ANALYSIS OF THE *ARABIDOPSIS* NAC GENE SUPERFAMILY IN PLANT DEVELOPMENT

A Dissertation

by

VERIA ALVARADO CHAVEZ

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2007

Major Subject: Molecular and Environmental Plant Sciences

ANALYSIS OF THE *ARABIDOPSIS* NAC GENE SUPERFAMILY IN PLANT DEVELOPMENT

A Dissertation

by

VERIA ALVARADO CHAVEZ

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Terry Thomas
Committee Members,	Tom Mcknight
	Alan Pepper
	Dorothy Shippen
Chair of Interdisciplinary Faculty,	Jean Gould

December 2007

Major Subject: Molecular and Environmental Plant Sciences

ABSTRACT

Analysis of the *Arabidopsis* NAC Gene Superfamily
in Plant Development. (December 2007)
Veria Alvarado Chavez, B.S. UNALM, Peru;
M.S., University of California Davis
Chair of Advisory Committee: Dr. Terry L. Thomas

There are a vast number of transcription factors that regulate plant growth and development. The NAC gene superfamily is one of the largest families of transcription factors in the plant kingdom. NAC gene expression profiles using Affymetrix ATH1 gene chips were obtained for different plant organs: heart embryo, mature embryo, leaf, root and flower. NAC gene expression profiles proved to be very complex, except for one NAC gene detected only in floral tissue, *At1g61110*. *At1g61110* was shown to be specifically expressed in the anther tapetum of A*rabidospis*; therefore, its name was changed to *TAPNAC*. *TAPNAC* became the focus of our studies.

We identified a *tapnac* T-DNA knockout (KO) line, SALK_069450. A molecular phenotype was observed. Several oligopeptide, sugar and metal transporters were differentially expressed. Coincidentally, a wheat NAC gene, named *TaNAM-B1* for its high sequence similarity to *ATNAM*, *TAPNAC* and *At3g15510* was found to be involved in nutrient remobilization. *PHOSPHOLIPASE Dα1* (*PLDα1*) was also found to be down-regulated in the *tapnac* KO. PLDα1 is an enzyme which hydrolyzes phospholipids that are part of tapetal cell membranes and tapetal lipid bodies. Once these tapetal cell structures are disrupted, the

secretion of the compounds that form part of the pollen coat (*i.e*. proteins, flavonoids and lipids) into the anther locule is facilitated.

Promoter deletion analysis using a GUS reporter and later GUS immuno-localization confirmed the findings of Wellmer and others. *TAPNAC* is a tapetal specific gene. The *cis*-regulatory sequence that enhances tapetal expression in the *TAPNAC* promoter was identified. The consensus motif TCGTGT increased tapetal expression of a GUS reporter gene, only when flanked by the *TAPNAC* minimal promoter region (-217 bp to +51 bp).

In summary, TAPNAC transcription factor has been characterized and data indicates that it could play a role in nutrient remobilization from the tapetum to the pollen grains, particularly during late floral stages. Also, important information on tapetal specifcation *cis*-regulatory sequences was discovered. The consensus motif TCGTGT, present in TAPNAC promoter, was shown to enhance tapetal expression of a GUS reporter gene.

ACKNOWLEDGMENTS

I would like to acknowledge all the people that have been part of my life as a graduate student at TAMU and without whom the completion of this dissertation would not have been possible. First, I would like to thank Dr. Terry Thomas for his endless support, encouragement and patience, for allowing me to be independent in my work and letting me grow as a scientist. I also would like to thank my other committee members for helpful comments and suggestions: Dr. Dorothy Shippen, Dr. Tom McKnight and Dr. Alan Pepper.

I would like to express my sincere gratitude to all the lab members, former and present: Dr. Mona Damaj, for her kindness and mentoring, Jung In Hur, for her friendship and active discussion on my research, Dr. Andy Tag for his companionship and endless help to any imaginable problem; Dr. Phil Beremand, for his expertise in microarray analysis and insightful discussions, Yichun Yang for her sweet comments on my research, and Kelly Bungard and Rami Weaver for their help with lab work and experiments. Also, I would like to acknowledge the companionship we share with members of Dr. McKnight's lab, in particular Dr. Ketan Patel and Dr. Shuxin Ren.

Last, but not least, I would like to thank my family for their infinite support, patience, and love. Without them I would not have been able to accomplish this personal goal.

TABLE OF CONTENTS

CHAPTER Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

CHAPTER I

INTRODUCTION: NAC TRANSCRIPTION FACTORS AND THEIR ROLE IN PLANT GROWTH AND DEVELOPMENT

Transcription factors are DNA binding proteins that regulate gene expression either inducing or repressing RNA polymerase activity. There are a vast number of transcription factors that regulate plant growth and development. NAC transcription factors along with the MYB, AP2/EREBP and bHLH proteins are the largest families of transcription factors in the plant kingdom (Riechmann et al., 2000).

NAC transcription factors were first identified in petunia embryos that carried a mutation in *no apical meristem* (*nam*) that disrupted the formation of the shoot apical meristem (Souer et al., 1996). The DNA binding domain is called the NAC domain based on the three founder genes: *NAM, ATAF1-2 and CUP SHAPED COTYLEDON 2* (*CUC2*) (Aida et al., 1997). The NAC gene superfamily contains 109 gene members. NAC proteins often display functional redundancy, as is the case of *CUP SHAPED COTYLEDON 1-2*, and their roles in plant growth and development are quite diverse. However, there are still several NAC genes with unknown functions. This dissertation focused on the functional analysis of a floral specific NAC gene.

This dissertation follows the style of Plant Cell.

Transcription factors

Transcription factors are proteins that bind specific sequences of DNA (*cis*-regulatory sequences) and are capable of activating and/or repressing transcription (Riechmann et al., 2000). Understanding the structure and DNA-binding properties of these proteins has helped elucidating how genetic information is utilized for regulation of cell development, differentiation and cell growth. Transcription factors have the characteristic of being modular proteins, they contain a DNA binding domain and a transactivation domain.

Based on their DNA binding domain, transcription factors have been classified into several classes including basic helix-loop-helix, zinc finger, leucine zipper or high mobility group (Pabo and Sauer, 1992). The basic helix-loop-helix structure contains two amphipathic α-helices with highly conserved basic residues in the amino-terminal side and several hydrophobic residues in the carboxy-terminal end. The helices are linked by amino acid sequences of variable length, which form reverse turns and loops and the entire motif mediates homo- and heterodimerization, which favors DNA-binding through the basic domains. Zinc finger transcription factors must recruit zinc in order to bind to DNA. There are two zinc finger motifs that have been identified. The first consists of 30 amino acids, including two cysteine-histidine pairings that coordinate tetrahedral binding to a single zinc atom. The second zinc finger motif displays a partnership between cysteine-cysteine residues to direct zinc chelation. The leucine zipper transcription factors bind DNA as dimers. A leucine zipper is formed by two α-helices, one from each monomer. The helices are held together by hydrophobic interactions between leucine residues, which are located on one side of each helix. The high mobility group box defines a class of transcription factors. It binds to a 20 bp span of DNA and distorts DNA structure. This motif is also a feature

of many structural and non-chromosomal proteins in the nucleus and can mediate bending, wrapping, spacing and coiling of DNA.

The transactivation domain is involved in organizing additional proteins involved in activating transcription. Transactivation domains may act directly or they may recruit coactivator proteins that possess activation properties and an ability to interact with the basal transcription machinery.

The *Arabidopsis* genome encodes at least 1500 transcriptional regulators, corresponding to ~6.0% of its estimated total number of genes. These regulatory proteins belong to different classes or families. The three largest families of transcription factors in *Arabidopsis* are AP2/EREBP (APETALA2/ethylene responsive element binding protein), MYB-(R1)R2R3, and bHLH (basic helix-loop-helix). There are many transcription factor families that are present only in plants. These include the AP2/EREBP, NAC (Aida et al., 1997), and WRKY families; also the trihelix DNA binding proteins; the auxin response factors (ARFs); the Aux/IAA proteins (which do not bind to DNA by themselves, but interact with the ARF proteins); and other smaller families (Riechmann et al., 2000).

NAC domain genes

NAC genes constitute one of the largest families of plant-specific transcription factors. This gene family is present in a wide range of plant species such us *Arabidopsis thaliana* (Aida et al., 1997), pumpkin (*Cucurbita maxima*) (Ruiz-Medrano et al., 1999); potato (Collinge and Boller, 2001); rice (Kikuchi et al., 2000); and wheat (Uauy et al., 2006). Most AtNAC genes contain three exons, the first two exons encode the NAC DNA binding domain and the last exon

encodes an activation domain. A group of NAC genes contain additional terminal exons that may encode protein domains that expand the range of biological functions of the NAC domain superfamily (Duval et al., 2002).

The recent determination of the DNA and protein binding NAC domain structure offers insight into the molecular functions of the protein family (Olsen et al., 2005). The structure of the NAC proteins display a unique transcription factor fold consisting of a twisted β-sheet surrounded by a few helical elements (Ernst et al., 2004; Olsen et al., 2004). A number of proteins use β-sheet structures for DNA binding (Pabo and Sauer, 1992; Luscombe et al., 2000), therefore it is not unexpected to see this type of protein fold for DNA-protein interactions. NAC domains are highly conserved among members of the gene family; it is thought that differences in the NAC domain likely affect the DNA-binding specificity for recognition of the *cis*-regulatory elements of the target genes, i.e. *CUC* genes (Taoka et al., 2004).

In *Arabidopsis*, there are 109 NAC genes identified through computational analysis. One study pointed out the similarity between NAC genes in the monocot *Oryza sativa* (rice) and the NAC genes in the crucifera *Arabidopsis thaliana* subdividing them based on the protein motifs. NAC genes were classified into two groups (Group I and Group II) and 18 subgroups, with the idea that NAC genes that belong to the same subgroup might have similar functions (Ooka et al., 2003).

Functions of characterized NAC genes

The roles of NAC genes in plant growth and development are diverse. The pioneering studies on *NO APICAL MERISTEM (NAM)* in petunia (Souer et al.,

1996) and the *CUC1-2* genes in *Arabidopsis* (Aida et al., 1997), revealed that mutations in these genes affect shoot apical meristem formation. *CUC1* regulates SAM formation by acting on the *SHOOT MERISTEMLESS* (*STM*) gene (Aida et al., 1999; Takada et al., 2001). Furthermore, when *CUC1* is ectopically expressed using a *35S* CaMV constitutive promoter, it can induce the formation of adventitious shoots (Hibara et al., 2003). Later, a third *CUC* related gene named *CUC3* was found. This gene is important for the establishment of boundaries that contain cells with low proliferation and/ or differentiation rates, *i.e* meristematic cells (Vroemen et al., 2003).

NAC genes have also been discovered to play roles in secondary wall thickening and vascular tissue formation. Mutations in *NAC SECONDARY WALL THICKENING PROMOTING FACTORS: NST1* (*At2g46770*) and *NST2* (*At3g61910*), result in loss of secondary wall thickening in anther endothecium. These NAC genes are functionally redundant for the process of anther dehiscence (pollen release) (Mitsuda et al., 2005). The *VASCULAR-RELATED NAC-DOMAIN* gene subfamily includes 7 *NAC* genes, of which *VND6* (*At5g62380*) and *VND7* (*At1g71930*) specifically inhibit metaxylem- and protoxylem-like vessel elements, respectively. These genes were identified in both *Arabidopsis* and poplar (Kubo et al., 2005). *NAC* genes involved in vascular tissue formation have been identified in other plant species, such as rice. Specifically, *OsNAC300* have been detected in developing and mature vascular tissues of the leaves and roots in rice (Kusano et al., 2005).

Specific NAC genes are not only expressed in the phloem or xylem, but they can also be transported and accumulated in vegetative, root and floral meristems. The initial discovery was made by Ruiz Medrano, who studied a *Cucurbita maxima* NAC gene, named *CmNACP*. This mRNA moved from a pumpkin plant

into a grafted cucumber plant, accumulating in the phloem and apical tissues (Ruiz-Medrano et al., 1999).

A number of NAC genes are involved in stress responses. The *ATAF* genes in the NAC gene superfamily are induced by wounding, and a gene in S*olanum tuberosum*, *StNAC*, was found to have similarity in structure and function to the *ATAF*s genes (Collinge and Boller, 2001). NAC genes are also involved in drought stress responses, as is the case of the NAC genes: *At1g52890*, *At3g15500* and *At4g27410* (*RD26*) (Taji et al., 1999). Specifically, RD26 can activate ABA-inducible gene expression under abiotic stress in *Arabidopsis* plants (Fujita et al., 2004). Furthermore, eight *Brassica napus* NAC genes were induced during abiotic (mechanical wounding, cold shock, and dehydration) and biotic stress (flea beetle feeding and *Sclerotinia sclerotorium* infection) in leaves of *Brassica napus* (Hegedus et al., 2003).

Finally, a recent discovery in the long list of NAC gene functions was done on *TaNAMB1*, a wheat NAC gene, which plays a role in mobilization of nutrients from leaves to the developing wheat grains (Uauy et al., 2006). This new role is associated with the role of NAC genes in processes of senescence.

Previously, we have been studying the NAC gene superfamily, focusing on one gene member *ATNAM* (*At1g52880*). This NAC transcription factor was isolated from an *Arabidopsis* developing seed library, and it was molecularly characterized, using DNA binding and transactivation assays. The activation domain was mapped to the protein C terminus and the DNA binding domain was mapped to the protein N terminus (Duval et al., 2002).

Using genomic approaches combined with bioinformatic tools, a NAC gene with high sequence similarity to *ATNAM* was identified. *At1g61110*, was found to be

floral specific, and later it was discovered to be expressed exclusively in tapetal cells (Wellmer et al., 2006). *At1g61110* was named *TAPNAC* for its anther tapetal specific expression. Because, TAPNAC displays a simpler pattern of expression when compared to *ATNAM*, it was decided to do a functional characterization of *TAPNAC* and to later determine if the same functions can be found in *ATNAM*.

Below, a review of flower and anther development to set the stage for this study is presented.

Arabidopsis **floral development**

Arabidopsis floral development is the best understood developmental process in plants. In *Arabidopsis* and most dicot plants, flowers are composed of four distinct organ types arranged in four concentric rings or whorls. From outer to inner the whorls are: sepals, petals, anthers and carpels (Krizek and Fletcher, 2005). Flower development is acropetal, meaning that the oldest flowers are at the base of the inflorescence, and the new flowers are at the tip. *Arabidopsis* flower development has been divided into 13 floral stages; each stage is characterized by a specific morphological marker, described in Table 1.1 (Smyth et al., 1990).

The molecular mechanisms that occur during flowering are very well known. More than 17 years ago the first genes involved in this process were cloned from *Arabidopsis thaliana* and *Antirrhinum majus* (Schwarz-Sommer et al., 1990; Coen and Meyerowitz, 1991), and today has been discovered another level of gene regulation via microRNAs in the development of flower organs (Krizek and Fletcher, 2005).

Table 1.1. Description of the morphological characteristics that distinguish the 13 *Arabidopsis* floral stages.

Stage	Morphology	Duration (h)	Age of flower (days)
1	Inflorescence arise	24	1
2	Floral primordium forms	30	2.25
3	Sepal primordium forms	18	3
4	Sepals overlie floral meristem	18	3.75
5	Petal and stamen primordia arise	6	4
6	Sepals enclosed the bud	30	5.25
7	Long stamen primordia stalked at base	24	6.25
8	Locules appear in long stamen	24	7.25
9	Male meiosis begins	60	9.75
10	Petals are level with short stamen stage	12	10.25
11	Stigmatic papillae appear	30	11.5
12	Petals are level with long stamen	42	13.25
13	Anthesis (pollen dehiscence)	6	14

The approximate duration of the stage and the age of flower at the end of the stage are indicated in the table.
Adapted from (Smyth et al., 1990).

Although much is known about floral development in a broad molecular context, there is still much to learn about the molecular genetic and the biochemical details of floral organ development. This is certainly the case for stamen development.

Stamen and anther development

The flower of *Arabidopsis thaliana* has six stamens, four medial (long) and two lateral (short); each stamen consists of an anther and a filament. Mature male gametophytes (pollen grains) are produced in the anthers through meiotic and mitotic events. In this review, I will focus on anther development, specifically tapetum biogenesis and pollen grain formation.

Anther development has been divided into 14 stages using morphological and cellular features from light microscopy (Sanders et al., 1999). These stages belong to two phases: Phase 1, cell division and differentiation that results in the formation of anther tissues (stages 1-8), and Phase 2, microspore development and pollen release (stages 9-14).

At stage 1, divisions in the L1, L2, and L3 layers of the floral meristem produce the anther primordium; subsequent anticlinal cell divisions in L1 will form the **epidermis** and in L3 will form the connective and vascular tissues. During stage 2, four clusters of archesporial cells are formed from periclinal cell division of the L2 layer. The archesporial cells then divide at stage 3 to produce both the primary parietal layer and the **sporogenous cells**. The primary parietal layer divides to form two secondary parietal layers (SPL). The external SPL further divides and differentiates at stage 4 into an **endothecium layer** and a **middle layer**. During this stage, cells of the internal SPL divide and develop into the **tapetum** layer. Stage 5 is characterized by the four lobed morphology of the anther. Each lobe contains four non reproductive layers: the epidermis, the endothecium , the middle layer, and the tapetum. These somatic tissues surround the microspore mother cells (MMC) which will undergo meiosis to

produce haploid cells during stage 6. Callose deposition on the wall of the meiotic cells occurs in stage 6, and by stage 7, meiosis is completed and the formation of tetrad cells can be observed (Fig. 1.1). Subsequently, callase is secreted from the tapetum into the anther locule to degrade callose (Piffanelli et al., 1998; Ma, 2005). The degradation of callose will cause the release of the microspores from the tetrads at stage 8.

Microspores then develop into pollen grains during stages 9-12, as shown in Fig. 1.1. In the process of pollen grain development, the tapetal cells disintegrate releasing their contents into the anther locule (stages 10 and 11), and pollen mitotic divisions occur producing tricellular haploid pollen grains (stage 12). Afterwards, the anther dehisces and pollen is released (stage 13) followed by the shrinkage of anther cells (stage 14) completing Phase 2 (Ma, 2005).

Meiosis and pollen development

Meiosis has already occurred during anther stage 6. Meiosis is a conserved cell division process that is essential for eukaryotic sexual reproduction and produces haploid cells from diploid parental cells. Following a single round of DNA replication, there are two rounds of nuclear division, meiosis I and meiosis II. Meiosis I involves the segregation of homologous chromosomes, whereas meiosis II is similar to mitosis and results in the segregation of sister chromatids. Following meiosis II, the cell undergoes cytokinesis to produce four haploid cells (Ma, 2005).

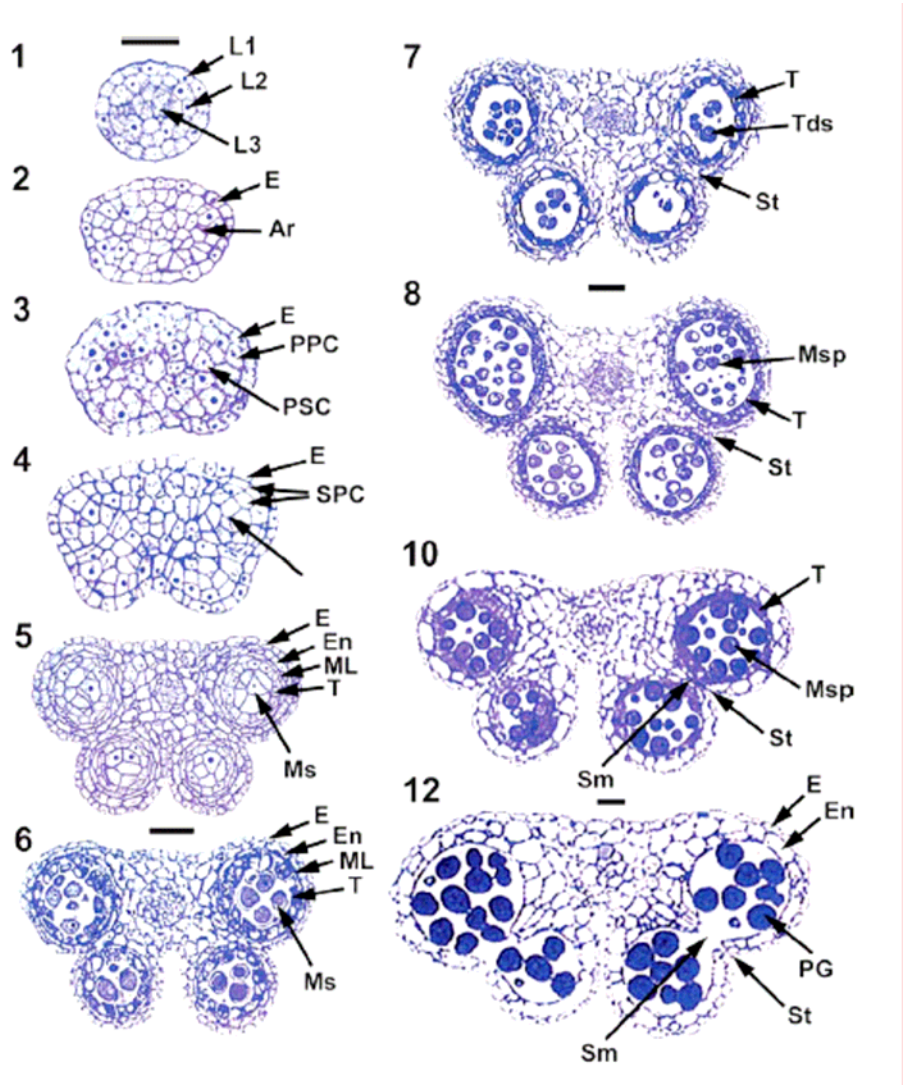


Figure 1.1 Anther cell differentiation in *Arabidopsis.**

Shown are the stages of wild-type *Arabidopsis* thaliana anther development. The numbers indicate stages. Bar = 25μm and stages 1 to 5, 6 and 7,8 and 10, and 12 alone have same sized bars. Ar, archesporial; E, epidermis; En, endothecium; L1,L2,L3, Layer 1,2,3; ML, midle layer; Ms, microsporocytes; Msp, microspore; PG, pollen grain; PPC, primary parietal cell; PSC, primary sporogenous cell; Sm, septum; SPC, secondary parietal cell; St, stomium; T, tapetum; Tds, tetrads.

* Reprinted, with permission, from the Annual Review of Plant Biology, Volume 56 (c) 2005 by Annual Reviews (Ma, 2005)

Each free haploid cell is characterized by a large centralized vacuole with the nucleus located to one side of the cell. Two mitotic divisions will occur in the case of *Arabidopsis* plants, the first asymmetric mitotic division produces a large vegetative cell with a dispersed nucleus and a small generative cell with highly condensed chromatin that later is engulfed by the vegetative cell; and the second mitotic division of the generative cell will produce two sperm cells. (Ma, 2005).

Microsporangial development

Development of the four non reproductive layers: the epidermis, the endothecium, the middle layer and the tapetum, and the sporogenous cell layer, proceeds at varying rates. The microsporocytes and tapetal cells enlarge rapidly and are regulated differentially at the level of gene expression. The endothecial layer does not differentiate; the middle layer is crushed by the expansion of the microsporocytes and the tapetum (Fig. 1.2). Thereafter, plasmodesmata are breached resulting in loss of cytoplasmic contact between the tapetum and the microsporocytes. Microspore developmental synchrony is accomplished by cell to cell communication through “cytomictic channels” (0.5 μm in diameter) described first by (Heslop-Harrison, 1966). Later in tapetal development, walls between the tapetal protoplasts disintegrate and interconnecting plasmodesmata enlarge to form irregular channels. Postmeiotic development thus involves crosstalk between tapetal cells and microsporocytes that are sealed within the microsporangium by a lipid/sporopollenin peritapetal membrane (Dickinson, 1970 cited by (Scott et al., 2004).

