shorter constructs were examined solely in *N. tabacum*.

**B.** *ACTI* L intron deletion construct. Based on the p310 construct, the intron was perfectly removed from the *ACTI* 5’ flanking region to create p310-ivs.

**C.** Full-length *ACTII*-GUS construct. Shown is the *ACTII* major transcription start site, intron L, and both the translational start sites for *ACTII* and GUS in construct p1500.
Figure 3-2. Expression analysis of *A. thaliana* reproductive class of actin in heterologous species.

A.- C. Expression patterns of 2.2kb *ACT1*-GUS construct.

A. Strong mature pollen expression in p2200.

B. Strong expression in stigma in p2200.

C. Strong expression in developing ovules in p2200.


D. Strong mature pollen expression in –310.

E. Medium expression in the stigma and stylar tissue in -310.

F. No expression in developing ovules is present in -238.

G. – H. Expression of full length *ACT1* promoter construct p1500 in tobacco.

G. Strong mature pollen expression.

H. Strong expression in stigma and stylar tissue.

I. Strong expression in developing ovules.

J. – L. Expression of p2200 with full length *ACT1* promoter in rice.

J. Strong expression of pollen within the anther sac.

K. Mature pollen expression.

L. Strong expression in root vascular tissue and root tip cells.
This dissertation demonstrates the use of two methodologies to examine both the likelihood and importance of conserved promoter function in essential cytoskeletal genes. One methodology focused on the used of evolutionarily distant plant species as a measurement of conserved cis-elements and trans-factors. \textit{ACT1}, a highly expressed \textit{A. thaliana} actin pollen-specific promoter was introduced as a 5' flanking region reporter gene construct into transgenic \textit{N. tabacum} and \textit{O. sativa}. The expression pattern in both heterologous species suggested that although \textit{ACT1} is also highly expressed in immature ovules and rapidly dividing tissue such as organ primordia these organs may have conserved transactivating factors, thus over evolutionary time the \textit{ACT1} protein was most likely not necessary for plant viability in these areas. Another possibility is that a number of activation domains are needed for the expression seen in \textit{A. thaliana} and in heterologous species may be activated but simply not visible. Other \textit{ACT1} promoter derivative constructs were tested in transgenic \textit{N. tabacum}. These constructs demonstrated that a small area of the 5' flanking region, 310 base pairs upstream of the start site of transcription, and other areas including a noncoding exon with intervening intron and the first 19 amino acid codon positions. 5' flanking regions containing larger truncations lost most or all expression, while a deletion of the leader intron retained expression in tobacco mature pollen. Thus the \textit{ACT1} promoter as found in \textit{A. thaliana} is capable of guiding high levels of expression in mature pollen in both a distant dicot and
Therefore transcriptional machinery has been conserved throughout land plant evolution that may have served to guide the evolution of the plant cytoskeleton.

The second approach that was used to determine elements of the *ACT1* 5' flanking region that may have been important for *ACT1* expression during land plant evolution was to undergo a rigorous series of mutations on the region, looking for loss-of-expression phenotypes in reporter constructs. Although the *ACT1* and closely related *ACT3* gene differ by only one conservative amino acid sequence and are estimated to be diverged by only 30 – 60 million years the noncoding 5' flanking area shows no conservation of sequence. The one exception is a 55 base pair area directly 5' to the TATA region of both genes. Because *ACT1* and *ACT3* share identical expression patterns, the conserved promoter area was first focused on as possibly responsible for the large level of expression seen in mature pollen. A series of 5' truncations and site-directed mutagenesis constructs of the 55 base pair region were introduced into transgenic *A. thaliana*. Other constructs involving rearrangements and a deletion of the *ACT1* leader intron were also examined. Qualitative analysis showed that a 5' deletion construct containing only the 55 base pair region along with the remaining flanking region was sufficient to guide mature pollen expression. Closer examination of this construct using a quantitative measurement of GUS expression in mature pollen showed that in comparison with a control full length promoter construct with level of expression was greatly reduced. Thus, though the construct bearing the conserved area is sufficient
for mature pollen expression areas upstream on this region are needed for the high levels noted in the control construct.

Future directions regarding understanding and prediction of reproductive actin and cytoskeleton expression will involve isolation of transativating factors and tying these factors into both known and novel signal transduction cascades. It will also be of extreme interest to note how actin and actin binding protein proteins are involved in a feedback scenario, whereby signals from an active cytoskeleton responding to environmental cues signal multiple cytoskeleton signal cascades.