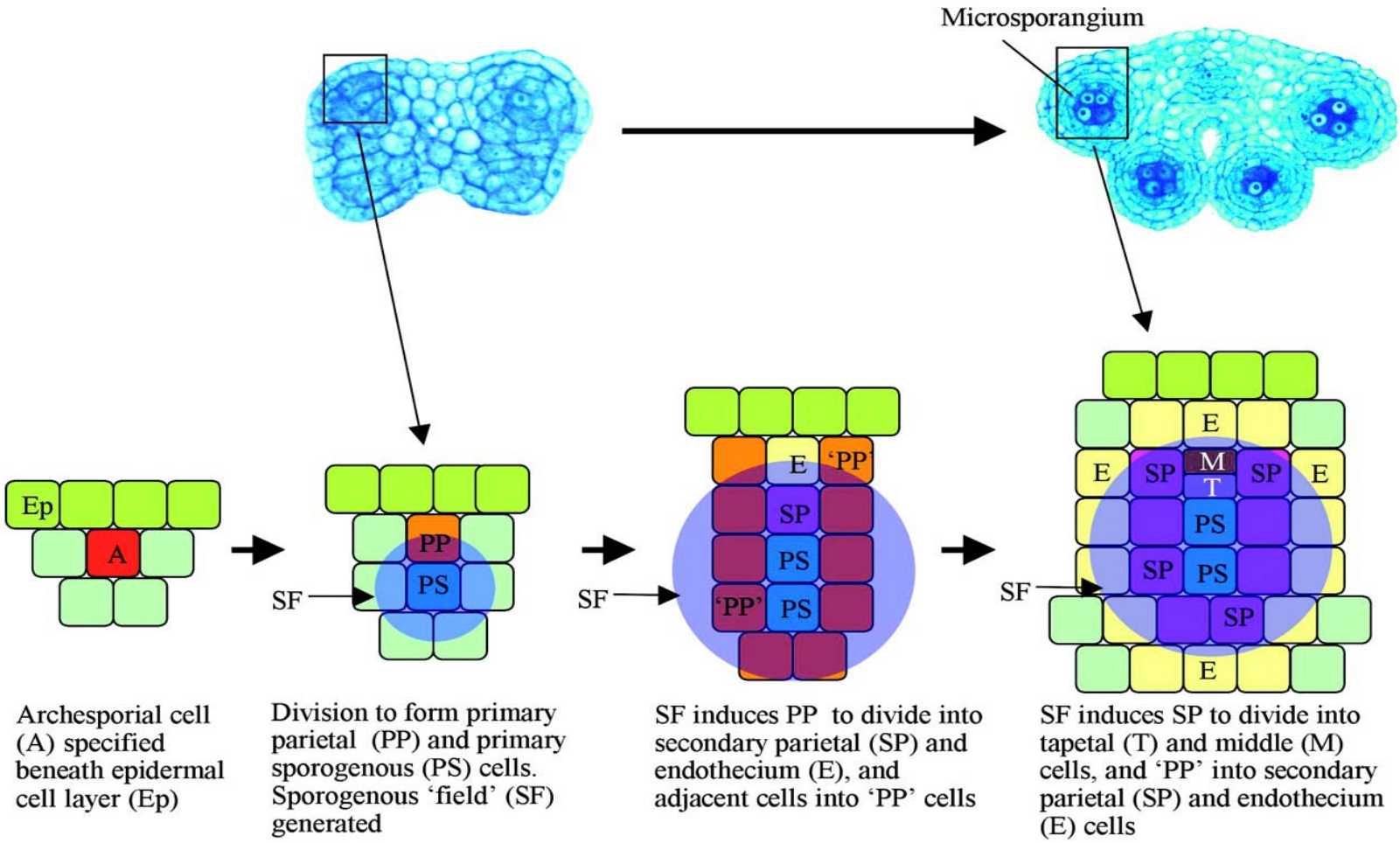


Figure 1.2. Model for the differentiation of the microsporangial cell layers in *Arabidopsis.**

Adapted from (Scott et al., 2004) copyrighted by the American Society of Plant Biologists and is reprinted with permission.

Tapetal development and pollen coat formation

The tapetum begins to divide prior to meiosis of the pollen mother cells; the *Arabidopsis* tapetum will go through several mitotic divisions that will produce polyploid tapetal cells. The first division is usually mitosis without cytokinesis. The doubling of the ploidy level after the second cell division could be caused by endomitosis or restitution mitosis. A third mitotic division has been observed in *Arabidopsis* tapetal cells (Weiss and Maluszynska, 2001).

The tapetum is a short lived and specialized tissue that provides specific compounds for the developing pollen grains in the anther locule. Therefore, becoming polyploid is a more efficient way to amplify genes that are required for the rapid production of specific compounds such as lipids and phenolic derivatives in such a short life span (from anther stage 5 to 12) (Leitch, 2000). In petunia, tapetal nuclei can attain DNA levels of 8C, requiring an intensity of nucleic acid synthesis that is probably unique to the tapetum and that, when combined with high levels of protein synthesis early in development, must create an exceptionally high demand for energy (Liu and Dickinson, 1989).

Tapetal cells can be either secretory, releasing its contents into the locule, or ameboid, moving into the locule and combining with the developing microspores (Pacini, 1990). The *Arabidopsis* tapetum is of the secretory type. Besides its role in pollen wall formation, the tapetum also plays a part in the formation of the lipid rich exine coating in several plant species; for example in *Lilium*, the tapetal cells secrete carotenoids, flavonoids and lipids (pollenkitt) (Reznickova and Dickinson, 1982); in brassicas, the tapetum contributes to the accumulation of lipids, glycolipids and proteins required for pollen-stigma interactions (Dickinson et al., 2000).

The pollen coat is important for desiccation tolerance, hydration of the pollen grain while landing on the stigmatic tissue (Preuss et al., 1993) and for self-incompatibility interactions (Dickinson et al., 2000). The pollen coat of *Arabidopsis thaliana* contains long and short–chain lipids along with a small set of proteins, including six lipases and eight glycine rich oleosin proteins that contain a lipid binding protein, calcium binding proteins and proteins with similarity to the extracellular domains of receptor kinases (Mayfield et al., 2001; Fiebig et al., 2004).

Pollen wall structure, synthesis and patterning

The pollen wall is a complex multilayered structure composed of a pectocellulosic intine surrounded by a sporopollenin-based exine that has two layers, nexine (internal) and sexine (external) terminology as given by Erdtman, 1969 (Blackmore and Barnes, 1990). The most important compound in the pollen wall is the sporopollenin; it confers physical strength, chemical inertness, and resistance to pathogen attacks. Sporopollenin is made of long-chain fatty acids and a minor component of phenolic compounds (coumaric acid) reviewed by (Domínguez et al., 1999; Scott et al., 2004). The phenolic monomers are coupled by ester bonds characteristic of polyphenolics with protective functions, such as lignin and suberin (Scott et al., 2004)

Anther dehiscence

The mature pollen should be released from the anther in time for pollination to occur; this is a very well coordinated process that apparently can take place independently of pollen grain formation in tobacco, implying a lack of signaling

between pollen grains and the anther cells (Goldberg et al., 1993). The dehiscence process begins with the degeneration of the middle layer and the tapetum; consequently the endothecial layer expands, and wall thickening of the endothecial and connective cells occurs. The septum is degraded producing a bilocular anther, and later the stomiun also disappears (Scott et al., 2004).

At the molecular level, there is a better understanding of the gene regulation behind the process of dehiscence. MYB26 a putative R2R3-type transcription factor (Steiner-Lange et al., 2003) may activate the phenylpropanoid pathway in endothecial cells to provide lignin residues for cell secondary wall thickening (Yang et al., 2007). Jasmonic acid (JA), a lipid derived plant hormone, is also involved in the process of dehiscence. Mutants in the JA biosynthetic pathway display the same phenotypes, *i.e. delayed anther dehiscence* / *opr3* that can be rescued by the application of MeJA to the flowers (Sanders et al., 2000; Stintzi and Browse, 2000) .

The role of JA in male fertility is supported by the JA signal transduction mutant *coronatine-insensitive1*, which is male sterile and insensitive to JA treatment (Feys et al., 1994; Xie et al., 1998) Another discovery related to JA is the gene *DEFECTIVE IN ANTHER DEHISCENCE (DAD1)* which synthesizes a phospholipase A1 that catalyzes the initial step of JA biosynthesis. *DAD1* is only expressed in the anther filament prior to flower opening, indicating that the filament is the main source of JA within the flower (Ishiguro et al., 2001). Based on this discovery, a model has been proposed where JA regulates water transport in the stamens and petals that results in the coordination of flower opening, filament elongation and anther dehiscence (Ishiguro et al., 2001) Further support for this hypothesis is that AtSUC1, a plasma membrane H1 sucrose symporter, increases water uptake by transporting sucrose (Stadler et al., 1999). *AtSUC1* accumulates in the connective cells surrounding the vascular

tissue at the final stages of anther development. An alternative model for JA action in dehiscence is the regulation of programmed cell death in the anther (Zhao and Ma, 2000).

Exine secretion and formation

The tapetum is the main tissue that provides compounds to form the pollen coat. *no exine formation1 (nef1)* mutant plants show defects in exine formation. A reduction of the lipid content of the tapetum plastids and abnormal structures in the chloroplast can be observed in this mutant. NEF1 may facilitate the transport of compounds needed for exine production. Disruption of NEF1 causes abnormal development of plastids in the tapetum and leaves (Ariizumi et al., 2004).

Other players in proper pollen coat formation are: i) *ms33 (male sterile 33*) a mutant defective in formation of intine and the deposition of tryphine (Fei and Sawhney, 2001), ii) *AtGPAT1* gene encodes a membrane bound glycerol -3-phosphate acyltransferase and when mutated results in an abnormal tapetum and defects in the pollen wall formation (Zheng et al., 2003), iii) *DEX1* is a membrane calcium-binding protein that helps anchor the sporopollenin into the exine; mutants in this novel protein exhibit a defective exine with irregular sporopollenin patterns (Paxson-Sowders et al., 2001). Furthermore, *FACELESS POLLEN 1* (*FLP1*), encodes a protein with sequence similarity to proteins important in cuticular wax deposition, mutations on this gene causes defective sporopollenin and aberrant tryphine (Tohru et al., 2003) in the pollen coat.

Conclusion

The tapetum is a very important nutritional tissue that plays a key role in pollen development. It provides specific compounds for pollen coat formation, such as glycine-rich oleosin proteins, lipases and phenolic compounds. The mixture of the compounds secreted by the tapetum form the sporopollenin. The sporopollenin is a biofilm that confers to pollen grains physical strength, chemical inertness and resistance to pathogen attacks. All of the processes involved in the production of these compounds must be finely regulated in the tapetal cells. Transcription factors must be major players in these processes. For example, MYB26 activates the phenylpropanoid pathway in endothecial cells of the anther tissue providing lignin residues for secondary wall thickening (Yang et al., 2007).

The role of NAC genes in the anther needs to be elucidated; eleven NAC transcription factors are expressed in anthers (Wellmer et al., 2004), including our gene of interest, *TAPNAC* (*At1g61110*). Therefore, the aim of this dissertation is to determine the function of *TAPNAC* in anther development.

CHAPTER II

EXPRESSION PROFILING OF THE NAC GENE SUPERFAMILY

Introduction

The *Arabidopsis* NAC gene superfamily is a group of novel transcription factors only found in plants. NAC genes share an amino acid sequence domain, called the NAC domain (NO APICAL MERISTEM, ATAF, CUP SHAPED-COYLEDON) that in specific instances has been shown to be a DNA binding domain (DBD; Duval et al, 2002). Genetic analysis of mutant alleles and expression assays proved that NAC domain proteins comprise at least one component of the shoot apical meristem (SAM) specification network. Several genes that contain this NAC domain have been identified, including *CUC1-3* (Vroemen et al., 2003; Taoka et al., 2004) *NAM* (*NO APICAL MERISTEM*) (Souer et al., 1996), *ATAF1* (Ping-Li et al., 2007), *ATAF2* and *NAP* (Guo and Gan, 2006).

There are more than 100 genes in the *Arabidopsis* genome that share conserved regions with the NAC domain (DBD). The recent discovery of the functions of several additional NAC genes has demonstrated their diverse role in plant growth and development. These functions vary and include: SAM formation, *CUC1-3* (Vroemen et al., 2003; Taoka et al., 2004); secondary wall thickening, *NST1* and *NST2* (Mitsuda et al., 2005); drought stress, *RD26* (Fujita et al., 2004); mechanical wounding, *ATAF1*; plant defense (Collinge and Boller, 2001); and even functions related to nutrient remobilization, *TaNAM-B1* (Uauy et al., 2006).

Despite progress in understanding the functions of several NAC genes, there

are many more NAC genes for which we do not know the function. Our initial interest was to study the function of *ATNAM* (*At1g52880*), a NAC gene characterized earlier in our lab (Duval et al., 2002). *In silico* analysis of the protein sequence of 109 NAC genes found in the *Arabidopsis* data base (TAIR) revealed a close relationship between *ATNAM*, *At1g61110* and *At3g15510* along with additional NAC genes that share significant sequence similarity. We used a focused genomics approach to define the gene expression patterns of NAC genes in *Arabidopsis*.

Microarray gene expression offers the possibility to obtain a global understanding of biological processes in living organisms. We can observe in parallel several thousands of genes and their responses to different stimulus or mutations of specific genes. We used Affymetrix ATH1 microarray chips to analyze NAC gene expression profiles in specific tissue types: heart embryo, mature embryo, root, leaf and flower, using. The expression of some of the NAC genes was validated with qPCR. Besides NAC genes, other tissue specific genes were identified. The information obtained through the ATH1 chips will help in the understanding of the NAC gene roles in specific tissues.

Materials and methods

Multiple sequence alignment

NAC protein sequences were obtained throught BLAST. We used ATNAM (At1g52880) protein sequence as a query in Washington University BLAST (WU-BLAST) version 2.0 program available at “The *Arabidopsis* Information Resource” (TAIR). Default parameters were used for this BLAST search.

The NAC protein sequences obtained were analyzed using MEGA version 4.0 (Tamura et al., 2007). We performed a phylogeny reconstruction analysis on the *Arabidopsis* NAC protein sequences using the neighbor-joining (NJ) method and the Poisson correction model. Pairwise deletion method was used to correct for gaps or missing data. Finally, a bootstrap / phylogenetic test analysis (1000 replicates) was included.

Affymetrix ATH1 gene chip analysis

Affymetrix ATH1 chips representing 22,814 genes of the *Arabidopsis* genome were used to assess the expression of five different plant tissues. Heart and mature embryo, roots, leaves and flowers were collected from *Arabidopsis* plants on December 2002 and April 2003. These two sets of biological samples were processed for hybridization, scanning and intial data processing, following standard Affymetrix protocols. Resulting raw data was analyzed using GeneSpring software (Agilent). Per chip and per gene normalizations were applied. Each measurement was divided by the 50th percentile of all measurements in that sample; then, the expression value of each gene was divided by the median of its measurements in all samples. A cross gene error model was also used to account for the lack of technical repeats; this model generates a t test p-value. A list of NAC genes that were present in the chips and have p-values equal or less than 0.1 was generated.

Since *At1g61110* was a floral specific NAC gene, we also generated a list of genes that are detected only in the flower. The following normalizations and filters were used: genes with transcripts present in flower but absent in the other tissues were submitted to ANOVA analysis with p-value cutoff of 0.05 or less. In addition, Benjamini and Hochberg false discovery rate were also used.

Quantitative real-time RT-PCR (qPCR)

2 μg total RNA was reverse transcribed with RT reaction mix including 1 X TaqMan RT buffer, 5.5 mM MgCl2, 500 μM dNTP mixture, 2.5 μM random hexamers, 40 U RNase inhibitor and 125 U Multiscribe™ reverse transcriptase using the TaqMan® reverse transcription kit (Applied Biosystems, Foster City, CA) or SuperScript™ First-Strand Synthesis System for RT-PCR (Invitrogen Life Technologies, Carlsbad, CA) following the manufacturer's instructions. Primers were designed by the Primer Express version 1.5 software (Applied Biosystems, Foster City, CA). qPCR was performed in an optical 96-well plate with a GeneAmp 7500 sequence detector (Applied Biosystems, Foster City, CA), using Power SYBR® Green to monitor dsDNA synthesis. Three identical reactions were repeated on the plates. Reactions contained 10 μl 2 X Power SYBR® Green Master Mix reagent (Applied Biosystems, Foster City, CA), 20 ng cDNA and 5 μM of each forward and reverse gene-specific primers in a final volume of 20 μl. Data was analyzed using the SDS 1.7 software (Applied Biosystems, Foster City, CA). In order to compare data from different PCR runs or cDNA samples, Ct values for the genes were normalized to the Ct value of 18S RNA. Quantification of the abundance of each transcript was determined using the comparative Ct method. The amount of target, normalized to 18S RNA is given by: $2^{-(\Delta Ct)}$ (user bulletin 2, ABI Prism 7500 sequence detection system; Applied Biosystems).

For comparison with the Affymetrix microarray values, the transformed Ct values $2^{-(\Delta Ct)}$ from the qPCR were analyzed with GeneSpring software (Agilent). The only normalization performed was a per gene normalization. Each gene measurement was divided by the median of its measurements in all samples.

Results and discussion:

In silico analysis of NAC genes

The NAC protein sequences for 109 NAC genes were identified through a WU BLAST search from the TAIR website, using as query the protein sequence of ATNAM (AGI At1g52880). MEGA 4.0 software was used to align the NAC protein sequences through ClustalW algorithm (Thompson et al., 1994). The alignment was then submitted for phylogenetic analysis using the neighbor-joining (NJ) method and bootstrap analysis.

The alignments generated a NJ tree of the protein sequences with the bootstrap numbers indicated for every cluster (Fig. 2.1). Clusters containing NAC proteins with high sequence similarity were observed. Since the amino acid sequence of a protein determines protein structure and subsequently protein function, we can infer that proteins that have high sequence similarity, probably fold in the same fashion, and could perform similar functions.

Sequence comparisons using the NJ method identifiedNAC genes that are similar and thus may have similar function. For example CUC1 and 2 NAC proteins (Aida et al., 1997) were grouped together in the same cluster. Also, the NAC genes involved in secondary wall formation NST1 (At2g46770) and NST2 (At3g61910) (Mitsuda et al., 2005) were grouped together. Furthermore the vascular related NAC domain proteins VND1, VND2, VND3, VND4, VND5, VND6 and VND7 (Kubo et al., 2005) clustered together (Fig. 2.1).

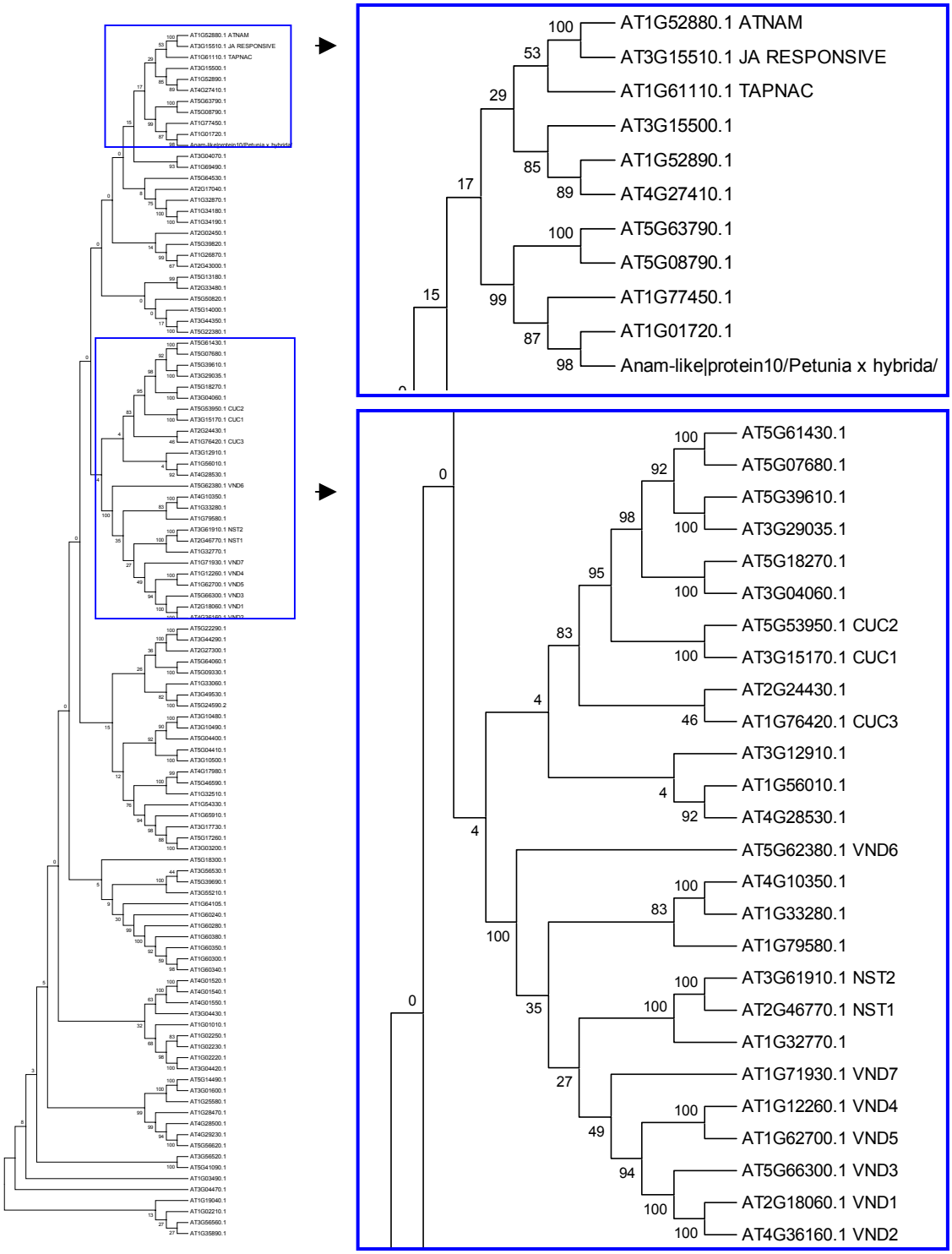


Figure 2.1. Phylogenetic analysis of NAC proteins*.

Bootstrap numbers supporting this neighbor-joining tree are indicated for every cluster
* Appendix figure A.1 shows the complete NJ-tree

Based on this alignment, we identified two NAC genes that were closely related at the protein level to *ATNAM* (*At1g52880*); these were *At1g61110* and *At3g15510* (putative jasmonic acid responsive). We knew that *ATNAM* is abundantly expressed in the embryo, we did not know where the other two NAC genes were expressed, and we did not know the expression profiles of most of the other NAC genes. We used a focused genomics approach to identify expression profiles for the NAC genes.

Expression profile of NAC genes

The Affymetrix ATH1 *Arabidopsis* gene chip was used to asses the global expression patterns of the majority of AtNAC genes. This is a near genome wide chip, representing 22,814 entries; approximately 90% of the NAC genes are represented on the ATH1 chip. Microarray gene expression offers the possibility to obtain a global understanding of biological processes in living organisms. We can observe in parallel thousands of genes and their responses to different stimulus or mutations on specific genes.

Genome wide expression was analyzed using total RNA from developing seeds (heart and mature embryo), root, leaf and flower. Two biological repeats, collected December 2002 and April 2003, were used. RNA samples were processed, labeled, hybridized and data obtained using standard Affymetrix protocols (http: // www. affymetrix.com/ support/ technical/ manual/ expression_ manual .affx). Affymetrix data was initially analyzed using Affymetrix Microarray Suite 5.0. The data was then ported to workstations running GeneSpring 4.2.1 for further analysis.

NAC gene expression profiles were obtained once the two biological repeats were combined and normalized using per gene and per chip normalizations. In

addition, a cross gene error model was used to account for the lack of technical repeats within samples. Of the more than 90 NAC genes on the ATH1 chip, consistent data with the two biological repeats was only obtained for 19 genes. *ATNAM* did not pass the restriction of t-test p-value cut off of 0.1 or less, therefore it is absent from the list of consistently expressed NAC genes. The relative expression of these genes is shown in Fig. 2.2, and the associated levels of expression are presented in Table 2.1.

The expression of *At1g61110* was remarkably high in floral tissue, 126 relative units when compared to the median of the genes (1 relative unit) expressed on the chip. *At1g61110* is a very abundant NAC gene, and its expression was detected only in the flower. This is a very different expression pattern compared to *ATNAM* (Fig. 2.1) which is mainly expressed in the developing embryo (Duval et al., 2002). Another remarkable expression pattern was displayed by *At3g29035*, transcripts of this gene are abundant in leaf and root tissue. Transcripts of *At2g02450* and *At1g69490* (*NAP)* (Guo and Gan, 2006) are both expressed at high levels in leaf and floral tissue (Table 2.1). The rest of the NAC genes do not display a specific pattern of expression; they are found at variable levels in the different tissues tested. Although, several NAC genes were abundantly expressed in leaf tissue.

Validation of the Affymetrix data was performed by qPCR using RNA from mature embryo, root, leaf and flower. Primers for six of the 19 identified NAC genes were designed (sequences are shown in Appendix table A.1), and the expression levels in the different tissues were evaluated.

Figure 2.3 shows the relative expression levels of these six genes. *At1g61110* was confirmed to be the most abundantly expressed NAC gene in the flower. The others had expression levels comparable to those in the Affymetrix experiments (Table 2.2).

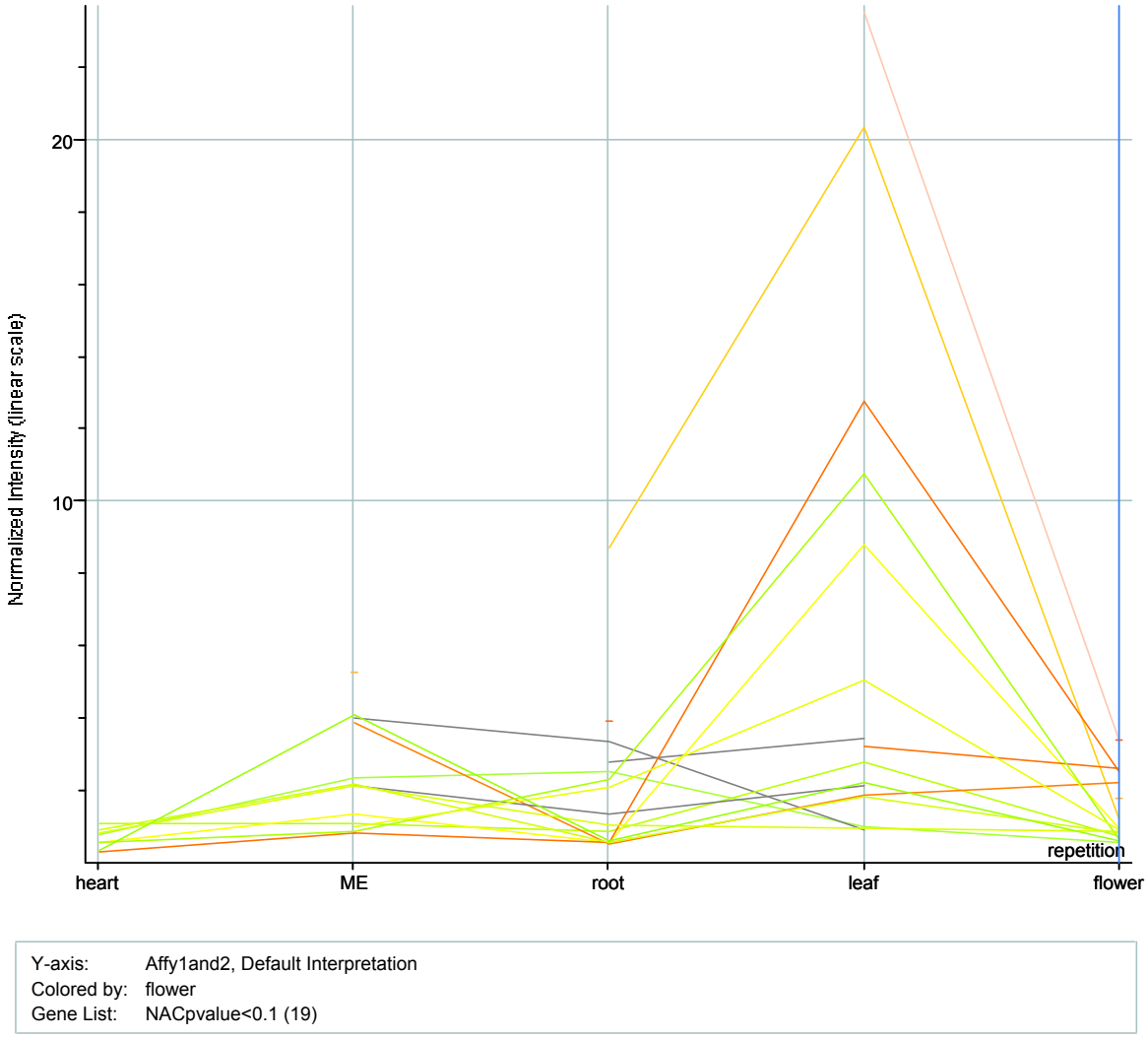


Figure 2.2. Expression profile of NAC genes.

Amplified RNA corresponding to heart embryo, mature embryo, root, leaves and flowers was biotin labeled and hybridized to GeneChip *Arabidopsis* ATH1 Genome Array. The graph shows the expression of 19 NAC genes that pass a t-test p value cut off of 0.1. The AGI numbers and associated expression levels are presented in Table 2.1.

Table 2.1. Relative abundance of 19 NAC genes found in the ATH1 Affymetrix chip.

Heart embryo		Mature embryo		Root		Leaf		Flower	
Norm	AGI	Norm	AGI	Norm	AGI	Norm	AGI	Norm	AGI
1.1	At4g01550	5.3	At5g07680	8.7	At3g29035	23.5	At2g02450	126.1	At1g61110
0.9	At2g33480	4.1	At4g27410	3.9	At4g28530	20.4	At3g29035	3.4	At4g28530
0.8	At3g49530	4.0	At2g24430	3.3	At2g24430	12.7	At1g69490	3.4	At2g02450
0.8	At1g79580	3.9	At3g15500	2.8	At2g43000	10.8	At5g13180	2.6	At1g52890
0.6	At5g13180	2.4	At1g79580	2.5	At1g79580	8.8	At5g39610	2.5	At1g69490
0.6	At5g39610	2.2	At2g33480	2.3	At5g13180	5.1	At5g63790	2.2	At3g15500
0.4	At4g27410	2.1	At1g33060	2.1	At5g63790	3.4	At2g43000	1.8	At5g07680
0.3	At1g52890	2.1	At3g49530	1.3	At1g33060	3.2	At1g52890	1.2	At3g29035
		1.4	At5g39610	1.0	At3g49530	2.8	At4g01550	1.0	At5g39610
		1.1	At4g01550	0.9	At4g01550	2.2	At4g27410	0.9	At5g63790
		1.0	At5g63790	0.6	At4g27410	2.1	At1g33060	0.9	At3g49530
		0.9	At5g13180	0.6	At5g39610	1.9	At3g15500	0.8	At2g33480
		0.8	At1g52890	0.6	At1g69490	1.8	At2g33480	0.7	At4g01550
		0.8	At1g69490	0.5	At2g33480	1.0	At1g79580	0.7	At5g13180
				0.5	At3g15500	1.0	At3g49530	0.6	At4g27410
						0.9	At2g24430	0.6	At1g79580

The table shows a list of NAC genes sorted from the highest expression values to the lowest for each tissue type. Norm, normalized values, AGI, *Arabidopsis* genome initiative number. t test p-value cut off <0.1

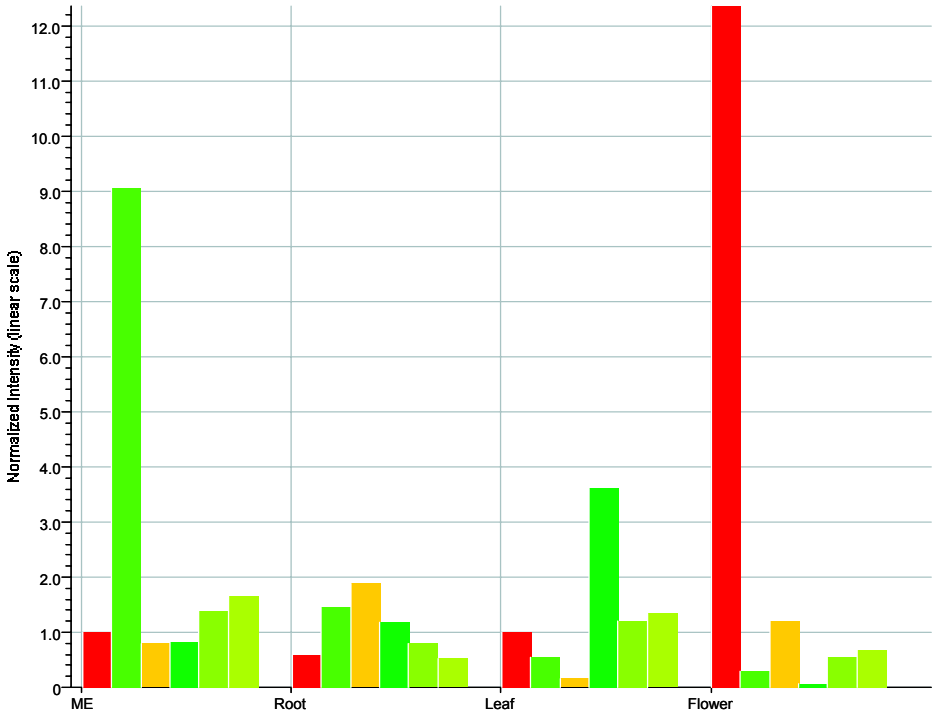


Figure 2.3. qPCR analysis of selected NAC genes for validation of Affymetrix expression profile.

The bar graph obtained with GeneSpring software (Agilent) shows the relative expression of six NAC genes with statistically significant expression in a specific tissue type. The AGI numbers of the genes are as follow, from left to right the bars represent: *At1g61110*, *At1g79580*, *At3g15500*, *At1g52890*, *At4g27410* and *At5g07680*.

Table 2.2. Comparison of qPCR and Affymetrix expression values on selected NAC genes.

	RT PCR				Affymetrix ATH1			
AGI	ME	Root	Leaf	Flower	ME	Root	Leaf	Flower
At1g61110	1.0	0.6	1.0	981.1				126.1
At1g 79580	9.1	1.5	0.5	0.3	2.4	2.5	1.0	0.6
At3g15500	0.8	1.9	0.2	1.2	3.9	0.5	1.9	2.2
At1g52890	0.8	1.2	3.6	0.1	0.8		3.2	2.6
At4g27410	1.4	0.8	1.2	0.5	4.1	0.6	2.2	0.6
At5g07680	1.7	0.5	1.3	0.7	5.3			1.8

Floral specific genes

An initial list of floral specific genes (336 genes) was generated from the Affymetrix data. An ANOVA analysis was used to identify 166 genes that passed the filtering restrictions (Appendix table A.2). The floral expression of these genes was corroborated with the AtGen Express Visualization Tool, a compilation of microarray experiments publicly available at http://jsp.weigelworld.org. Pollen specific genes were also found in this list.

The resulting genes were classified based on their associated functions (Table 2.3). The main processes identified were cell wall biosynthesis and carbohydrate metabolism, mainly related to the growth and expansion of new floral organs. Several lipases and LEA / storage proteins were also identified, and they must be part of the tapetum biogenesis. The tapetum is a nutritious tissue rich in lipids and proteins that secrets the sporopollenin, a compound that is part of the pollen coat. The tapetum also provides the steryl esters, tryacylglycerols and oleosins, compounds that help pollen grains with dessication tolerance and rehydration during pollen tube germination.

Some other genes identified in this microarray analysis fell into functional categories that include kinases and transcription factors, apparently related to the regulatory networks of flower development.

Conclusion

The NAC protein sequences for 109 NAC genes were obtained. MEGA 4.0 software was used to align the NAC protein sequences with the ClustalW

Table 2.3. Different functional categories of floral specific genes identified in Affymetrix ATH1 gene chip experiment.

Functional category	Number of genes
Acid phosphatase activity	2
Atp binding	2
Atpase activity	3
Calcium related	5
Carbohydrate metabolism	12
Cell wall activity	18
Contains a signal peptide	5
Cytochrome P450	6
Cytoeskeleton	5
DNA binding	2
Electron transport	2
Expressed protein	9
Glycine/hydroxiproline rich protein	4
Hormone responsive	7
Ionic homeostasis	6
Kinase activity	8
Lea protein related / storage	4
Lipase activity	4
Lipid binding	8
Lipid metabolism	2
Lipid transporter	1
Other functions	13
Oxidoreductase activity	3
Peptidase related	2
Phosphatase activity	2
Pollen and tapetum specific protein	1
Pollen specific proteins	6
Secondary metabolism	2
Secretion	4
Self incompatibility	3
Serine carboxypeptidase activity	2
Transcription factor activity	9
Transferase activity	4

algorithm (Thompson et al., 1994). The alignment was then submitted for phylogenetic analysis using the neighbor-joining (NJ) method and bootstrap analysis.

Clusters containing NAC proteins with high sequence similarity were observed. This was the case for clusters of *CUC1-2* (Aida et al., 1997), and *NST1* and *NST2*, genes involved in secondary wall formation (Mitsuda et al., 2005) and for *VND6* and *VND7*, vascular related NAC domain genes (Kubo et al., 2005). We identified a cluster containing ***ATNAM***, the embryo abundant NAC gene that initiated our interest on NAC genes, ***At1g61110***, a floral specific NAC gene and ***At3g15510***, a putative jasmonic acid responsive gene. We speculated that the functions of these three genes must be similar, and we decided to explore the role of these genes in *Arabidopsis*.

The Affymetrix microarray analysis allowed us to characterize gene expression profiles of NAC genes in different plant organs. We realized that the majority of NAC genes displayed complex patterns of expression, a feature that could reflect their plasticity, their ability to compensate for loss of function of a closely related NAC gene member (functional redundancy). From all the NAC genes analyzed, only *At1g61110*, displayed the simplest and strongest pattern of expression; qualities that would be very useful in facilitating the discovery of *At1g61110* gene function. We decided to start with functional analysis of *At1g61110* to later extrapolate our findings on the other two NAC gene members of the cluster, *ATNAM* and *At3g15510*.

CHAPTER III

FUNCTIONAL CHARACTERIZATION OF AN ANTHER TAPETUM SPECIFIC NAC GENE

Introduction

The tapetum is a single cell layer that surrounds the anther locule and is derived from periclinal cell divisions of the archesporial cells in the floral meristem (Ma, 2005). Pollen development and maturation depends on the secretion of compounds from the tapetum. The *Arabidopsis* tapetal cells have a short life span (from floral stage 8 to 12) and are the only cells within the anther tissue that are polyploid. Tapetal cells experience up to three cycles of mitosis without cytokinesis (Weiss and Maluszynska, 2001). This is an efficient way to amplify genes that are required for the rapid production of specific compounds such as lipids and phenolic derivatives in such a short life span (Leitch, 2000). Molecular ablation of the tapetal cells results in pollen abortion and male sterility (Mariani et al., 1990) demonstrating the importance of the tapetal cells in pollen biogenesis

Microscopic studies revealed that mature tapetal cells do not contain a well-developed primary cell wall and are filled with secretory organelles on surfaces exposed toward the locular region of the anthers (Wu and Cheung, 2000). The tapetum also releases enzymes that degrade substances such as callose (β-1,3-glucan) and pectic molecules that attach pollen grains during the pollen tetrad stage, (Rhee and Somerville, 1998).

Furthermore, two new tapetal organelles have been described: the elaioplast and the tapetosome (Hsieh and Huang, 2004). Their contents are secreted into the anther locule when the tapetum is degraded through programmed cell death (PCD) (Papini et al., 1999). The released compounds will form part of the pollen coat and will confer specific characteristics to the pollen grains. The elaioplast contains steryl esters that confer the pollen with dessication tolerance, and the tapetosome contains triacylglycerols and oleosins (amphipatic proteins) important in pollen water uptake when landing on the stigmatic tissue (Hsieh and Huang, 2005).

The regulatory mechanisms involved in the production, storage and secretion of compounds required for pollen development, pollen coat formation and pollen grain maturation are unknown. *At1g61110*, named *TAPNAC* for its tapetal specific expression, could play an important role in the regulation of one or more of these processes. Based on the promoter analysis (chapter IV) we have confirmed the tapetum specific gene expression of *TAPNAC*. Consequently, we needed to know the temporal expression of *TAPNAC* during floral development. Using qPCR techniques we identified floral stage 11 as the time point where *TAPNAC* is maximally expressed in the tapetum, indicating that *TAPNAC* may be involved in pollen maturation and pollen coat formation.

We performed several experiments to try to gain some insights into the function of TAPNAC in the *Arabidopsis* anther tapetum. Phenotypic analysis of plants containing loss of function, over expression and dominant repressor *TAPNAC* constructs were conducted. Affymetrix based transcriptome analysis of the loss of function *tapnac* transgenic line (SALK_060459) was performed in order to identify TAPNAC target genes. *TAPNAC* overexpression transgenic plants using a *35S* CaMV promoter were generated and a morphological phenotype was detected. Plants carrying a single copy of the transgene displayed shorter

stature and produced 50% less seeds than wild type plants. Additionally we evaluated the nutrient composition (nitrogen, zinc and iron) of the seeds set by the over expression lines and by the wild type plants. The results of these experiments will be presented in this chapter.

Materials and methods

Plant materials and plant growth condition

Plants of *Arabidopsis thaliana* Columbia ecotype wild type and loss of function *tapnac* (SALK_060459) were grown in REDI earth soil. Seeds were pre chilled for 3 days at 4°C to break dormancy and then transfered to 21°C / 18°C in a Conviron CMP 3023 chamber. Plants were grown under long day conditions, 20 h light and 4 h darkness.

DNA extraction

DNA isolation was done using an *Arabidopsis* quick genomic DNA prep protocol, published online by Elliot Meyerowitz (2002). One small leaf per plant was harvested and placed in 500 ul of DNA extraction buffer (100 mM Tris pH 8, 50 mM EDTA pH 8, 500 mM NaCl and 10 mM β-mercaptoethanol) the tissue was ground with a small pestle, 35 µl of 20% SDS was added to the mix and placed at 65°C for 5 minutes, then 130 µl of 5 M potassium acetate was added to the mixture and placed on ice for 5 min. The mixture was centrifuged at 15000 g for 10 min and the supernatant was mixed with 60 ul of 3 M sodium acetate and 640 µl of isopropyl alcohol. The sample was placed at -20°C for 10 min to allow nucleic acid precipitation, followed by a 15000g centrifugation for 10 min. The

DNA pellet was washed with 70% ethanol and then resuspended in 50 µl of nuclease free water (Ambion) containing RNase (1 µl RNase per 50 µl of water). 1 µl of this DNA suspension was used for PCR applications. For higher yields of DNA the same protocol was scaled up accordingly.

SALK- TDNA line screening

Searches of the database for T-DNA insertion lines through the SALK Research Institue dedicated website. (http://signal.salk.edu/cgi-bin/tdnaexpress), identified two insertional mutagenesis lines SALK-060447 and SALK-060459. analysis of the *tapnac* T-DNA lines was performed as described below.

PCR screening and primer design: Two *TAPNAC* gene specific primers covering the region where the T-DNA was inserted: 1G61110 LP: ATGGAAAACA TGGGGGATTCG and 1G61110 RP:TTTTGAGAGCATGCCTTCCATT and one primer in the left border of the T-DNA insert, Left T-DNA:CCGATTTCG GAACCACCATC were designed. Two PCR reactions were conducted using either 1g61110 LP and RP primers or 1G61110 RP and Left T-DNA primers. PCR conditions: 94°C 3 min; 94°C 45 sec, 63°C 45 sec, 72°C 1 min for 30 cycles and 72°C 4 min. The fragment obtained with the second pair of primers, was subcloned into pGEMT plasmid and then sequenced to determine the exact location of the T-DNA insert in the *TAPNAC* gene.

Segregation of an unlinked morphological male sterile phenotype: This procedure was performed by backcrossing the *tapnac* KO to *A. thaliana* Columbia ecotype wild type plants. F2 generation plants were segregated base on morphological phenotype, a homozygous T-DNA insertion in the *TAPNAC* gene was verified by PCR using the same conditions described above.

Southern blot analysis to determine T-DNA copy number: The pROK2 vector containing the T-DNA insert was analyzed with Mac Vector software to identify restriction enzymes that will not cut the T-DNA. We selected XhoI, NruI, SalI, NdeI and BglII restriction enzymes to digest 500 ng of genomic DNA, the digested products were resolved in a 0.75% agarose gel and then transferred by blotting to Hybond-N^{+} ™ membrane (Amershan , Arlington Height, IL) with 0.4 N NaOH. Hybridization was performed in a Hybridization caseine base solution (1% casein (Oxoid-Remel LP0041), 7% SDS, 1mM EDTA and 0.25 M disodium phosphate) following standard procedures. The hybridization was performed at 65°C and the stringent washing steps (2XSSC 0.5%SDS and 0.2xSSC 0.1%SDS) were done at the same temperature.

RNA extraction

Total RNA from leaves was isolated following a phenol/SDS method for plant RNA isolation (Ausubel et al., 1990) and total RNA from floral tissue was done using the pine tree RNA isolation method (Chang et al., 1993), a method that is used for tissues containing high amount of polysaccharides and/or phenolic compounds.

Quantitative real time PCR (qPCR)

2 μg total RNA was reverse transcribed with RT reaction mix including 1 X TaqMan RT buffer, 5.5 mM $MgCl_2$, 500 μM dNTP mixture, 2.5 μM random hexamers, 40 U RNase inhibitor and 125 U Multiscribe™ reverse transcriptase using the TaqMan® reverse transcription kit (Applied Biosystems, Foster City, CA) or SuperScript™ First-Strand Synthesis System for RT-PCR (Invitrogen Life

Technologies, Carlsbad, CA) following the manufacturer's instructions. Primers were designed by the Primer Express version 1.5 software (Applied Biosystems, Foster City, CA). A list of the primers used for all the qPCR experiments discussed in this chapter can be found in the Appendix table A.5. qPCR was performed in an optical 96-well plate with a GeneAmp 7500 sequence detector (Applied Biosystems, Foster City, CA), using Power SYBR® Green to monitor dsDNA synthesis. Three identical reactions were repeated on the plates. and the reactions contained 10 µl 2X Power SYBR® Green Master Mix reagent (Applied Biosystems, Foster City, CA), 20 ng cDNA and 5 µM of each forward and reverse gene-specific primers in a final volume of 20 µl. Data was analyzed using the SDS 1.7 software (Applied Biosystems, Foster City, CA). In order to compare data from different PCR runs or cDNA samples, C_t values for the genes were normalized to the C_t value of 18S RNA. Quantification of the abundance of each transcript was determined using the comparative C_t method. The amount of target, normalized to 18S RNA, is given by: $2^{-(\Delta Ct)}$ (user bulletin 2, ABI Prism 7500 sequence detection system; Applied Biosystems).

Affymetrix ATH1 gene chip analysis

Affymetrix ATH1 gene chips (Affymetrix, Inc., Santa Clara, CA) representing 22,814 genes of the *Arabidopsis* genome, were used to assess the expression of TAPNAC loss-of-function plants compared to wild type. *Arabidopsis* flowers at stage 11 were collected for both samples. 100 plants per treatment were used, and tissue collection was performed at the same time of the day in three sets each of 30 plants. cDNA synthesis, cRNA labeling, hybridization and image analysis were performed by Codon BioSciences, LP (Houston, Texas). Resulting raw data was normalized and analyzed using GeneSpring 7.0. Normalization was performed as follows. First, values below 0.01 were set to 0.01. Each

measurement was divided by the 50th percentile of all measurements in that sample. Specific samples were normalized to one another and all the samples were normalized against the median of the control samples. Each measurement for each gene in those specific samples was divided by the median of that gene's measurements in the corresponding control samples. Fold changes and t-test p-values measuring statistical significance of differential expression were calculated for each gene. The significance level was further analyzed by applying multiple testing correction of Benjamini and Hochberg false discovery rate. Genes that showed higher than 1.7-fold difference with $p < 0.05$ were further analyzed. One-way ANOVA was used to identify statistically significant differences in gene expression between wt and *tapnac* floral tissues.

Bioinformatics analysis

Weeder (Pavesi et al., 2004; Tompa et al., 2005) was used to identify consensus motifs present in the up- and down-regulated genes in the *tapnac* KO. Weeder is a general purpose motif discovery tool (Tompa et al., 2005). The software identifies putative motifs of up to 12 bp length containing a maximum of four mistmatches. The parameters used were: R 50, percentage of sequences that must contain the motif (50); M , indicates that a motif can occur more than once in each sequence; S , process both strands of the input sequences; and T 15 , report the number of highest scoring motifs of the run (15). The same set of sequences were randomized using the shuffle sequence program from the X-Blast TAMU server and run under the same parameters described previously.

Overexpression of *TAPNAC* cDNA

The *TAPNAC* coding region was amplified from a pUNI51 vector containing the *TAPNAC* cDNA (ABRC- *Arabidopsis* Biological Resource Center). We used NcoI AT1G61110 primer: CCGGCCATGGAAAACATGGGGGATTCGAGC combined with NheI stop primer: CATGGCTAGCAGATTATGAGTGCCAGTTCA TGTTAG Tm 53°C. The PCR product was cloned into TOPO 2.1 and then sequenced to verify its identity. NcoI and SacI restriction enzymes were used to clone the cDNA fragment into pRTL2 vector. The pRTL2 cassette was then transferred to pCAMBIA 23000 and pCAMBIA 3300 using HindIII. Controls were also generated with the pRTL2 cassette without cDNA in pCAMBIA 2300 and pCAMBIA 3300. The binary vector was transformed into *Agrobacterium* cells (GV3101) and then *Arabidopsis* plants were transfected with these cells. The seeds produced by the transgenic plants were screened for kanamycin or BASTA resistance. Northern and Southern blot analyses were conducted to identify plants that over express *TAPNAC* cDNA and have a single T-DNA insertion.

Phenotype analysis: Homozygous overexpression lines (OE8 and OE23) were planted in Redi-earth soil (Sun Gro Horticulture Canada Ltd.) along with wild type control plants. Plants were grown in Conviron chambers set at 21°C/18°C 20h day cycle with 90 $\mu mol/m^2.s^{-1}$ light intensity. After their life cycle was completed (~2 months), the plants were evaluated; height and the number of siliques on the main stem were recorded. The seeds produced by these plants were harvested in three sets of 30 plants each. The seeds were then sent to The Office of Texas State Chemist (College Station, Texas) for analysis of nitrogen, iron and zinc contents. The protein analysis was done by combustion using AOAC Method 990.03 and the zinc and iron were analyzed by atomic absorption spectroscopy using AOAC method 968.08.

TAPNAC dominant supressor (DS) analysis

The pUNI51 plasmid that carries the coding region (cDNA) of the *TAPNAC* gene was used to amplify a chimeric *TAPNAC–DS* gene. The primers used were NcoI AT1G61110 primer: CCGGCCATGGAAAACATGGGGGATTCGAGC and DS no stop primer: CGATGGTACCGGATCCGCTAGCTTAAGAGAAACCCAAA CGGAGTTCTAGATCCAGATCCAGTGAGTGCCAGTTCATGTTAGGAAGCTGA AA, the PCR product was digested with NcoI and KpnI and subcloned into pRTL2 vector. The *35S* CaMV promoter of the pRTL2 was exchanged with the *TAPNAC* promoter region of -1800bp. Briefly, the promoter region contained in the TOPO 2.1 vector was PCR amplified with the 5' promoter and the EcoRV 3' primers, the PCR product was cloned into the pRTL2 cDNA-DS vector with PstI and EcoRV replacing the *35S* with the *TAPNAC* promoter region. After sequence verification, the chimeric cassette was finally cloned into pCAMBIA 3300 BASTA resistance vector using HindIII restriction enzyme. The final construct was electroporated into GV3101 and then transfected into *Arabidopsis* Columbia ecotype wild type or *tapnac* KO backgrounds. The transgenic plants were screened with BASTA herbicide (FINALE, Aventis) solution containing 240 ug/ml BASTA in 0.005% silwet L-77.

Yeast two hybrid analysis

cDNA library construction: Total RNA from anthers of flowers at stage 10 to 12 were used for cDNA library construction. RNA samples were DNase treated and then evaluated on a Bioanalyzer (Agilent, Santa Clara CA). Total RNA was then used for library construction with Matchmaker Two Hybrid System 3 (Clontech, Palo Alto, CA) according to the supplied manual. Synthesis of first strand cDNA used an oligo dT primer (CDSIII primer) and then synthesis of the

complementary strand required Smart III primer. The amplified ds cDNA was size fractionated with CHROMA SPIN+TE-1000 columns to enrich for longer or complete cDNA transcripts, the sized cDNA was then transformed into AH109 yeast cells.

cDNA library screen: The *TAPNAC* cDNA was cloned into pGBKT7 vector and the introduced sequence was verified. *TAPNAC* transcript in Y187 seems to be toxic so cells don't grow in liquid media but they grow on solid media. In order to do the mating a single colony was picked and resuspended in 1ml DO-trp liquid media, and plated on DO-trp plates then the cells were harvested to initiate mating. Yeast two hybrid screening was done following recommended standard procedures. Briefly, overnight yeast mating of TAPNAC-BD (Y187 cells) with the AH109 yeast cells containing the anther cDNA library was performed along with the positive control mating of AH109[pGADT7-RecT] x Y187[pGBKT7-53] and the negative control mating of AH109[pGADT7-RecT] x Y187[pGBKT7-Lam]. Mated cells were then spread in QDO (quadruple dropout) media that lacks adenine, histidine, leucine and tryptophan. After four days, the cells that grew on these plates were streaked on fresh QDO plates containing 4mg / ml of 5-bromo-4-chloro-3-indolyl α-D-galactopyranoside (XαGal). XαGal is a chromogenic substrate used to demonstrate α-galactosidase activity. We used this substrate to identify real protein-protein interactions and eliminate false positives. Colony growth and blue color indicates an interaction between the two hybrid proteins.

Characterization of positive controls: Positive clones (blue colonies) were grown overnight in 3 ml of YPDA (20 g/L peptone, 10 g/L yeast extract and 0.2% adenine hemisulfate) media containing ampicillin 50 µg/ml (for selection of the pGADT7 library plasmid) at 30°C. The plasmid was recovered with a "smash and grab protocol" (Hoffman and Winston, 1987). The lysis buffer is composed

of 2% triton X-100 (Sigma), 1% SDS, 100 mM NaCl, 10 mM Tris pH 8.0 and 1 mM EDTA. The overnight cultures were collected by centrifugation and resuspended in 0.2 ml of lysis buffer, 0.2 ml phenol:chloroform: IAA (25:24:1) and 0.3 g of 0.45-0.5 glass beads (Sigma). The suspension was vortexed at high speed for 2 min. and then centrifuged for 10 min at 14000 rpm. The supernatant containing the plasmid DNA was precipitated with 1/10 volume of 3 M sodium acetate and 2.5 volumes of 100% ethanol. Thereafter, 1 ul of the plasmid prep was electroporated into DH10B electrocompetent cells (Invitrogen) and plated on LB carbenicillin 100 ug/ml. The plasmids were then sequenced with the 3' AD sequencing primer: AGATGGTGCACGATGCACAG.

Verification of protein–protein interaction: The plasmids obtained from the positive clones were transformed into the yeast AH109 cells and then selected on SD –leu media. The surviving colonies were used for mating with the Y187 strain yeast cells containing: i) pGBKT7 self ligated, ii) pGBKT7-*TAPNAC* cDNA and iii) pGBKT7-Lam. The yeast matings were performed following recommended standard procedures.

Results

Analysis of a T-DNA insertion line in *TAPNAC* gene

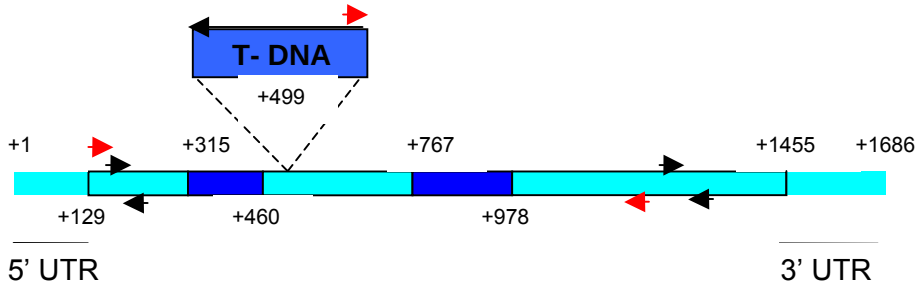
PCR screening methods were used to identify a T-DNA insertion in the *TAPNAC* gene using SALK-060459 line (ABRC). Briefly, two *TAPNAC* gene specific primers (left and right) combined with a T-DNA left border (T-DNA-L) primer were used. Two PCR experiments were run, one with gene specific primers and the other with T-DNA-L and right primers. T-DNA homozygous insertion lines showed no amplified product when using gene specific primers but clearly

amplified a PCR product with the T-DNA-L and the right primers. The PCR product containing the junction within the T-DNA and the *TAPNAC* gene was sequenced to precisely locate the T-DNA insertion in the *TAPNAC* genomic region. The T-DNA was inserted at nucleotide 499 downstream of the 5' UTR in the second exon (Fig. 3.1–A).

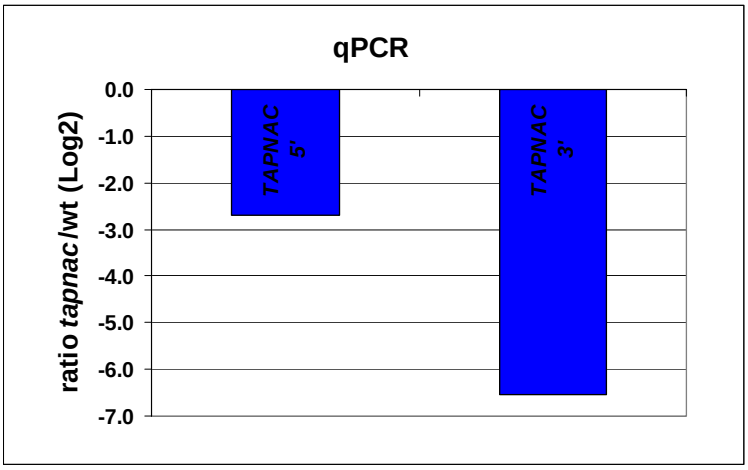
A homozygous line displayed a male-sterile phenotype in less than 5% of the plants (1 in 30 plants). The phenotype was caused by a second T-DNA insertion, which was segregated by backcrossing the SALK-060459 line to wild type, and screening the F2 generation. The male sterile phenotype was segregated out, and the homozygous insertion in the *TAPNAC* gene was recovered. No other morphological phenotype was observed in the *tapnac* KO under the conditions studied.

Once the homozygous T-DNA insertion line in the *TAPNAC* gene was obtained, a qPCR analysis was performed to evaluate*TAPNAC* transcript levels in the T-DNA insertion line. Whole inflorescences were collected from the *tapnac* KO and wild type plants. Total mRNA was extracted, reverse transcribed and analyzed by qPCR. Two sets of primers were used, one pair close to the 5' UTR preceding the T-DNA insertion and the other pair close to the 3' UTR, following the T-DNA insertion. Levels of *TAPNAC* transcript in the *tapnac* KO varied depending on the primer set used. When the 5' primer set was used only a seven fold reduction in *TAPNAC* transcript was detected; however, when the 3' primer set was used a reduction of more than 100 times was reported (Fig. 3.1-B,C).

A



B



C

Region amplified	Ratio t*apnac*/wt	Stdv
TAPNAC 5'	0.1548	0.0404
TAPNAC 3'	0.0107	0.0061

Figure 3.1. Analysis of *TAPNAC* T-DNA insertion line.

A. Physical map of the *TAPNAC* gene. The position of the T-DNA insertion is located in the 2nd exon of the NAC DNA binding domain. The red arrowheads indicate the positions of the gene specific primers: left (+129bp) and right (+1085bp) and the T-DNAL primers use for PCR screening of the homozygous T-DNA insertion line. Positions of the qPCR primers used for the 5' (152-218bp) and 3'ends (1166-1225bp) of the *TAPNAC* gene are shown as black arrowheads. **B**. qPCR analysis of whole inflorescences using 5' and 3' primers (panel A). The ratios were calculated between *tapnac* KO line and wild type, the values are presented as a lo2 transformation. **C**. Original qPCR ratios of the graph in panel B.

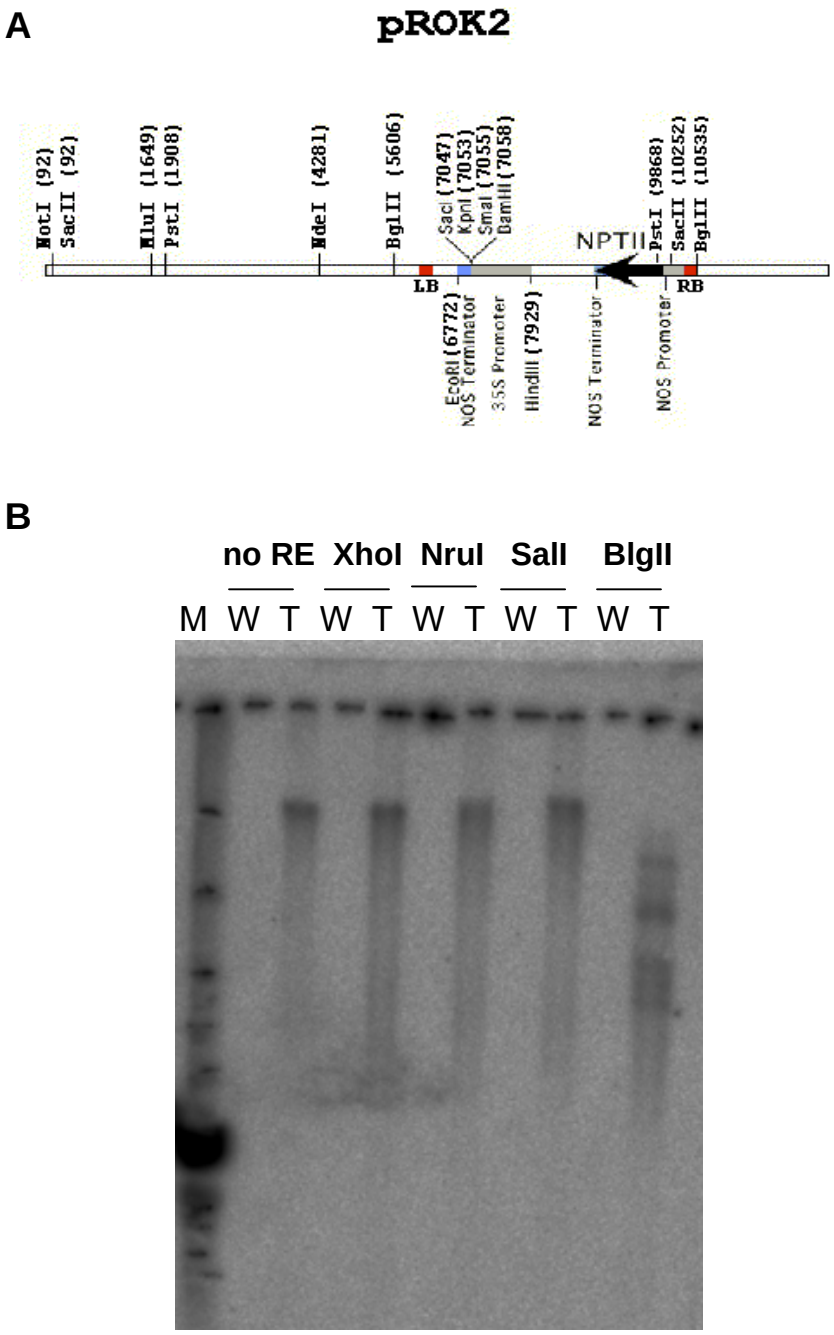


Figure 3.2 Southern blot analysis of T-DNA insertion line in *TAPNAC* gene.

A. Diagram of the vector used in the SALK T-DNA insertion lines, **B.** Southern blot analysis with RE that do not cut the T-DNA insert, except for BlgII that cuts at the end of the T-DNA right border.M, marker; W , wild type and T , *tapnac* T-DNA KO.

Southern blot analysis in the segregated line revealed a single T-DNA insertion when using restriction enzymes that do not cut the T-DNA insert (Fig. 3.2-B). However, when using restriction enzymes that cut within the T-DNA insert, we found that the T-DNA was inserted as tandem repeat (data not shown).

Temporal expression of *TAPNAC*

Although there is much information concerning the events that take place during tapetum development, the regulatory networks involved in these processes are not understood. TAPNAC, A NAC transcription factor, may play a role in the regulation of tapetum development. Therefore precise knowledge of *TAPNAC* temporal expression is crucial to elucidate its role in tapetum development.

Different floral stages of *Arabidopsis thaliana* Columbia ecotype, wild type and *tapnac* KO were collected. Total RNA prepared from stages <10 through stages >12 was used in qPCR experiments to determine the temporal pattern of *TAPNAC* expression in the different genetic backgrounds. The temporal expression of *TAPNAC* is shown in Fig. 3.3. The *TAPNAC* transcript is barely detected at floral stage 10 then increases at floral stage 11 decreasing again at floral stage 12. By floral stage >12 the tapetal cells that harbor *TAPNAC* expression are degraded and consequently TAPNAC expression is not detected.

Two other important transcripts, whose mutations cause male sterility in *Arabidopsis* plants, were also evaluated. The genes were: *OPR3* that encodes an isozyme of 12-oxophytodienoate reductase (Stintzi and Browse, 2000) also called *DELAYED DEHISCENCE1* (Sanders et al., 2000) and *ABORTED MICROSPORES (AMS)* that encodes a MYC class transcription factor

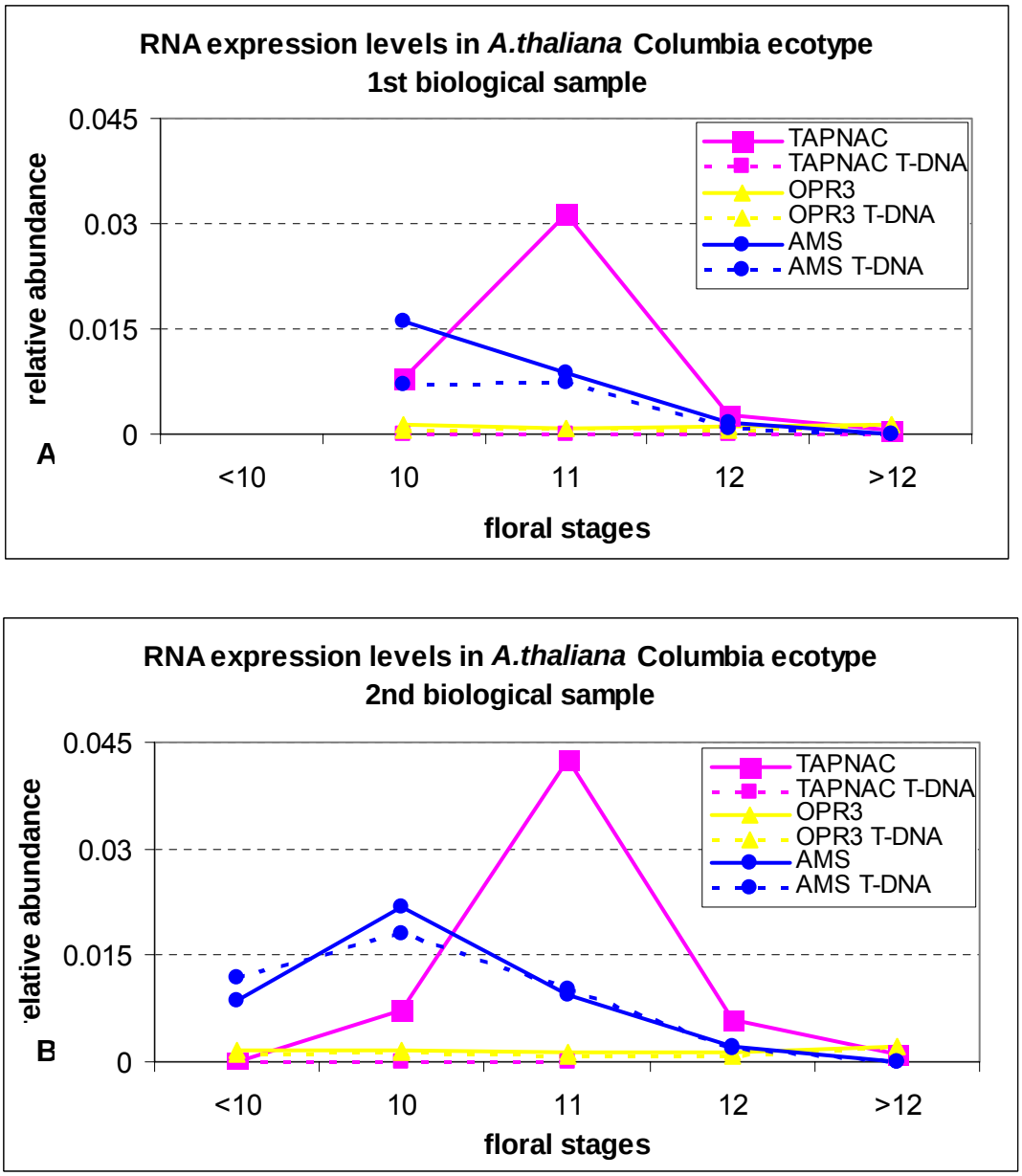


Figure 3.3. Temporal gene expression of *TAPNAC OPR3* and *AMS.*

Wild type and *tapnac* KO mutants plants were grown at two different times in conviron chambers. Flowers corresponding to floral stages <10 (whole inflorescence up to stage 9), 10, 11, 12 and >12 (late floral stage 12) were collected for both genetic backgrounds, total RNA from these samples was assayed for expression levels of *TAPNAC*, *OPR3* and *AMS.* A and B display the temporal expression profile for the 1st and 2nd biological repeats respectively.

(Sorensen et al., 2003). The temporal expression of these genes was different from each other and different from *TAPNAC*. Briefly, *AMS* transcript was abundant in early floral stages and the levels decreased with later floral stages in both genotypes. The *OPR3* transcript was overall less abundant and there were no differences within the mutant and wild type plants

Affymetrix microarray analysis of the *tapnac* mutant

The Affymetrix ATH1 *Arabidopsis* gene chip was used to assess global expression changes within the *tapnac* KO and wild type plants. This molecular phenotype profiling is based on the assumption that knocking down the expression of a transcription factor will have an impact on the transcriptome, since genes regulated by this transcription factor may change their expression profile once the regulatory protein is no longer expressed.

Three biological repeats (30 plants each) were used in this analysis. Flowers corresponding to floral stage 11, the stage with the highest level of *TAPNAC* expression, were collected and RNA extracted. Total RNA was processed, labeled and hybridized by Codon biosciences, LP (Houston, Texas). The raw data was then ported to workstations at TAMU running GeneSpring 7.0 software (Agilent Technologies, Inc., Santa Clara, CA) for further normalization and statistical analysis.

Changes in the transcriptome were evident. A total of 94 genes showed differential expression when the *TAPNAC* transcript was disrupted. Genes that were down- or up-regulated at $p<0.05$ are shown in Figures 3.4 and 3.5 respectively. The list of 52 down-regulated genes including *TAPNAC* and the list

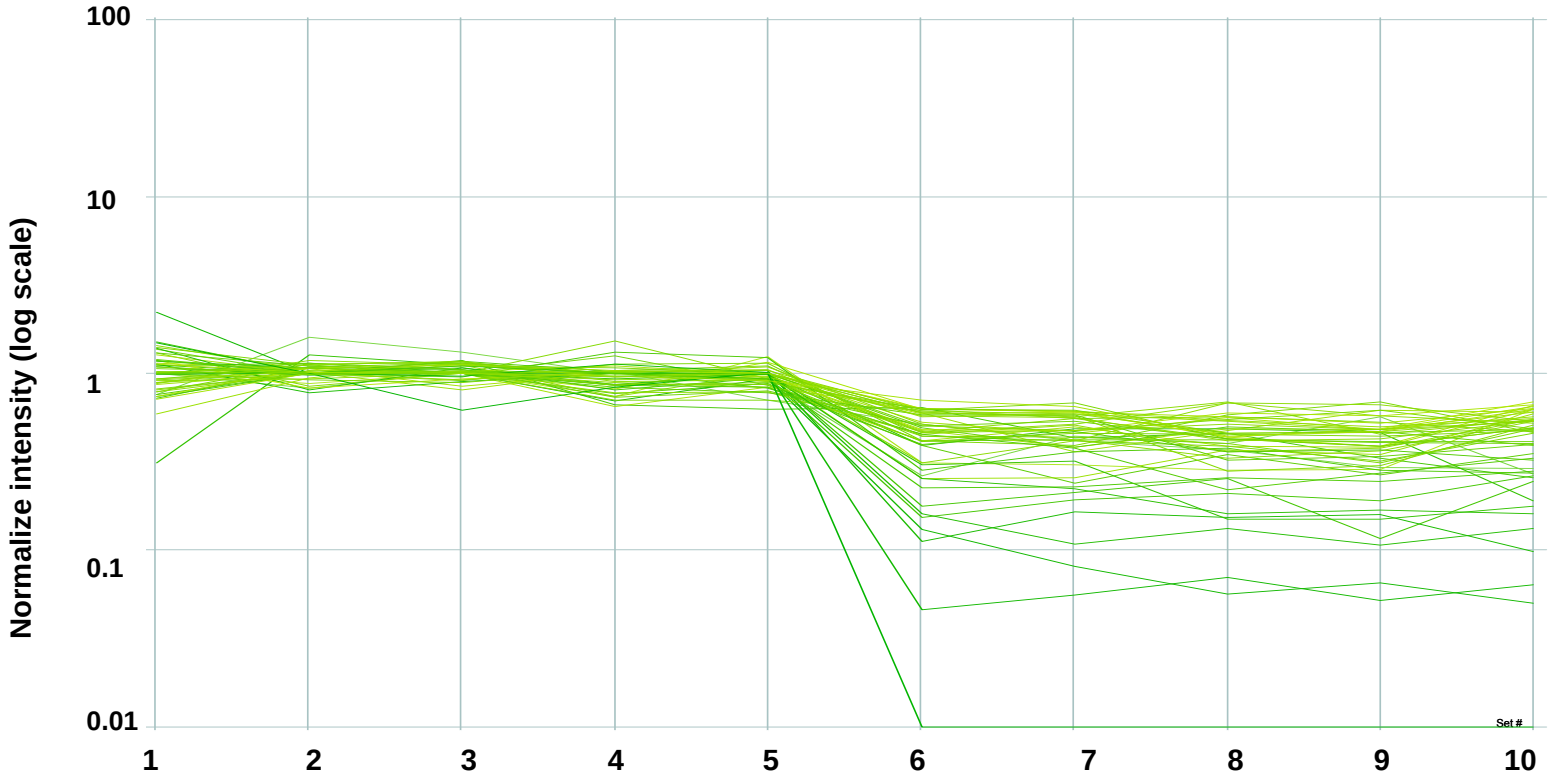


Figure 3.4. Down-regulated genes in the *tapnac* mutant background.

Graphic representation of the down-regulated genes in the floral stage 11 of *tapnac* KO plants (6-10) when compared to wild type control plants (1-5). Normalizations were performed with Genespring 7.0 (Agilent);per gene and per chip normalizations were applied, then all the samples (1-10) were normalized against the median of the wild type control samples (1-5). Only genes that have t- test p-values <0.05 and pass the one way ANOVA test are represented in this graph.

Table 3.1. List of down-regulated genes (52) in the *tapnac* KO when ATH1 Affymetrix chips were used.

Functional category	Number of genes
Calcium-dependent phospholipid binding	2
Carbohydrate metabolism	1
Carotenoid biosynthetic process	1
Catalytic activity	2
Cell wall metabolism	1
DNA binding	1
Electron carrier	1
Endonuclease	1
Expressed protein	9
Hormone responsive	3
Hydrolase	2
Kinase	3
Membrane protein	2
Metal binding protein	1
Other functions	7
Oxidase	2
Phospholipase	1
Transaminase	1
Transcription factor	6
Transferase	2
Transporter	3

The genes depicted in the graph above were categorized by their predicted molecular function.

of 42 up-regulated genes in the mutant background are shown in Appendix tables A.3 and A.4 correspondingly.

The down- and up-regulated genes were classified into different functional categories shown in Tables 3.1 and 3.2. Twelve percent of the down-regulated genes belong to the category of transcription factors, among these regulatory proteins another NAC domain gene was identified, *At5g61430*. This NAC gene is mainly involved in senescence (Lin and Wu, 2004). The category containing the highest number of genes (17%) was the expressed protein category, where

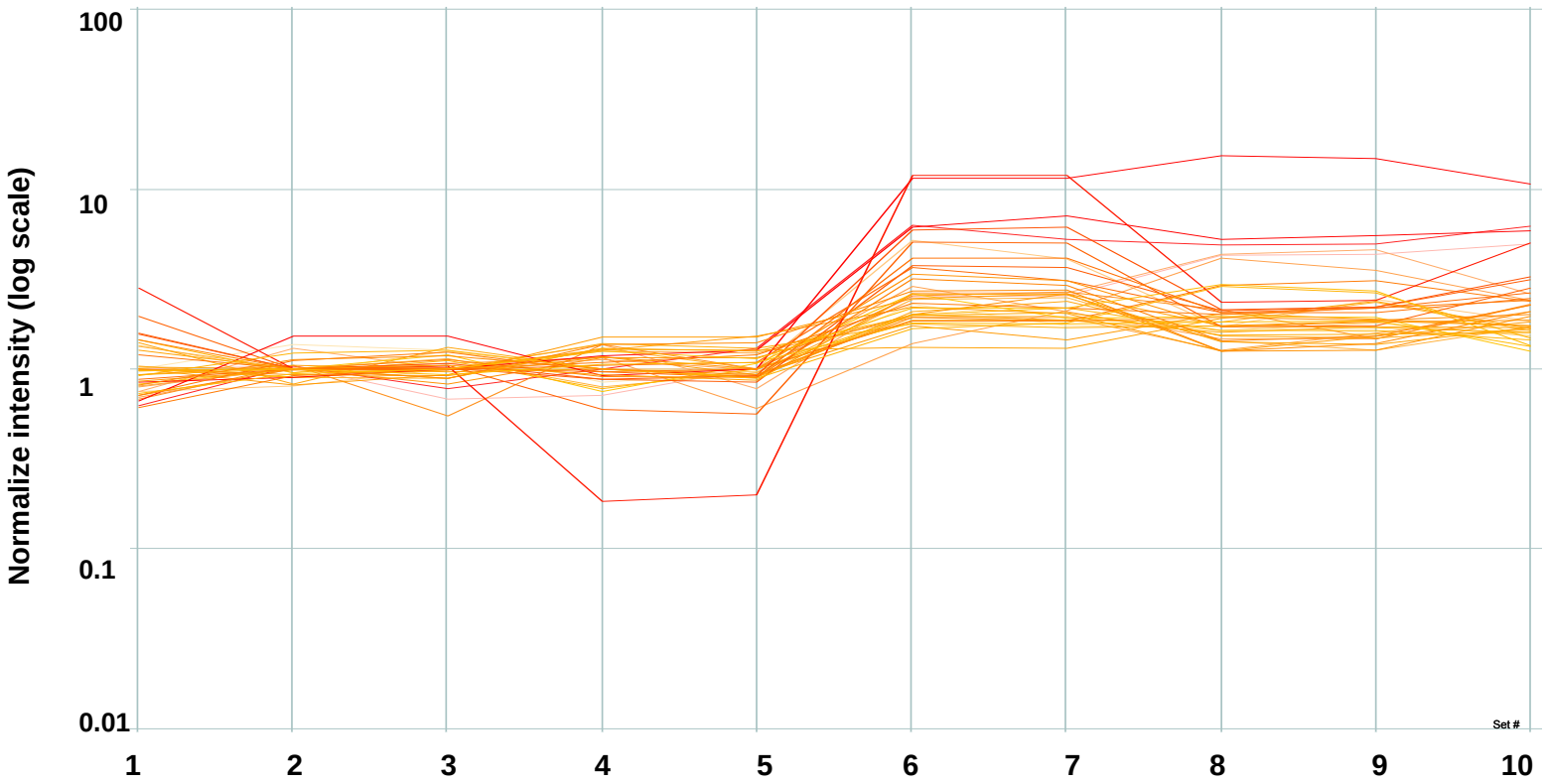


Figure 3.5. Up-regulated genes in the *tapnac* mutant background.

Graphic representation of the up-regulated genes in the floral stage 11 of *tapnac* KO plants (6-10) when compared to wild type control plants (1-5). Normalizations were performed with Genespring 7.0 (Agilent Technologies Inc.); per gene and per chip normalizations were applied, then all the samples (1-10) were normalized against the median of the wild type control samples (1-5). Only genes that have t- test p-values <0.05 and pass the one way ANOVA test are represented in this graph.

Table 3.2. List of up-regulated genes (42) in the mutant background when ATH1 Affymetrix chips were used.

Functional category	Number of genes
Carbohydrate metabolism	2
Cell wall biogenesis	2
Chloroplast specific	3
Defense response	2
Expressed protein	8
Ion homeostasis	5
Kinase	1
Ligase	1
Other functions	7
Proteolysis	1
Stress response	1
Transcription factor	2
Transferase	2
Transporter	4

Genes are categorized base on their predicted molecular function.

no function can be addressed to these genes base on protein sequence similarity. The most down-regulated genes, besides the *TAPNAC* gene in the mutant background were *PHOSPHOLIPASE Dα1, PLDα1* (*At3g15730*) and *RbohE* (*At1g19230*), a respiratory burst oxidase protein. These genes were respectively 20 and 14 times down-regulated when compare to the wild type plants.

From the list of up-regulated genes in the mutant background (Appendix table A.4), we identified a gene that is highly up-regulated, 12-fold up. This gene *At3g30720*, encodes a 59 amino acids protein with low sequence similarity to an oligopeptide transporter and to a catalytic ligase. This gene produces a very small transcript with half of the transcript encoding for the 5' UTR and lacking a 3' UTR. We need to note that this gene is located close to the centromeric region of chromosome 3, which encodes putative non-protein encoding transcripts.

Three other genes displayed a high level of induction: *At5g16920* (fascilin containing domain), *At3g61930* (gene involved in N-terminal protein myristoylation, related to membrane interaction of various proteins) and *At3g25260*, classified as a nitrate transporter and oligopeptide transporter. Only two up-regulated genes belonged to the transcription factor category, and seven genes were involved in ion homeostasis. As in the case with the down-regulated genes, 19% up-regulated genes do not have an ascribed function.

Table 3.3. Comparison of microarray ratios with qPCR to validate the difference in transcripts abundance within *tapnac* KO and WT control plants on selected genes.

AGI	microarray ratio	qPCR ratio	Function
At1g19230	0.08	0.04	Respiratory burst oxidase protein E **RbohE**
At1g15580	0.14	0.10	AUXIN induced
At3g15730	0.06	0.11	**Phospholipase Dα1**
At2g19500	0.22	0.13	Putative cytokinin oxidase
At5g61620	0.32	0.16	MYB transcriptional activator
At5g47330	0.19	0.21	Palmitoyl-(protein) hydrolase activity
At1g59740	0.25	0.25	**Oligopeptide transporter**
At5g63850	0.36	0.52	**Amino acid transporter**
At5g61430	0.49	0.63	NAC transcription factor activity
At5g10180	1.91	1.48	**Sulfate transporter**
At4g04760	3.28	1.57	**Putative sugar transporter**
At1g22990	1.83	1.70	Metal binding protein "putative"
At1g52410	2.38	1.99	thylakoid membrane Ca ion binding
At1g13990	2.99	2.47	Chloroplast molecular_function
At5g48850	3.07	3.26	MS5 family
At3g61930	6.01	4.13	N-terminal protein myristoylation
At3g25260	5.53	5.69	**Nitrate transporter**
At3g30720	12.81	9.58	**Oligopeptide transporter**

R= 0.98. The highlighted genes are potentially involved in nutrient remobilization or cell membrane degradation.

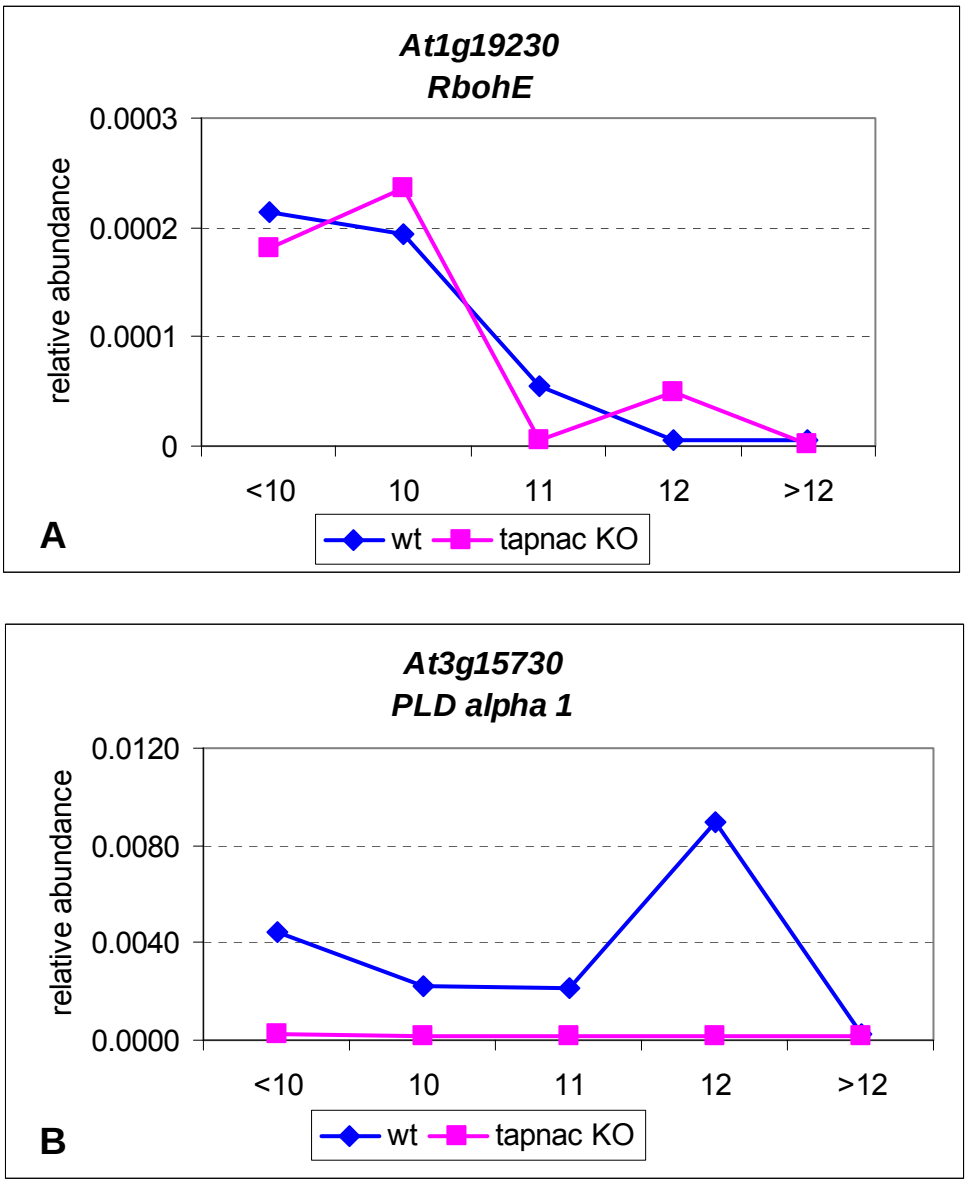


Figure 3.6 Temporal expression of *RbohE (At1g19230)* and *PLDα1 (At3g15730)* transcripts in *Arabidopsis* wild type and *tapnac* KO.

Flowers of different floral stages (<10 to >12) were collected and the levels of specific transcripts were evaluated with qPCR. The relative abundance of the transcript obtained by $2^{-(\Delta Ct)}$ was graphed for every floral stage in the wild type and mutant plants.

qPCR

Based on the list of down- and up-regulated genes, we chose 18 genes that were differentially expressed in the *tapnac* mutant. These genes were analyzed by qPCR in order to validate the microarray data. qPCR results confirmed that these genes change expression levels when the TAPNAC protein was not present. The qPCR data correlated very well with the microarray results with a correlation coefficient of 0.98 (Table 3.3). qPCR was also performed on samples taken from other time points during floral development to determine the kinetics of these genes' expression in relation to *TAPNAC* gene expression.

Total RNA from different floral stages (from <10 up to >12) was assayed by qPCR. Two of the most down-regulated genes, *PLDα1* and *RbohE* (Table 3.3), were evaluated. *PLDα1* shows a genuine down regulation in the *tapnac* KO over the time points evaluated when compared to wild type plants. Remarkably, *PLDα1* expression peaked at floral stage 12 in the wild type background and the transcript disappeared at a later stage when the tapetum began to be degraded. The *RbohE* transcript is down-regulated specifically at floral stage 11 (Fig. 3.6), when *TAPNAC* transcript levels are highest.

Genes classified as transporter genes were also analyzed through the different floral stages. They were chosen because other NAC genes had been shown to play a role in nutrient remobilization (Uauy et al., 2006). The genes *At1g59740*, *At3g30720* and *At3g25260* are oligopeptide transporters. As shown in Fig. 3.7 each gene displayed a different kinetic behavior. *At1g59740* was consistently down-regulated in the absence of the *TAPNAC* transcript (Fig. 3.7 A-B), and *At3g30720* and *At3g25260* were up-regulated with different kinetics (Fig. 3.7 C-D and E-F).

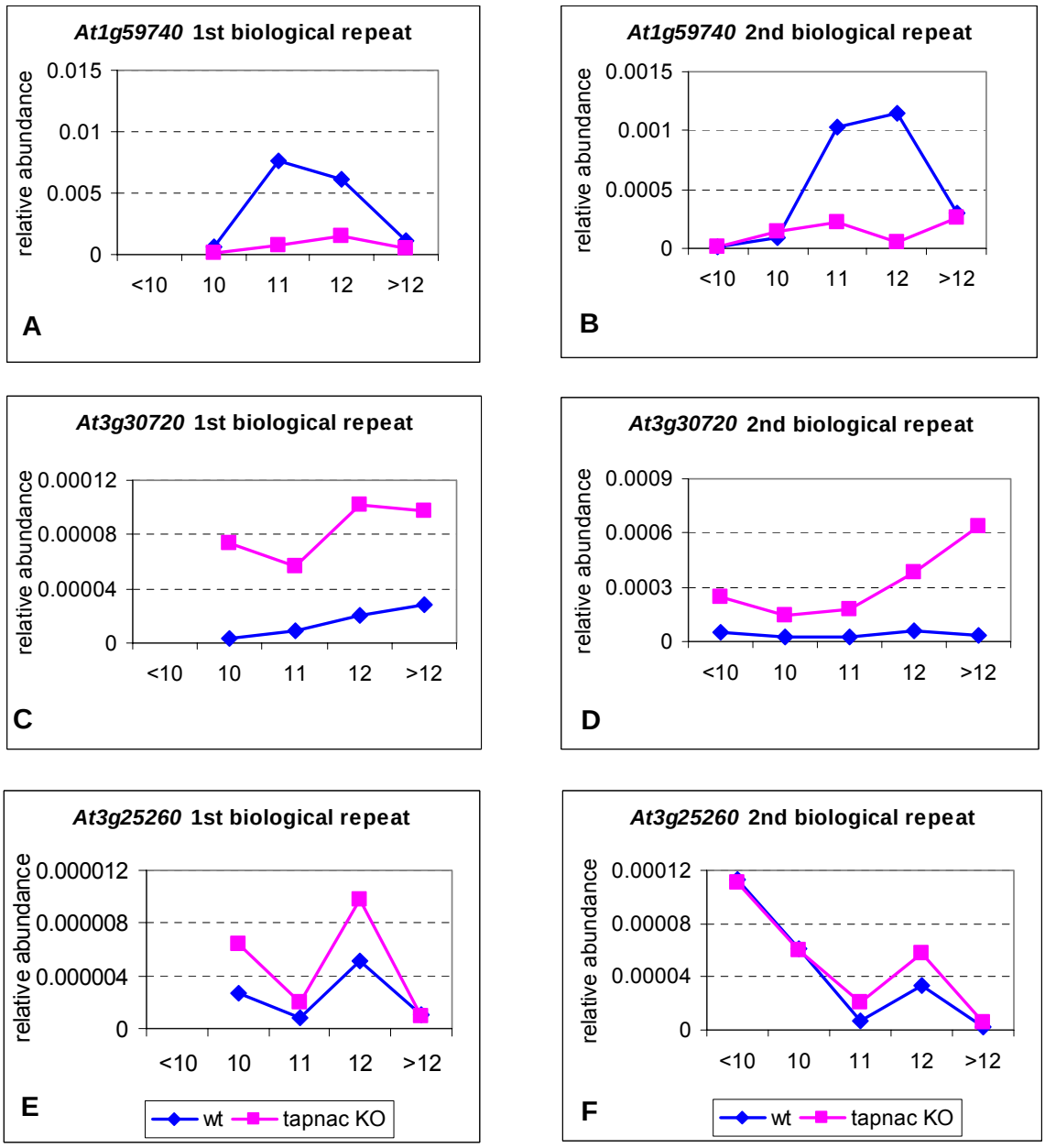


Figure 3.7. Temporal expression of oligopeptide transporter genes in *Arabidopsis* wild type and *tapnac* KO.

qPCR was used to assess the transcript levels of oligopeptide transporters across five floral stages. Results of two independent biological repeats per each gene analyzed are shown in the graphs.

More genes were selected for qPCR analysis: *At2g19500*, a putative citokinin oxidase; *At5g61620*, a MYB transcriptional activator; *At5g47330*, Palmitoyl-(protein) with hydrolase activity; At3g61930, an N-terminal protein myristoylation; and *At1g13990*, located in the chloroplast with high sequence similarity to a zinc ion binding protein.

The first three genes were down-regulated in *tapnac* KO (Fig. 3.8 A-F), and the other two were up-regulated (Fig. 3.8 G-J). Over the course of different floral developmental stages, stage 11 seems to be the most critical for regulation of these genes. This floral stage is characterized by the elevated expression of TAPNAC transcription factor, which may affect directly or indirectly (through interactions with other transcription factors and proteins) the expression of these genes.

In silico analysis of promoter regions of TAPNAC regulated genes

Bioinformatics tools were used to analyze the promoter regions of the *TAPNAC* regulated genes. We wanted to identify *cis*-regulatory sequences that are shared by *TAPNAC* regulated genes (Appendix tables A.3 and A.4). Upstream sequences of 1000 bp with their respective 5' UTR for all the TAPNAC regulated genes identified in the ATH1 chip were collected from the TAIR website. The sequences were analyzed with Weeder, a general purpose motif discovery program (Pavesi et al., 2004). The putative motifs identified with Weeder were examined with PLACE, a database of plant *cis*-acting regulatory DNA elements (Higo et al., 1999) to determine if there were known motifs within these putative regulatory regions. The motifs identified through Weeder software are shown in Table 3.4.

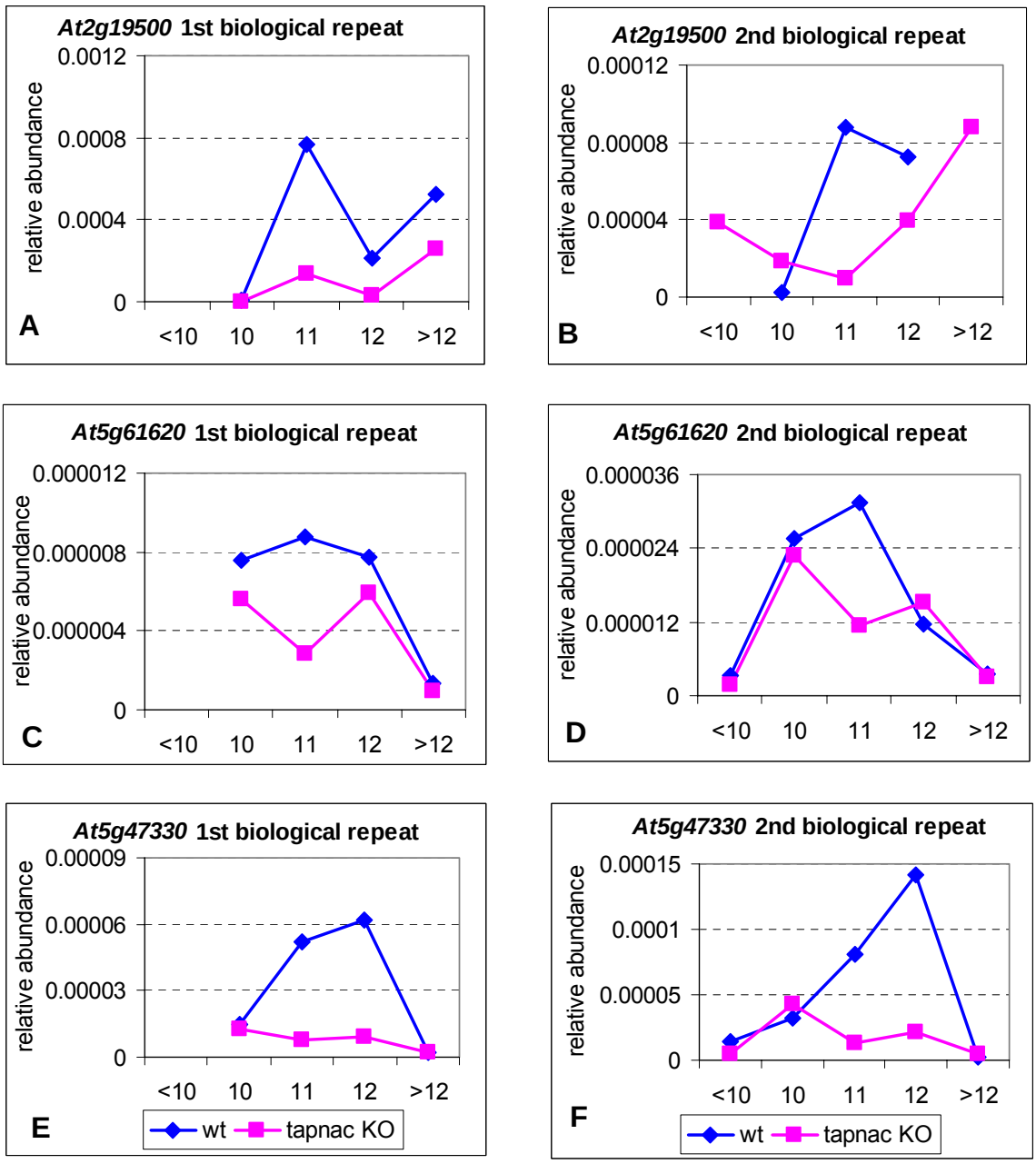


Figure 3.8 Expression profile of selected genes during floral development in *Arabidopsis* wt and *tapnac* KO plants.

qPCR expression analysis of **A-B,**: *At2g19500* (a putative cytokinin kinase); **C-D**, *At5g616201* (a MYB transcriptional activator); **E-F**, *At5g47330* (palmitoyl protein with hydrolase activity); Results of two independent biological repeats per each gene analyzed are shown in the graphs.

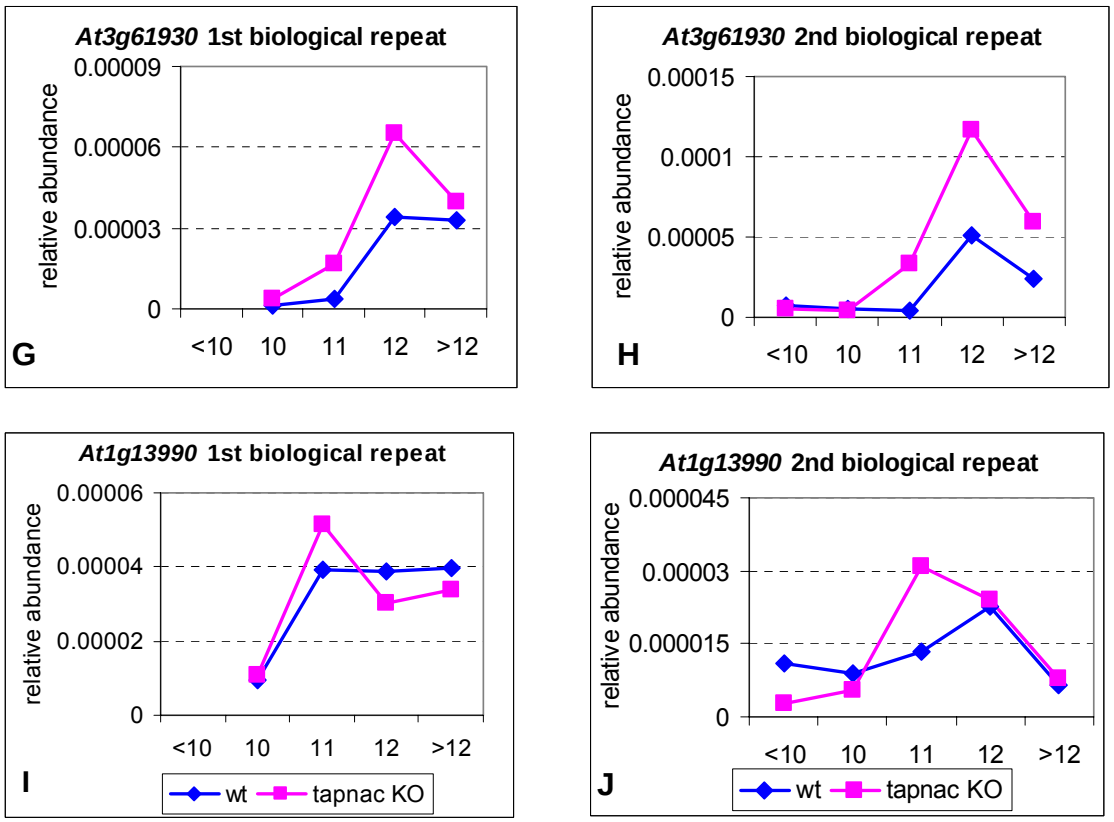


Figure 3.8 Continued.

G-H, *At3g61930*, an N-terminal protein myristoylation; and **I-J,** *At1g13990*, a protein similar to zinc ion binding protein.

All up-regulated genes in the *tapnac* KO contained an ACGT core motif. This core motif could be regulated by ABA signaling and drought stress. The up-regulated genes also had two motifs regulated by micronutrients. Motif 2 contains the IRO2OS regulatory sequence; this is a motif found in a rice transcription factor (OsIRO2) that is induced exclusively by iron deficiency. Motif 3 contains the GTAC sequence that is a copper-response element (CURECORECR) found in cytochrome c6 (Cyc6) and coprogen oxidase (Cpx1) genes in *Chlamydomonas* (Quinn and Merchant, 1995; Quinn et al., 2000).

Table 3.4. Putative *cis*-regulatory elements identified in the *TAPNAC* regulated promoters.

	Motifs	Genes containing particular motif (%)	Associated function/related motif
	Up regulated genes		
Motif 1	ACGTGG CCACGT	100%	ACGT core
Motif 2	CCACGTGA TCACGTGG	76%	ACGT core **GTGANTG10** **IRO2OS**
Motif 3	TGCGTGGTAC GTACCACGCA	49%	**CURECORECR**
Motif 4	AGTTAGGTGCAC GTGCACCTAACT	56%	MYBPLANT
	Down regulated genes		
Motif 1	GTGTGA TCACAC	100%	**GTGANTG10**
Motif 2	TAAGGTAC GTACCTTA	86%	**CURECORECR**
Motif 3	CGTGTAGGGC GCCCTACACG	37%	Novel motif
Motif 4	GACGAGAGGACT AGTCCTCTCGTC	45%	Novel motif

The promoter regions of all down-regulated genes display a *cis*-regulatory sequence that is similar to a late pollen gene g10 (GTGANTG10) found in tobacco. Promoters of 86% of the down-regulated genes contained the copper-response element (CURECORECR) motif. Moreover, two novel motifs were identified, motif 3: CGTGTAGGGC and motif 4: GACGAGAGGACT, were present in 37% and 45% of the genes respectively; these motifs have not been associated with specific gene regulatory functions (Table 3.4).

Overexpression of TAPNAC

Ectopic expression of transcription factors can often provide insights into their biological function if it results in a visible or assessable phenotype. We have taken a similar approach to investigate whether ectopic expression of *TAPNAC* results in any visible phenotype that may provide some indication of its biological function.

The *TAPNAC* gene under the control of a *35S* CaMV promoter was introduced into *Arabidopsis* Columbia ecotype. Resulting transgenic plants were analyzed for transgene expression, T-DNA copy number and phenotypic abnormality. 30 independent lines were generated, but only five plants expressed the *TAPNAC* transgene in leaves. The expression levels differed between lines, presumably due to positional effects and transgene copy number (Fig 3.9-A). Southern blot analysis of these five overexpressing lines identified three lines bearing a single T-DNA insertion (Fig 3.9-C). Coincidentally, these were the same lines that displayed high levels of *TAPNAC* cDNA transcripts in leaves.

Phenotype analysis: Plants with a single T-DNA overexpression (OE) insertion were propagated and the T3 generation was used to determine if the plants were homozygous for the T-DNA insertion. Once homozygous plants were selected, they were planted in soil together with the wild type control plants. The OE lines germinated at the same time as the wild type plants; however, a difference in plant size was evident. Fig. 3.10 shows taller and more vigorous wild type control plants than the two OE lines, OE8 and OE23.

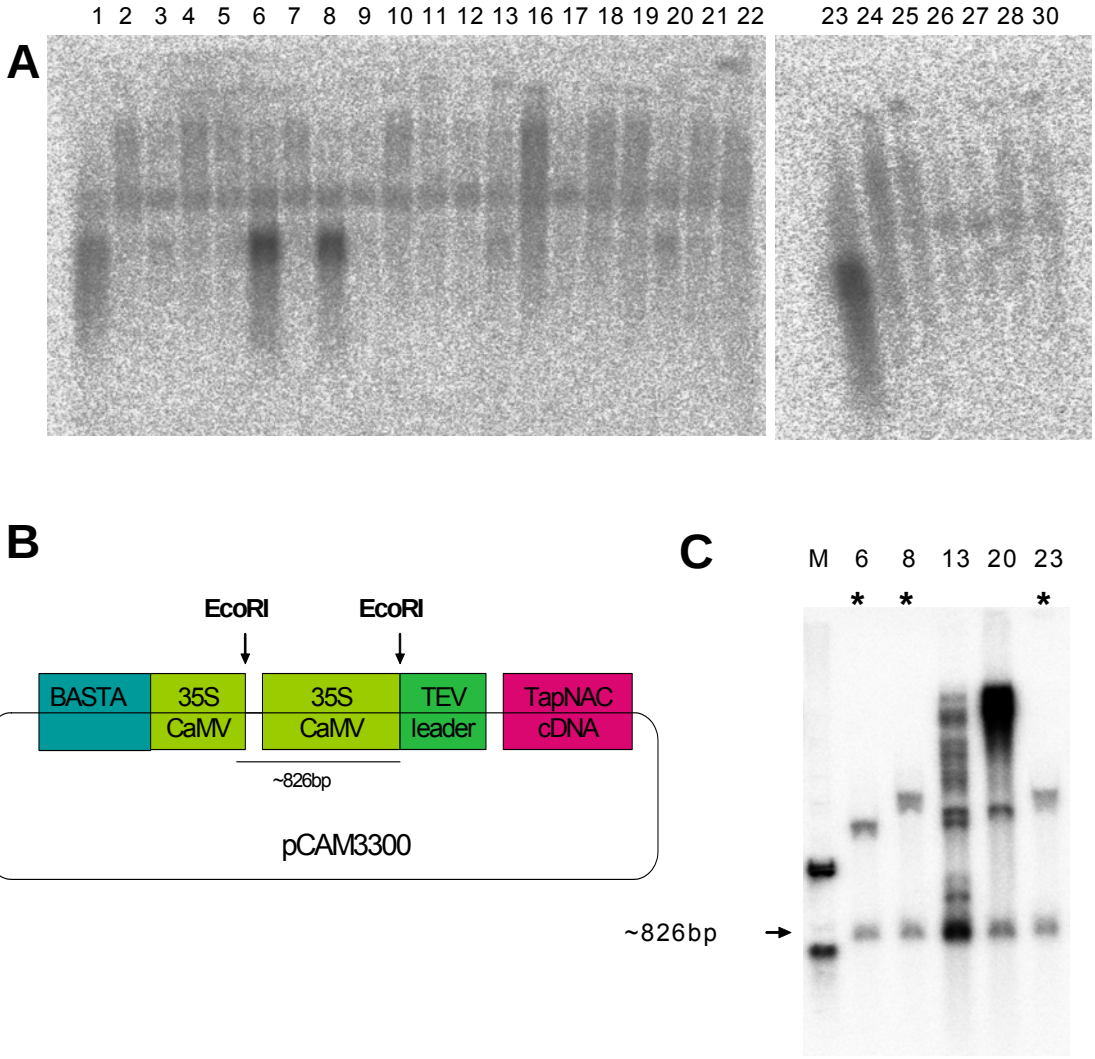
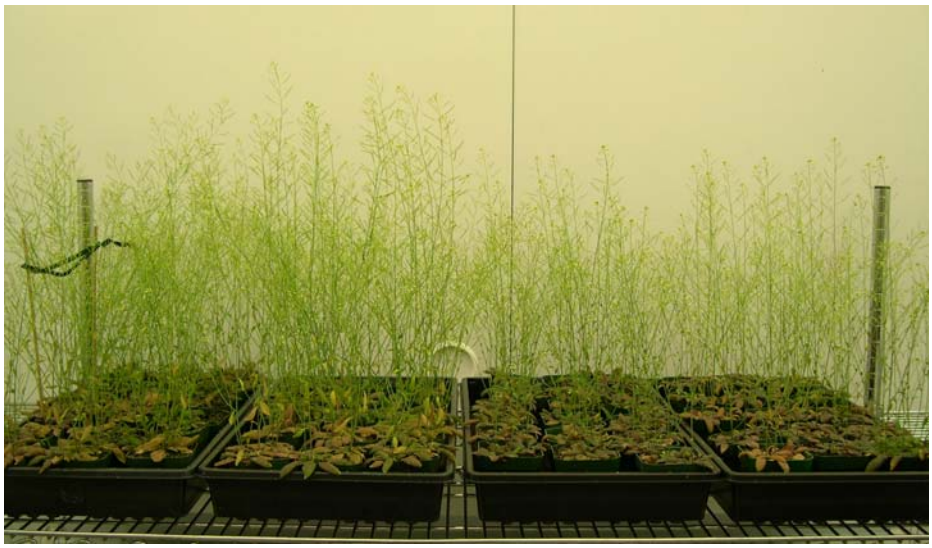


Figure 3.9 Ectopic expression of *TAPNAC* cDNA driven by *35S* CaMV promoter.

A. Northern blot analysis of BASTA resistant lines. 5 ug of leaf total RNA were used to examine the expression levels of the *TAPNAC* cDNA.
B. Diagram of the pCAMBIA 3300 binary vector containing the pRTL2-*TAPNAC* cDNA overexpression cassette.
C. Southern blot analysis of the lines selected from A. ~500 ng of genomic cDNA were digested with **EcoRI** and probed with the promoter region. *, indicates T-DNA single insertion lines.

A



WT OE

B

WT OE8

Figure 3.10. Morphological phenotype of the overexpression lines.

A. Wild type and OE lines shown as a group in the convironment chamber.
B. Wild type plants compared to OE line 8

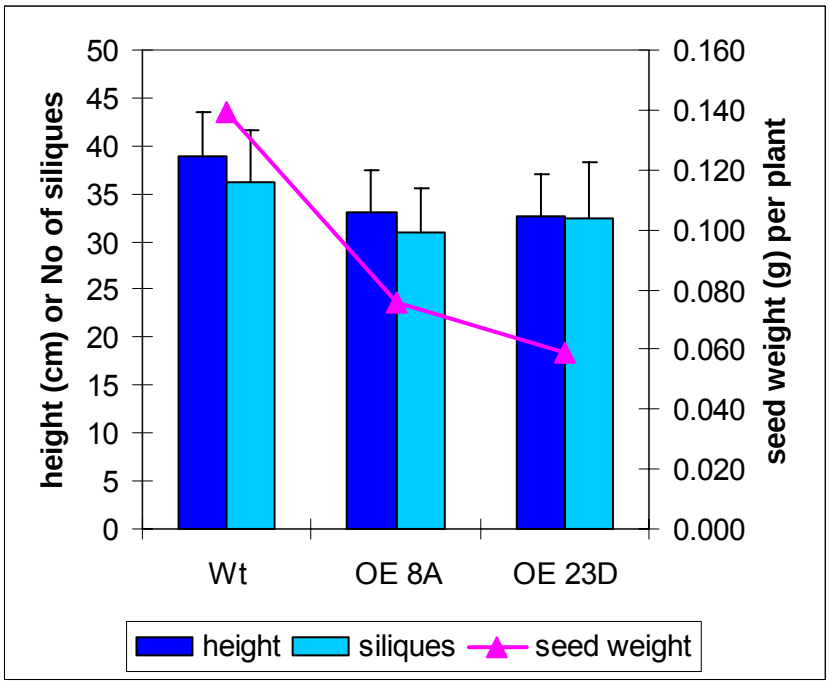


Figure 3.11. Morphological differences between wild type plants and overexpression lines.

Plant height (cm), silique number and seed weight per plant were recorded and the average values are represented in the graph.

Table 3.5. Analysis of nitrogen, zinc and iron contents in seeds of wild type plants and overexpression lines.

Sample	Protein (%)	stdv	Zinc (ppm)	stdv	Iron (ppm)	stdv
Wt	15.06	0.21	43.37	2.57	66.93	2.80
OE 8	15.38	0.14	44.00	2.66	66.97	3.94
OE 23	15.30	0.62	41.93	3.89	66.83	7.53

Protein was determined by combustion AOAC 990.03 method, and the zinc and iron content were determined with atomic absorption spectroscopy AOAC 968.08 method

Plants height was measured and the number of siliques in the main stem were counted. In addition seed weight per plant was determined. Wild type plants were apparently more vigorous than the transgenic lines. Wild type plants were 15% taller and had 14% more siliques than the overexpression lines, wild type plants produced 50% more seeds (Fig. 3.11).

We wanted to know if the change in plant height and amount of seeds set in the transgenic plants could be explained by a change in the amount of nutrients remobilized from leaves (source) to seeds (sink). Therefore, the nutrient content of the seeds was evaluated for nitrogen, zinc and iron. All the samples evaluated (wild type and transgenic lines) showed similar nutrient contents; the nitrogen content measured through protein level was on average 15%; the zinc and iron were 43 ppm and 66 ppm respectively (Table 3.5), indicating that these nutrients were not affected in the seeds by overexpression of TAPNAC protein.

Dominant suppressor

Repression domains of the class II ETHYLENE–RESPONSIVE ELEMENT BINDING FACTOR (ERF) and TFIII A-type zinc finger repressors of transcription that include SUPERMAN (SUP) contain the EAR motif (ERF-associated amphiphilic repression). Short peptides containing the EAR motif when fused to activators of transcription change the behavior of the activator into strong repressors (Hiratsu et al., 2003). Repression activity of the EAR motif was demonstrated by fusing the EAR-motif to known functionally redundant transcription factors (*i.e*. *CUP SHAPED COTYLEDON 1*), and these chimeric proteins acted as dominant repressors, effectively suppressing their target genes (Hiratsu et al., 2003). Based on the success of these experiments, we

fused the EAR motif to the 3' end of the TAPNAC protein and evaluated the effects of this chimeric protein in transgenic plants.

After several trials using different constructs we evaluated the effect of adding the dominant suppressor at the end of the cDNA coding region of TAPNAC driven by its own TAPNAC $-_{1800}$ promoter region. We obtained approximately 18 lines that were BASTA resistant, of those only eight lines expressed *TAPNAC* mRNA in the floral tissue (data not shown). The transgenic plants were analyzed by segregation of BASTA resistance gene and three lines were found to contain a single T-DNA insertion. The selected chimeric transgenic plants were grown along with wild type plants under the same environmental conditions. We inspected them weekly through their life cycle, but no morphological differences were detected under the conditions studied.

Identification of proteins that interact with TAPNAC

We characterized the spatial and temporal expression of *TAPNAC*. However, the use of *tapnac* KO and overexpression lines were not sufficient to understand the role of TAPNAC in the *Arabidopsis* anther tapetum. Knowing the proteins that interact with TAPNAC could help in bringing a better understanding of the function of this gene. Therefore, we attempted to identify potential interaction partners of TAPNAC protein using the yeast two hybrid system.

Two hybrid assays can identify protein-protein interactions in yeast. The system is based on the expression of a "bait" protein fused to the GAL4 DNA-binding domain, cDNAs encoding putative interacting proteins are fused to the GAL4 DNA-activation domain. When the bait and library fusion proteins interact, the DNA binding domain and activation domain activate four reporter genes: *HIS3,*

ADE2, lacZ and *MEL1*. Yeast cells that grow on media lacking histidine and adenine and turn blue at the same time, probably contain the interacting proteins.

The first step on the yeast two hybrid system is the construction of a cDNA library fused to the GAL4 DNA-activation domain. The library was made from total RNA of anthers that belong to floral stages 10 to 12, and the library transformation efficiency was 1.06 x 10^6 transformants / 3 µg pGADT7-Rec. The *TAPNAC* cDNA derived from the pUIN51 vector (ABRC) was cloned into the pGBKT7 vector. TAPNAC was the bait protein fused to the GAL4 DNA-binding domain. Screening was performed by yeast mating. The selection of positive clones was done on quadruple dropout (QDO) media and was then replicated in QDO media containing XαGal for identification of false positives.

The selected clones were screened by PCR to verify single plasmid transformation in yeast cells. Primers flanking the pGADT7 vector were used in this PCR screening; clones that yielded a single PCR product were sequenced to identify the putative interacting proteins. A list of the clones and their gene identity are listed in Table 3.6. The most abundant gene identified among the putative positive clones was the transcription factor *ARABIDOPSIS RESPONSE REGULATOR 2* (*ARR2*). We further analyzed this putative protein-protein interaction along with two other clones: clone 9, *At1g04080* and clone 28, *At1g61110* (Table 3.6).

We retested both the phenotype in QDO media containing XαGal and the interaction in yeast. When the phenotype was retested in QDO XαGal media, we noticed that several of the ARR2 clones did not grow in the fresh media, and several showed a mixture of white and blue colonies. In order to retest the interaction in yeast, selected plasmids were retransformed into AH109 yeast

cells. The new transformed cells were mated with three different plasmids: DNA-BD empty vector, DNA-BD bait and pGBKT7-Lam and then plated on media lacking leucine and tryptophan; all the diploid mated cells grew on these plates. A replicate of the diploid cells was done on QDO media containing XαGal, under these conditions only the diploid cells containing the "bait", TAPNAC, and the protein interacting partner should grow and display a blue color. Out of three clones tested, only one, clone 9 (At1g04080) interacted with TAPNAC protein (Fig. 3.12). At1g04080 contains a half tetratricopeptide repeat (HAT) helix domain. The HAT-repeat-containing proteins appeared to be components of macromolecular complexes that are required for RNA processing.

Table 3.6. Yeast two hybrid screening.

Clone number	Function
9	RNA-processing protein, HAT helix . At1g04080
10	ARR2 At4g16110
11	ARR2
12	ARR2
14	ARR2
16	ARR2
19	ARR2
21	ARR2
24	ARR2
28	TAPNAC At1g61110
46	Glutamate ammonia ligase ATGSKB6 At3g17820
47	ARR2
48	Aspartic protease At4g04460
52	Aspartic protease At4g04460
56	Aspartic protease At4g04460
57	Glutamate ammonia ligase ATGSKB6 At3g17820

Putative positive clones are listed.

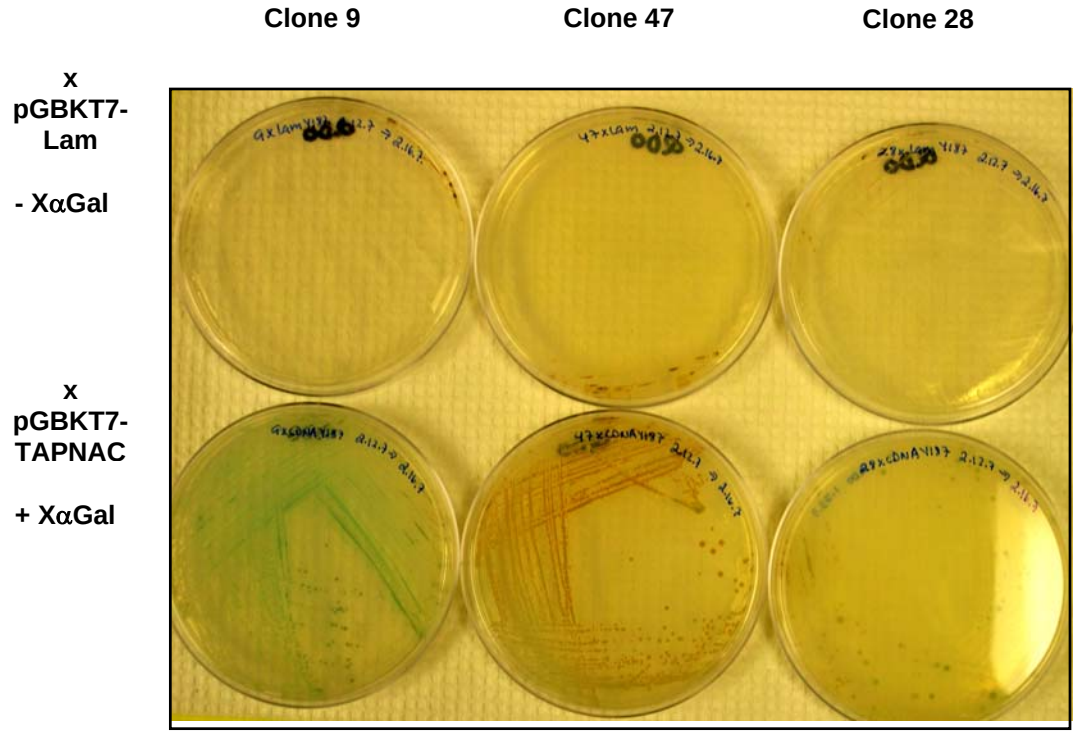


Figure 3.12. Retesting protein interactions in yeast.

Yeast mating was used to verify protein interactions. Selected clones isolated from the pGADT7/library were retransformed into AH109. Diploid cells of the yeast cell matings indicated (pGADT7 x pGBKT7-Lam and pGADT7 x pGBKT7 TAPNAC) were plated in QDO media to verify protein-protein interactions.

Discussion

There is substantial information on the morphological and structural changes that occur during tapetum development; however, the regulatory networks involved in these changes are not completely understood. TAPNAC, a NAC transcription factor, may play a role in the regulation of tapetum development; therefore, information on the precise temporal expression of *TAPNAC* gene is

fundamental to reveal its role in tapetum development. Furthermore, the use of mutant lines that disrupt the function of TAPNAC or overexpress *TAPNAC* may be important in the discovery of the function of TAPNAC in *Arabidopsis*.

A T-DNA insertion line that disrupts the *TAPNAC* was identified in the SALK Research Institute T-DNA collection. The SALK_060459 line was obtained, and homozygous single insertion plants were identified, propagated and used in subsequent analysis.

The temporal expression of *TAPNAC* in the mutant and wild type backgrounds was obtained by qPCR. The expression of the *TAPNAC* transcript in the wild type plants increases from floral stage <10 up to floral stage 11 (Fig 3.1) where it displayed its maximum level of expression. Thereafter, a decrease in expression was observed with later floral developmental stages. This decrease in expression correlates with the degradation of the tapetal cells, the same cells that harbor the expression of *TAPNAC*. This data implies that TAPNAC may be important in regulating events that will occur in the tapetum during floral stage 11 and 12. Given that *TAPNAC* starts being expressed at stage 10 and reaches maximum expression at stage 11, its effect should be exerted on proteins that are required at subsequent stages (stage 11 and 12).

Morphological changes during flower and anther development are well known and were described in detail in Chapter I. Briefly, there are two phases during anther development. Phase I involves cell division and differentiation that results in the formation of anther tissues (stages 1-8); and phase II is characterized by microspore development and pollen release (stages 9-14). The second phase will be the focus of our discussion. Phase II begins at floral stage 10 when *TAPNAC* transcripts start to accumulate in the tapetum, at this time the tapetal cells become active in the formation of the cytoplasmic lipid bodies: the

elaioplast and the tapetosome (Hsieh and Huang, 2005, 2007). The elaioplast is a plastid that houses the steryl esters, and the tapetosome, a cytoplasmic lipid body that contains triacylglycerols, oleosins and flavonoids. The steryl sters will help later in conferring pollen dessication tolerance, and the oleosins will facilitate water uptake for pollen germination on the stigmatic surface. Meanwhile thinning of the tapetal cells, whose cell walls are not well developed begins (Wu and Cheung, 2000), and this thinning aids in the secretion of sporopollenin precursors (a mixture of fatty acids and phenylpropanoid compounds) that will be deposited in the exine layer of the pollen grain. By the end of stage 10, the pollen grains look bigger, and the walls have thickened due to the formation of exine.

TAPNAC transcript levels increase to a maximum by stage 11. At stage 11 the pollen grains will go through their first mitotic division, the energy required for this process must be provided by the nutrients being secreted from the tapetum. Another morphological change that can be observed during floral stage 11 is pollen coat formation. The pollen coat starts to be formed as a result of the secretion of lipids, oleosins and flavonoids from the tapetal cells. The membranes of tapetal cells begin to disintegrate, and all the tapetal cells contents are finally released into the anther locule by the end of floral stage 12. At this stage, the levels of expression of *TAPNAC* transcript decreases and at a later stage the *TAPNAC* gene is not detected, while the tapetal tissue that harbors *TAPNAC* expression is disintegrated.

Since no morphological phenotype in the *tapnac* KO plants was detected under the environmental conditions tested, we decided to do a global transcriptome analysis in order to get more information on the genes potentially regulated by the TAPNAC transcription factor. ATH1 Affymetrix gene chip analysis on RNA from tissue of floral stage 11 of wild type and *tapnac* KO plants, revealed that 94

genes were differentially expressed (Figs. 3.4 and 3.5). From this list, 18 genes were selected for verification of gene expression via qPCR. Table 3.3 shows the ratios reported in the microarray experiments and the ratios obtained with qPCR. The correlation coefficient for these two set of values was 0.98, a good indication that both platforms were producing consistent results. Furthermore, qPCR was performed on samples taken from other time points during floral development (stage <10 to stage >12) to determine the kinetics of these genes in relation to *TAPNAC* gene expression. Based on the results obtained, we were able to identify potential TAPNAC target genes.

Several transporter genes (sugar, iron and nitrate) were found to be affected at the level of mRNA expression in the *tapnac* KO (Figs. 3.7 and 3.8). TAPNAC could be important for the regulation of nutrient remobilization at floral stage 10 and 11, since it is at these stages when the pollen grains go through their first mitotic division (high energy demand) and also start the maturation phase. NAC genes in wheat have recently been implicated in senescence and rapid nutrient remobilization from leaves to the developing grains (Uauy et al., 2006). The wheat NAC gene was named *TaNAM-B1* due to its high sequence similarity to a clade of three *Arabidopsis* NAC proteins that included ATNAM (Duval et al., 2002), TAPNAC and At3g15510. Consequently, one of the functions of TAPNAC could be nutrient remobilization in the tapetal cells. Additionally, the *in silico* analysis of the promoter regions of the genes affected by the disruption of *TAPNAC*, identified motifs regulated by micronutrients such as iron and copper (Table 3.4) supporting the idea that TAPNAC is implicated in nutrient remobilization in the tapetal cells by means still unknown.

Also, the overexpression of *TAPNAC* in the two transgenic lines showed a difference in plant growth and seed yield. The overexpressing lines were shorter and have 50% less seed production than the wild type control plants. Therefore,

we analyzed the seed's nutrient content of these overexpression lines with the assumption that overexpressing *TAPNAC* in the leaves could have changed the amount of nutrients transported from the leaves to the seeds, reflected in the amount of seeds set by the overexpressing lines. The presence of TAPNAC protein in the leaves could have impaired the normal function of *ATNAM* and *At3g15170*, the two other NAC genes present in leaves and whose putative role is nutrient remobilization of nitrogen, zinc and iron (Uauy et al., 2006). However, no changes in nitrogen, zinc or iron content were observed on the seeds of these transgenic lines. This result indicates that the nutrients affected by the overproduction of TAPNAC protein in the leaves could be others than the ones analyzed (nitrogen, zinc and iron). Other important nutrients that start to be accumulated in the tapetal cells at the time of *TAPNAC* expression are the steryl esters, triglycerides, flavonoids and oleosin proteins. Some of these compounds are also accumulated in *Arabidopsis* seeds, therefore we would have to analyze lipids and oleosin protein content in the seeds produced by the overexpressing lines.

PLDα1 was another putative target gene of TAPNAC. This gene displays a drastic change in expression when TAPNAC protein is disrupted in the mutant background. Figure 3.6 clearly shows the behavior of *PLDα1* in the wild type plants: its transcript is highly expressed at floral stage 12, following *TAPNAC* maximum level of expression (floral stage 11). We hypothesize that TAPNAC protein could activate expression of PLDα1 in the tapetum at floral stage 12. Therefore the PLDα1 enzyme could play a role at this point since the tapetal cell membranes will have to be degraded and the lipids and oleosins inside the cytoplasmic lipid bodies need to be released and deposited into the pollen coat.

Why could PLDα1 be so important in the regulatory networks between TAPNAC and tapetum development? PLDα1 is a hydrolyzing enzyme that belongs to the

Phospholipase D protein family (12 genes). PLD enzymes hydrolyze membrane lipids to produce phosphatidic acid (PA) and a free head group (*i.e*. choline). PLD enzymes have been implicated in a variety of stresses including osmotic stress and wounding. PLD can regulate cell function by different modes of action such as: (1) production of lipid mediators, i.e. PA, N-acylethanolamine, (2) direct interaction with other proteins such as actin and tubulin, (3) alteration of the lipid composition in the membranes and (4) membrane degradation.

We are mostly interested in the involvement of PLDα1 in membrane lipid composition and membrane degradation. A role in membrane lipid composition has been documented for PLDα1 in response to freezing (Welti et al., 2002) In this study, lipid profiles of wild type plants and PLDα1 deficient plants that were subjected to cold and freezing stresses were compared. The results revealed that PLDα1 selectively hydrolyzes phosphatidylcholine (PC) and alters the ratios of membrane lipids and contributes to membrane lipid degradation. Another study on the lipid profiling of wild type plants compared to a *pldα1* mutant KO (Devaiah et al., 2006), revealed that the flowers were the only organ in which the wild type genotype had a significantly higher level of total polar lipids than the mutant genotype (at $p<0.05$). Additional evidence for the involvement of PLDα1 in the degradation of membranes and changes in lipid composition is in the studies done on seed viability of *pldα1* KO and antisense lines. The results of these studies showed that seeds from *pldα1* knockdown and KO plants were more resistant to aging than the seeds from wild type plants, indicating that high levels of PLDα1 in *Arabidopsis* seeds are harmful since it will induce seed aging and deterioration. The study suggests that the formation of PA by PLDα1 initiates a series of reactions that will damage cell membranes and storage lipids in seeds. The activity of PLDα1 could also destabilize oil bodies that consist of triacylglycerol, coated by a phospholipid monolayer containing proteins (Devaiah et al., 2007).

Based on the expression profiling shown in Fig. 3.1 we know that *TAPNAC* gene expression begins at early stage 10, peaks at floral stage 11 and decays at stage 12. Interestingly, PLDα1 expression peaks at floral stage 12 (Fig. 3.6), the same stage when the tapetum starts to be degraded and its contents are released into the anther's locule. All of these facts suggest that PLDα1 could probably exert its phospholipase activity in the membranes of the tapetal cells or in the phospholipid layer embedded in the tapetosome. Additionally, the expression of PLDα1 is not detected in mature pollen grains (http://jsp.weigel-world.org), supporting the idea that is mainly expressed in the anther tapetum and probably regulated by TAPNAC protein.

There are other roles for PLDα1, especially in relation to drought stress. One study showed that PLDα1 derived PA interacts with ABI1 phosphatase 2C, a negative regulator of ABA responses (Zhang et al., 2004) promoting ABA signaling. Plants with lower levels of PLDα1 are insensitive to ABA and have impaired stomatal conductance (Sang et al., 2001). Besides its role in phospholipid degradation, PLDα1 could also be involved in signaling other responses through the PA molecule released as a consequence of its phospholipase activity.

Dominant repressor technique

Given that no morphological phenotype was observed under the conditions studied, probably due to functional redundancy, we expected to see an effect on the plants when using a chimeric dominant repressor (TAPNAC fused to the EAR-motif). This technique has been successfully used with another NAC gene, *CUC1*, where the chimeric plants were able to phenocopy the *cuc1cuc2* double mutant plants (Hiratsu et al., 2003). We hypothesized that the TAPNAC-EAR

motif would be able to repress TAPNAC target genes resulting in a visible morphological phenotype.

We fused the 12 aa sequence EAR-motif to the 3' end of TAPNAC and added a stop codon at the end of the EAR motif. Previous chimeric constructs were done with the cDNA including the 5'UTR of *TAPNAC* or a *35S* CaMV promoter region. The problem with these two constructs was that the 5'UTR included two initiation codons prior to the actual start codon, so this construct may not be very efficient in TAPNAC protein translation. The second construct was driven by the *35S* CaMV promoter region, a problem since the *35S* promoter was not able to drive expression of *TAPNAC* in the tapetum. Therefore, we had to create a construct that has the $TAPNAC_{-1800bp}$ promoter region (Chapter IV), and the *TAPNAC* coding region (cDNA) obtained through ABRC.

tapnac KO plants were transformed with these chimeric protein constructs, and the plants bearing a single T-DNA insertion were planted in soil together with wild type control plants. We examined plants specifically for male sterility and decrease in seed production under the assumption that impairing nutrient remobilization from the tapetum would affect the normal development of pollen grains. Pollen coat composition could also be affected to a point that pollen germination on the stigma will be impaired, leading to less seed production. However, no morphological phenotype was observed under the conditions studied. Future tests on pollen germination under different water potential conditions will be performed, since a change in the pollen coat composition could affect both pollen desiccation tolerance and pollen hydration.

The search for interacting proteins with the yeast two hybrid system identified *At1g04080*, a HAT helix domain / RNA processing protein, as one of the strongest interacting partners. Nevertheless, the most frequent clone identified

during the first screening, ARR2 transcription factor, resulted in a “false positive”, although mated yeast cells carrying these *ARR2* gene still grew in QDO media, but did not turn blue. Possible explanations for these findings are: multiple plasmids contained in individual yeast cells used in the first screening that confounded the results, streaking the cells two or three times in QDO media could help in isolating real interacting partners. It is also possible that rearrangements into the yeast genome, made possible the activation of the other two reporter genes (ADE2 and HIS3) but not the activation of MEL1. Moreover, since we are testing in this particular case two transcription factors (TAPNAC and ARR2), it may have been a third protein involved in the activation of all three reporters tested (*ADE2*, *HIS3* and *MEL1*) in the first screening, but once the third protein was eliminated in the second screening the putative protein–protein interaction between TAPNAC and ARR2*,* disappeared. More verification tests will be done on the remaining clones, and new screens will have to be performed.

In summary, we attempted to assess *TAPNAC* gene function via various reverse genetic approaches. Functional redundancy did not allow us to observe a morphological phenotype when the *TAPNAC* gene was knocked out or fused to a chimeric-EAR dominant repressor, although, we do not discard the possibility of finding a phenotype in a different environmental condition, *i.e* lower relative humidity. The molecular phenotype analyzed with the use of ATH1 affymetrix chips, resulted in the identification of potential target genes that are involved in nutrient transport and cell membrane lipid degradation (PLDα1). Given that *TAPNAC* is highly expressed during floral stage 11, when energy demands are high due to pollen division (mitosis) and pollen coat formation; it is reasonable to hypothesize that *TAPNAC* protein is involved in regulating nutrient remobilization both by regulating expression of nutrient transporter genes and by thinning the tapetal cell membrane and disrupting the lipids enclosed in the tapetosome.

CHAPTER IV

IDENTIFICATION OF TAPETAL SPECIFICATION *CIS*-REGULATORY MOTIFS IN THE TAPNAC PROMOTER

Introduction

The *TAPNAC* gene is a floral specific NAC gene that was identified through microarray and qPCR data analysis (Chapters II and III). Moreover, previous studies performed by Wellmer *et. al*. (2004) detected *TAPNAC* transcripts in the anther tapetum. The tapetum is the tissue that surrounds the pollen sac and is responsible for nurturing the pollen grains.

The importance of tapetal cells in the development of pollen grains has been demonstrated in several studies (Kaul, 1988). Mariani et.al, (1990, 1992) was able to induce plant male sterility using a transgenic approach. A bacterial ribonuclease gene was fused to the promoter of *TA29* gene (glycine-rich protein), a tobacco tapetal specific gene. The chimeric gene was introduced into tobacco plants, and the expression of this RNase caused developmental arrest of the tapetum and consequently a defect in pollen grain formation. The resulting transgenic tobacco plants were male sterile (Mariani et al., 1990). This technique proved to be a powerful tool that can be used in the production of hybrid seeds. Hybrid seed production is a labor intensive process that usually requires the emasculation of flowers (removal of the anthers) prior to pollination. Male sterile plants can alleviate the seed production labor and possibly reduce the cost of hybrid seeds.

Even though the tapetum is important in pollen grain formation and maturation, and besides the capability of using tapetal specific promoter regions to induce tapetal cellular ablation that results in male sterility; there is no report that clearly identifies a *cis*-regulatory sequence that confers tapetal specific gene expression. Identification of a tapetal *cis*-regulatory motif could increase our understanding of how tapetal expressed genes are regulated, and what are the *trans*-acting factor(s) associated with these motifs? Moreover, a well defined tapetal specification element could be used as a tool to drive expression of specific proteins in the tapetum. These proteins could confer specific characteristics to pollen garins *i.e*: drought resistance, self incompatibility (important for seed hybrid production).

The aim of this part of the study was to identify *cis*-regulatory sequences in the *TAPNAC (At1g61110)* promoter that specify tapetal-specific gene expression. First of all, a *TAPNAC* 1700 bp upstream promoter region was fused to the *β-glucuronidase* (GUS) reporter gene. This construct was designated $Pro_{TAPNAC\text{-}1700}$:GUS. Initial analysis showed that the 1700 bp *TAPNAC* promoter drove GUS expression in the anthers of different floral stages. GUS immunolocalization was used to show that the GUS protein was localized to the tapetum. Later, promoter serial deletion analysis was conducted on the *TAPNAC* promoter and a putative *cis*-regulatory motif that enhances GUS expression in the tapetum was identified.

Materials and methods

Construction of the GUS reporter constructs

Pro $_{TAPNAC\ -\ 1700}$: GUS (Table 4.1) was constructed by amplifying a 1700 bp upstream region (-1700 to +125) using a 5' forward primer, CCGG CTGCAGGACTTGCAAGCATTTTTGTCTTTTCT, and a 3' reverse primer, CCGGCCCGGGAACTTCTAATTTGATATAATGACGATA, which contained PstI and SmaI restriction enzyme sites at their 5' ends, respectively. PCR cycle: 94°C x 3min, 94°C x 50 sec, 59°C x 50 sec. and 72°C for 2min (29 cycles) 72°C 10 min and then 4°C. The PCR product was cloned into TOPO 2.1 (Invitrogen Life Technologies) for sequence verification purposes. The -1700 bp promoter region was then excised with BamHI and SmaI and subcloned into BamH1 and SmaI sites in the pBI101 binary vector (Clontech, Palo Alto, California).

Subsequently, smaller pieces of the *TAPNAC* promoter were studied. Three *TAPNAC* promoter GUS repoter cassettes were constructed: Pro$_{TAPNAC\ -962}$:GUS , Pro$_{TAPNAC\ -596}$:GUS and Pro$_{TAPNAC\ -237}$:GUS (Table 4.1) were fused to a GUS reporter gene in the pBI101 promoterless binary vector (Clontech). The -962 BamHI GGCCGGATCCTGGAGAGT TTGCTAACATAGTGT and -596 HindIII GGCCAAGCTTTGACCTCATATCTC CTGACTCCT primers were used in combination with the 3' reverse primer. The PCR product was cloned into TOPO 2.1 for sequence verification and then subcloned into pBI101 with BamHI-SmaI restriction enzymes for -962 bp promoter and with HindIII-SmaI for the -596 bp promoter. Pro$_{TAPNAC\ -237}$:GUS was constructed by digesting the original Pro$_{TAPNAC-1700}$ in TOPO 2.1 with EcoRI restriction enzyme. Then the digested plasmid was self ligated with T4 ligase (Promega) and then the -237 bp promoter region was sub cloned into pBI101 with BamHI and SmaI restriction enzymes.

Smaller promoter regions within the -237 promoter region were also analyzed. For this purpose we used pBI101 and pBIN19.64 vector, obtained by inserting a CaMV -64. 35S (-64 to +8):GUS fusion into the pBIN19 binary vector (Bevan, 1984). The primers and restriction enzymes used were as follow: -129 HindIII SacI : GGCCAAGCTTGAGCTCCCTGCCTTTTCTCCAA, -1 HindII SacI: GGCC AAGCTTGAGCTCTCTTCTCGCATCTTATAT, -51 HindIII SacI GGCCAAGC TTGAGCTCTTCCTACATATTCTCT, and the 3' end reverse primer. PCR products were digested with either HindIII and SmaI for pBI101 subcloning or with SacI and SmaI for pBIN19-64 subcloning.

Table 4.1. Promoter constructs used in the identification of tapetal *cis*-regulatory sequences.

Construct name	Promoter region
$Pro_{TAPNAC\ -1700}$:GUS	-1700 to +125
$Pro_{TAPNAC\ -962}$:GUS	-962 to +125
$Pro_{TAPNAC\ -596}$:GUS	-596 to +125
$Pro_{TAPNAC\ -237}$:GUS	-237 to +125
$Pro_{TAPNAC\ -217}$:GUS	-217 to +125
$Pro_{TAPNAC\ -189}$:GUS	-189 to +125
$Pro_{TAPNAC\ -150}$:GUS	-150 to +125
$Pro_{TAPNAC\ -129}$:GUS	-129 to +125
$Pro_{TAPNAC\ -1}$:GUS	-1 to +125
$Pro_{TAPNAC\ +51}$:GUS	-51 to +125
$Enh_{TAPNAC\ -962}$:GUS	-962 to -237
$Enh_{TAPNAC\ -596}$:GUS	-596 to -237
$Motif_{TAPNAC\ 20}$:GUS	-237 to -217

The 5'UTR promoter region (+1 to +125) was included in all the constructs used in the serial deletion analysis.

A finer deletion analysis between -237 bp and -150 bp of *TAPNAC* promoter was subsequently performed. The constructs generated were: $Pro_{TAPNAC\text{-}217}$:GUS, $Pro_{TAPNAC\text{-}189}$:GUS and $Pro_{TAPNAC\ -150}$:GUS. The primers used were: -217 HindIII SacI ggc CAAGCTTGAGCTCTTTATAAAATCTTACACAAAAAC, -189 HindIII SacI GGCC AAGCTTGAGCTCTGAGGTATGTGCGTGATG and -150 bp HindIII SacI ggccaagcttgagctccataaggcagttcgaaa aaac in combination with the 3' end reverse primer. The PCR amplified products were digested with HindIII and SmaI and subcloned into pBI101.

Enhancer identification

Putative *TAPNAC* enhancer regions $Enh_{TAPNAC\ -962}$:GUS and $Enh_{TAPNAC\ -596}$:GUS (Table 4.1) were fused to either the putative *TAPNAC* minimal promoter region, $Pro_{TAPNAC\ -150}$:GUS construct, or to the pBIN19.64 CaMV minimal promoter vector. The primers used were -962 5' SacI HindIII GGCCGAGCTCAAGCT TTGGAGAGTTTGCTAACATAGTGT and -596 5' SacI HindIII GGCCGAGC TCAAGCTTTGACCTCATATCTCCTGACTCCT and -237 SacI SmaI GGCCGA GCTCCCCGGGCTCTATAAATATATTTTA or -237 HindIII GGCCAAGCTTC TCTATAAATATATTTTA. The T_m used for the PCR amplification was 55°C. The PCR product was digested with SacI and SmaI and cloned into pBIN19.64 CaMV vector, or excised with HindIII to subclone the fragment into the $Pro_{TAPNAC\ -150}$:GUS construct. Verification of sequences was conducted using either primers pBI101_F TAGCTCACTCATTAGGC and pBI101_R CGCGATCCAGACTGAA for the pBI101 vector or the primers used in the initial PCR amplification (i.e. -962 5' SacI HindIII, 596 5' SacI HindIII and -237 HindIII) for the pBIN19.64 CaMV vector.

The 20 bp region within -237 and -217 bp promoter region, was also fused to the $Pro_{TAPNAC\ -150}$:GUS construct ($Motif_{TAPNAC20}$:GUS) to verify tapetal specificity. The primers used were HindIII-motif– 150: 5' GGCCAAGCTTGAATTCCCACA TTTCTAGTTCATAAGGCAGTTCGAA in combination with the 3' reverse primer, CCGGCCCGGGAACTTCTAATTTGATATAATGACGATA, which contained SmaI restriction enzyme site. The T_m used for the PCR cycle was 52°C and the $Pro_{TAPNAC\ -150}$:GUS construct was used as a template for the amplification. The fragment was cloned into pBI101 with HindIII-SmaI and then this vector was transfected into GV3101 *Agrobacterium* cells.

PCR mutagenesis

The region between -237 and -217 (GAATTCCCACATTTCTAGTTT) contains a potential tapetal specification *cis*-regulatory sequence. In order to identify this motif we mutagenized the sequence between -237 and -217 (GAATTCCCACATTTCTAGTTT) by randomizing the sequence using SHUFFLESEQ (http://xblast.tamu.edu/), while maintaining the same GC content.

Four different shuffled sequences were generated: all shuffle ***CCA TTTTGATACAGCTTTACT***, the first ***AACACTTGC***CATTTCTAGTTT, the second GAATT***AATCTCTCCC***TAGTTT and the last third GAATTCCCACA ***ATCTTTTGTT***. Forward and reverse primers containing 15 bp of unmodified sequence from the flanking regions of the -237bp construct in TOPO 2.1, were designed (Stratagene Quick Change Site directed mutagenesis protocol). The sequences of the primers used were: all) GCCGCCAGTGTGCTGC CATTTTGATACAGCTTTACTTTATAAAATCT, 1st) CGCCAGTGTGCTGAACAC TTGCCATTTCTAGTTTTT, 2nd) CAGTGTGCTGGAATTAATCTCTCCCTAGT TTTTATAAAATC, and 3rd) GTGCTGGAATTCCCACAATCTTTTGTTTTATAA

AATCTTACAC. *Pfu* DNA polymerase was used in the PCR reaction, followed by DpnI endonuclease digestion and electroporation of the digested reaction into DH10B cells. Sequence confirmation was done by using Big Dye reagent. $Pro^{m}_{TAPNAC\ -237}$:GUS reporter cassettes were then transferred to pBI101 using BamH1-SmaI.

Arabidopsis plant transformation and screening

Constructs were transferred into the *Agrobacterium tumefaciens* strain, GV3101, via electroporation. For transformation, a 250 ml of liquid broth [1% (w/v) tryptone, 1% (w/v) NaCl and 0.5% (w/v) yeast extract] containing 50 $ugmL^{-1}$ kanamycin and 50 $ugmL^{-1}$ gentamycin was inoculated with 5 mL overnight culture of *A. tumefaciens* strain GV3101 containing the binary vector with the *TAPNAC* promoter:GUS fusion. Cultures were grown for an additional 16-18 h at 28°C. Cells were harvested by centrifugation at 5000 rpm and then resuspended in 250 ml of 4.3 $mg.mL^{-1}$ Murashige and Skoog media, 5% sucrose, 0.02% (v/v) Silwet L-77 (Lehle Seeds, Round Rock, TX) and 0.001% of benzylamino purine stock solution (10 mg/ml in DMSO). The suspension was then used for transformation with the floral dip method (Clough and Bent, 1998). Seeds were harvested and dried for 1 day under vacuum and then stored at 4°C.

For screening, seeds were sterilized in 70% ethanol for 1 min and 50% (v/v) bleach solution containing 0.2% (v/v) triton for 7 min, followed by four subsequent washes with sterile water. Kanamycin resistant plants were obtained by incubating seeds on 4.3 $mg.mL^{-1}$ MS, 1% sucrose and 0.7% (w/v) Agar (Sigma-Aldrich, St Louis MO), 30 $ug.mL^{-1}$ kanamycin and 200 $ug.mL^{-1}$ carbenicillin.

GUS activity assay

β-glucuronidase activity was assessed by either fluorometric assays or histochemical staining (Jefferson et al., 1987). Protein content was assessed by the Bradford (1976) method, using bovine serum albumin (BSA) as a standard.

Histochemical GUS analysis: Whole inflorescences were collected from kanamycin resistant R0 plants (100% heterozygous) and immersed in the GUS reaction buffer (50 mM NaPO4, pH 7.0, 0.6 mM potassium ferrycianide, 0.6 mM potassium ferrocyanide, 2% dimethylformamide, 6 mM EDTA and 1 $mg.mL^{-1}$ of 1 X-Gluc (5-bromo-4-chloro-3-indolyl-β-d-glucuronide and 0.1% triton detergent). Samples were incubated for 24 h at 37°C, and the reaction was stopped by replacing the solution with a 50% ethanol solution to clear the tissue. Photographs were taken with a Zeiss Axiophot microscope or with a stereoscope SZ-PT Olympus, Japan.

Fluorometric GUS analysis: Flowers at late stage11/early stage 12 were collected from kanamycin resistant R0 plants (100% heterozygous), 5 or more lines per promoter construct were evaluated. Floral tissues (12-14 flowers) were homogenized in GUS extraction buffer (50 mM $NaPO_4$, pH 7.0, 10 mM EDTA, 10 mM β-mercaptoethanol, 0.1% (v/v) Triton X-100, 0.1%(v/v) lauryl sulfate sarcosine), and microcentrifuged for 20 minutes at 14000 rpm at 4°C. 10 μL of plant extract was mixed with 10 uL of substrate buffer (extraction buffer with 4 mM of 4-methylumbelliferyl β-D-glucuronide, triethylamine salt (4MUG)). The reaction was incubated for 60 min at 37°C and was terminated by the addition of 180 μl of the stop buffer (0.2 M Na_2CO_3).

Fluorescence levels were measured with a multi-detection microplate reader Synergy HT (Bio-Tek Instruments, Highland Park, Winooski, Vermont) with

excitation at 237 / 40, Emission at 528 / 20 and Sensitivity at 50. A 4-MU (4-methyl umbelliferone) standard solution was used for the elaboration of the standard curve. Protein concentration was determined by the Bradford method (1976) using the multi-detection microplate reader Synergy HT, in the absorbance mode (OD_{595nm}). BSA was used for the standard curve.

Immunohistochemistry

Floral apexes (10-12) were immersed in a fixative solution containing 4% formaldehyde, 5% acetic acid and 50% ethanol. Floral tissue was vacuum infiltrated, placed on ice for 30 min and finally vacuum infiltrated in a Pelco microwave oven for 2 min at 250 W power, 2 min off, 2 min at 250 W power, 37°C and vacuum to a 30 sec on/off cycle. The floral tissue was dehydrated in a series of ethanol concentrations: 50, 60, 70, 80, 90, 95 and 100% ethanol (3 times). The 95 and 100% ethanol contain eosin bluish CI 45400 (Polysciences, Inc. Warrington, PA) to counterstain the tissue.

Floral tissue was embedded immediately after the ethanol dehydration. Samples were placed at 37°C for 30 min. 1/3 volume of molten Steedman's wax (90% w//w PEG distearate and 10% w/w 1-hexadecanol homogenized at 55°-60°C) (Steedman, 1957; Vitha et al., 2000) was added, homogenized and incubated for 2 h at 37°C. This step was repeated three times. Finally, the ethanol/wax mix was replaced with pure wax and incubated at 37°C overnight; this step was repeated twice. The embedded tissue was then placed in silicon embedding molds. Afterwards, the samples were sectioned with a rotary microtome (MICROM Laborger≅te Gmb H Robert- Bosch-Strasse 49 D-6919 Walldorf) at 8 μm width. Silane coated slides (EMS 63411-01) were used to place the 8 μm sections. The ribbon was floated with water and air dry overnight at room temperature and placed at 4°C for future use.

Immunolocalization

Slides were dewaxed by washing them in 100% ethanol, followed by three washes in 95% ethanol. Following air drying for 15 min, slides were hydrated in 90% ethanol for 10 min in 50% ethanol in TBS, [tris buffered saline (25 mM Tris)] for 10 min, and TBS 10 min. Slides were blocked for 2h with TBST [0.1% (v/v) tween 20 and 3.75% BSA] at room temperature. Slides were then incubated 12 h in a humid chamber with 100 uL/slide of blocking buffer with the anti-GUS primary antibody (McKnight) 1:300. This was followed by three washes with TBST buffer and then a second incubation at room temperature for 2h with the secondary monoclonal antibody (anti-rabbit immunoglobulins clone RG-16 alkaline phosphatase conjugate , Sigma) 1:300, followed by three washes of TBST of 10 min each.

Slides were then covered with a developing solution containing Nitro blue tetrazolium (NBT) 0.3 $mg.mL^{-1}$ and (5-bromo 4-chloro 3 indolyl phosphate (BCIP) 0.15 $mg.mL^{-1}$. The developing reaction proceeded for 12 hours.

Bioinformatics analysis

Weeder (Pavesi et al., 2004; Tompa et al., 2005) was used to identify consensus motifs present in a set of tapetal specific genes: *At5g07230*, *At1g61110*, *At4g34990* (Myb32), *At5g07520* (Grp18), *At4g24972*, *At5g56110* Myb Family Transcription Factor and *At5g07530* (Grp17). Weeder is a general purpose motif discovery tool (Tompa et al., 2005). The parameters used were described in chapter 3. The same set of sequences was randomized using the shuffle sequence program from X-Blast TAMU server and run under the same parameters described above.

The Clover program was used to asses wether the consensus motifs identified by Weeder were statistically overrepresented in the tapetal specific sequence set. This method proceeds in two steps; first it calculates the raw score number that quantifies the degree of the motif's presence in the test sequence, and then estimates a p-value for this raw score. The p-value is the probability of obtaining the same or higher raw score value by chance in the background sequence set; i.e, a set of random promoters for the same organism (Frith et al., 2004).

WebLogo (Beta 3 version) was used to create a logo of the consensus motif identified by Weeder in the *TAPNAC* promoter gene. The 16 sequences used have 20 additional base pairs at the flanking regions of the consensus sequence; forward and reverse complement motifs were included.

In order to verify the positions of the consensus motif within the seven tapetum specific genes we used Motif locator. This program was designed for locating in a set of sequences a motif described with a consensus (Pavesi et al., 2004). The parameters chosen allowed for two substitutions and 88% similarity to the original sequence.

Results

The *TAPNAC* promoter confers tapetal-specific GUS expression

The main goal for this set of experiments was to identify *cis*-regulatory elements in the *TAPNAC* promoter that confers tapetum specific gene expression. Analysis of a 1700 bp upstream region of the *TAPNAC* gene, including a series of deletions within this region were done to locate tapetum specification *cis*-regulatory motifs in the *TAPNAC* promoter. We first evaluated the -1700 bp

region fused upstream of the GUS reporter gene (Pro$_{TAPNAC\text{-}1700}$:GUS) in a pBI101 based binary vector (Clontech). This construct was introduced into *Arabidopsis* by the floral dip method (Clough and Bent, 1998), and 14 independent transformants were recovered. Floral tissue was collected from each of the 14 independent transformants, and the tissues were assayed for histochemical GUS activity (Fig. 4.1).

The histochemical GUS assay is characterized by the production of an intense blue compound (Cl-Br) that precipitates upon formation, allowing precise tissue localization of β-glucuronidase enzyme activity (Gallagher, 1992). The tested samples were examined under a stereoscope, and clearly showed that the *TAPNAC* promoter drove GUS expression mainly in the anthers of these transgenic plants (Fig. 4.1). In addition, histochemical GUS assays were also conducted on other plant tissues of the same 14 transgenic lines (*i.e* leaf, stem and pollen grains), but no GUS activity was detected under the conditions used (Fig. 4.1, panels C-E). These results indicated the floral-anther specificity of the Pro$_{TAPNAC\text{-}1700}$:GUS.

However, precise localization within the anther organ was required to help understand the function of *At1g61110* in the flower. Therefore, GUS stained floral tissue was embedded in polywax and sectioned. The samples were examined under a light microscope, revealing the presence of the blue Cl-Br precipitant in the tapetal cells that surround the anther's locule (Fig. 4.2). We identified the tapetum as the main tissue expressing GUS activity.

Immunolocalization using an anti-GUS antibody (a gift from Dr. Tom McKnight at Texas A&M University) and a secondary monoclonal rabbit antibody AP conjugated (Sigma) on plants containing the Pro$_{TAPNAC\text{-}1700}$:GUS construct were used to determine the exact location of GUS protein within the anther. Two

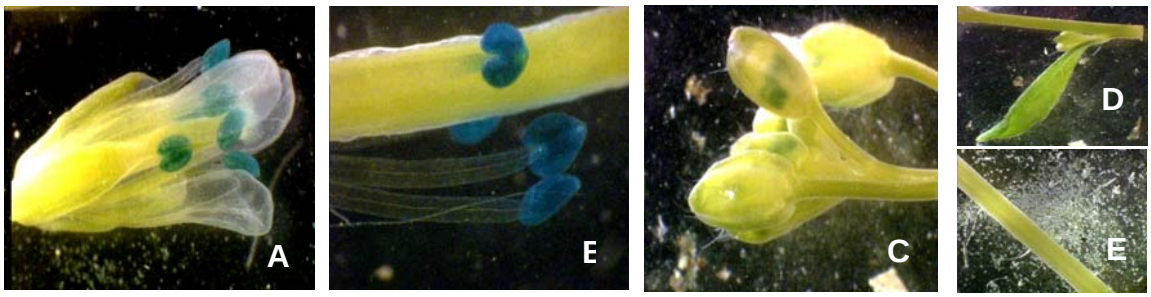


Figure 4.1. Pro$_{TAPNAC\text{-}1700}$ confers anther specific GUS expression.

A, B and **C**. Floral tissue of transgenic plants carryng Pro$_{TAPNAC\text{-}1700}$ construct. **D**. axilar bud and caulinar leaf and **E**. stem. All samples were reacted with X-Gluc

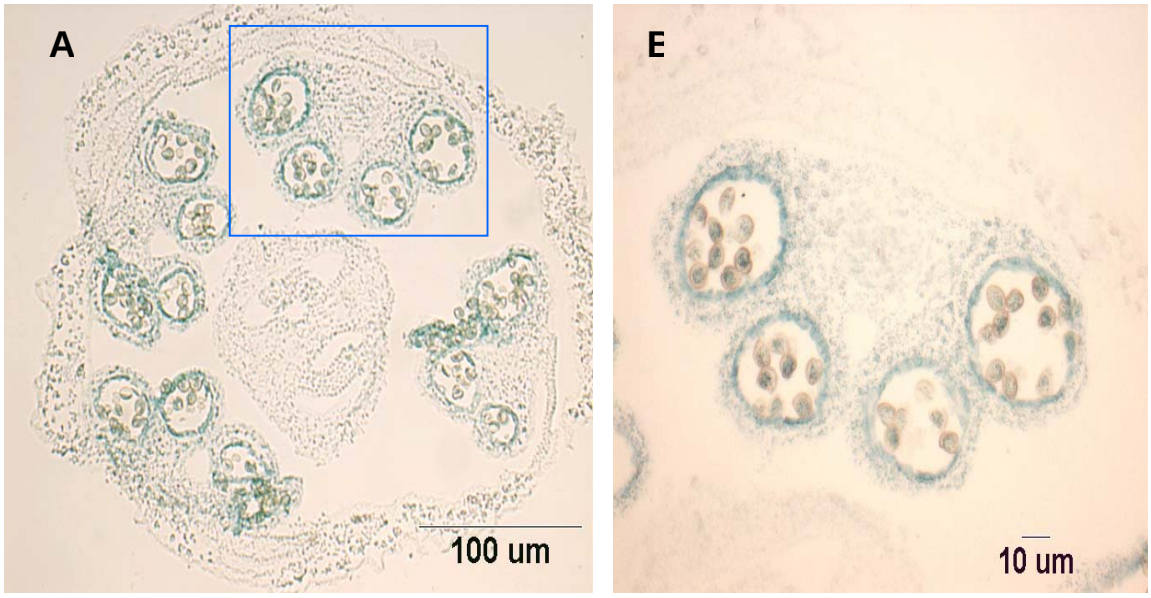


Figure 4.2. Pro$_{TAPNAC\text{-}1700}$ confers tapetum specific GUS expression.

A, B. Cross sections of immature flower buds of transgenic plants containing the Pro$_{TAPNAC\text{-}1700}$ construct stained with X-Gluc. **B.** Higher magnification of the boxed area in **A**.

different controls were used. The first was a negative control; wild type non transgenic plant material was reacted with both antibodies (not shown). The second control (Fig. 4.3) was done on the same transgenic line but omitting the primary antibody. This second control assayed the background levels of alkaline phosphatase activity in floral tissues.

Stained tissues were observed by light microscopy. The only samples that showed staining as a result of the alkaline phosphatase (AP) activity were the transgenic lines reacted with both primary and secondary antibodies. Staining was clearly located in the cell layer that surrounds the anther's locule, the tapetum. Therefore, we concluded that the NAC gene, *At1g61110*, was specifically expressed in the anther tapetum (Fig. 4.3).

Deletion analysis of the *TAPNAC* promoter reveals a 237 bp region that confers anther specific GUS expression

Initially, histochemical analysis was used to assess the effect of 5' deletions on *TAPNAC* promoter function. Plants containing GUS reporter constructs of Pro$_{TAPNAC\text{-}1700}$:GUS, Pro$_{TAPNAC\text{-}962}$:GUS, Pro$_{TAPNAC\text{-}596}$:GUS, and Pro$_{TAPNAC\text{-}237}$:GUS were assayed for GUS activity. In Fig. 4.4 panel I, Pro$_{TAPNAC\text{-}1700}$:GUS is compared with Pro$_{TAPNAC\text{-}237}$:GUS. Pro$_{TAPNAC\text{-}237}$ is sufficient to confer anther specific GUS expression (Fig. 4.4, Panel I, B-D).

A comprenhensive analysis of the *TAPNAC* promoter was required to identify important regulatory DNA sequences. Therefore, more 5' deletions were generated. The 5' deletions were done at positions -217, -189, -150, -129, -1 and +51 bp. We included 5' deletions within the 5' UTR because it is known that some genes contain important regulatory elements in this region. Some of these

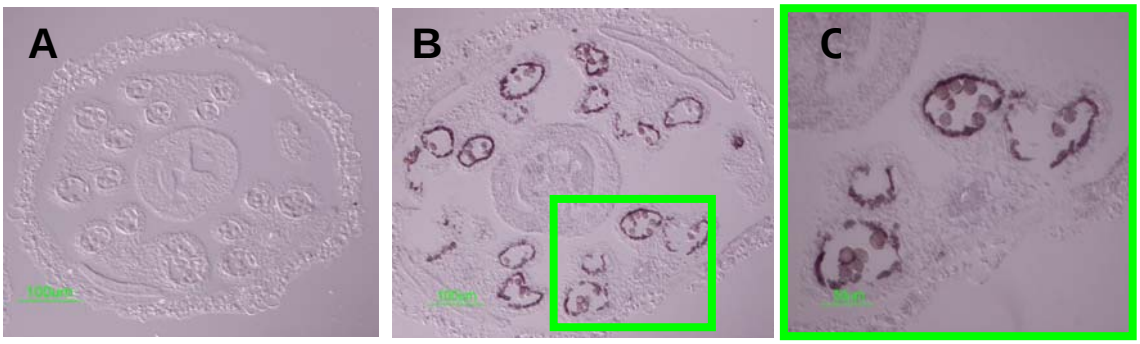


Figure 4.3. Immunolocalization of GUS protein expression driven by Pro$_{TAPNAC\text{-}1700}$.

Cross-sections of floral tissue reacted without (**A**) or with anti-GUS antibody (**B,C**) and then reacted with a secondary monoclonal antibody conjugated to alkaline phosphatase (**AP**) and stained for AP activity (**A-C**)
B=10x; C=20x

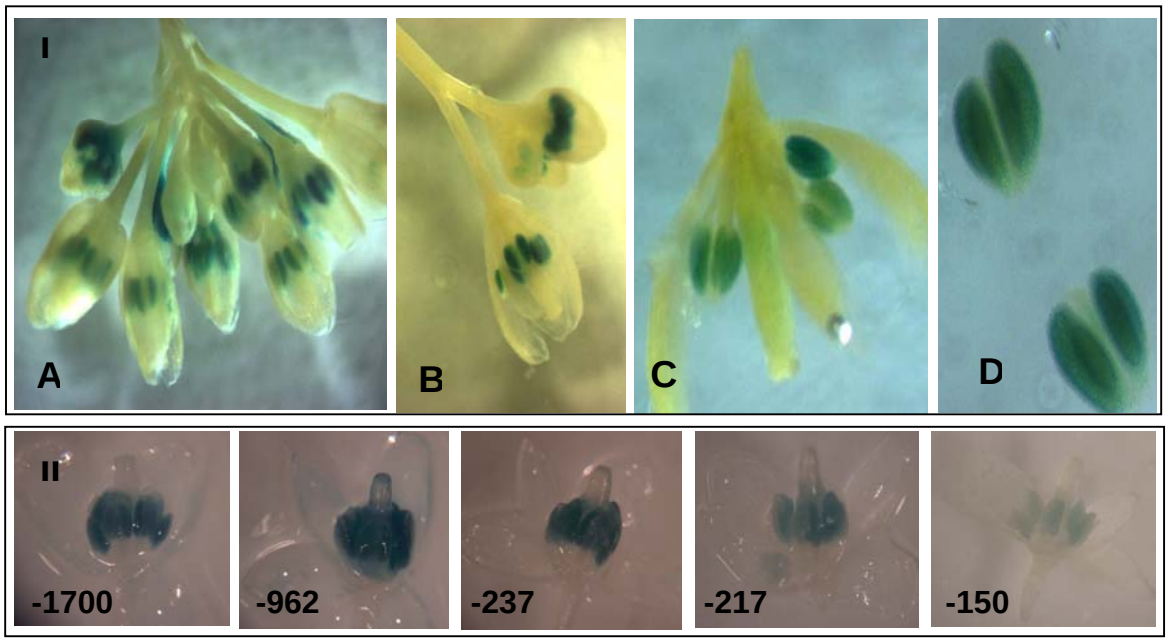


Figure 4.4. Effect of 5' deletions on *TAPNAC* promoter activity.

Histochemical analysis of floral tissue of transgenic plants containing *TAPNAC* promoter deletions are shown in panel I and II. **Panel I**. ***A***. Pro$_{TAPNAC\text{-}1700}$:GUS confers anther specific GUS expression. ***B-D.*** Pro$_{TAPNAC\text{-}237}$ is sufficient to confer anther specific GUS expression. **Panel II.** Effect of serial deletions on *TAPNAC* promoter function.

constructs, required the addition of a minimal promoter since a putative TATA box was removed from their sequence. Consequently, a different vector, pBIN19-64 CaMV, was used for Pro$_{TAPNAC\text{-}129}$, Pro$_{TAPNAC\text{-}1}$ and Pro$_{TAPNAC+51}$ constructs. This vector provided a *35S* CaMV minimal promoter containing a TATA box for binding of general transcription factors.

Changes in GUS expression were evident when the Pro$_{TAPNAC\text{-}217}$, Pro$_{TAPNAC\text{-}189}$ and Pro$_{TAPNAC\text{-}150}$ constructs were assayed for histochemical GUS activity. A decrease in the intensity of the blue Cl-Br indigo precipitant was observed in the plants carrying Pro$_{TAPNAC\text{-}189}$ and Pro$_{TAPNAC\text{-}150}$ constructs (Fig. 4.4 Panel II). Furthermore, no GUS activity was visualized in the anther tissue of plants transformed with Pro$_{TAPNAC\text{-}129}$, Pro$_{TAPNAC\text{-}1}$ and Pro$_{TAPNAC+51}$ constructs (not shown). Therefore, based on these histochemical assays we concluded that the promoter region required for tapetal GUS expression was located between -237 to -150 bp (Fig. 4.4, Panel II).

Quantitative analysis of *TAPNAC* promoter:GUS fusions

Histochemical assays revealed little difference in GUS staining between Pro$_{TAPNAC\text{-}1700}$:GUS, Pro$_{TAPNAC\text{-}962}$:GUS and Pro$_{TAPNAC\text{-}596}$:GUS. Quantitative fluorometric GUS assays were performed to evaluate *TAPNAC* promoter activity on identical constructs. All of the transgenic plants assayed were the plants derived from the first kanamycin screening. In other words, all the plants were heterozygous for the T-DNA insertion. The fluorometric assays revealed significant differences in GUS enzymatic activity (Fig. 4.5). There was a drastic reduction in GUS activity when the full promoter region, Pro$_{TAPNAC\text{-}1700}$:GUS (32.16 pmole.ug^{-1}min^{-1}) or the Pro$_{TAPNAC\text{-}962}$:GUS (69.32 pmole.µg^{-1}min^{-1}) were compared to Pro$_{TAPNAC\text{-}237}$:GUS (2.17 pmole.ug^{-1}min^{-1}).

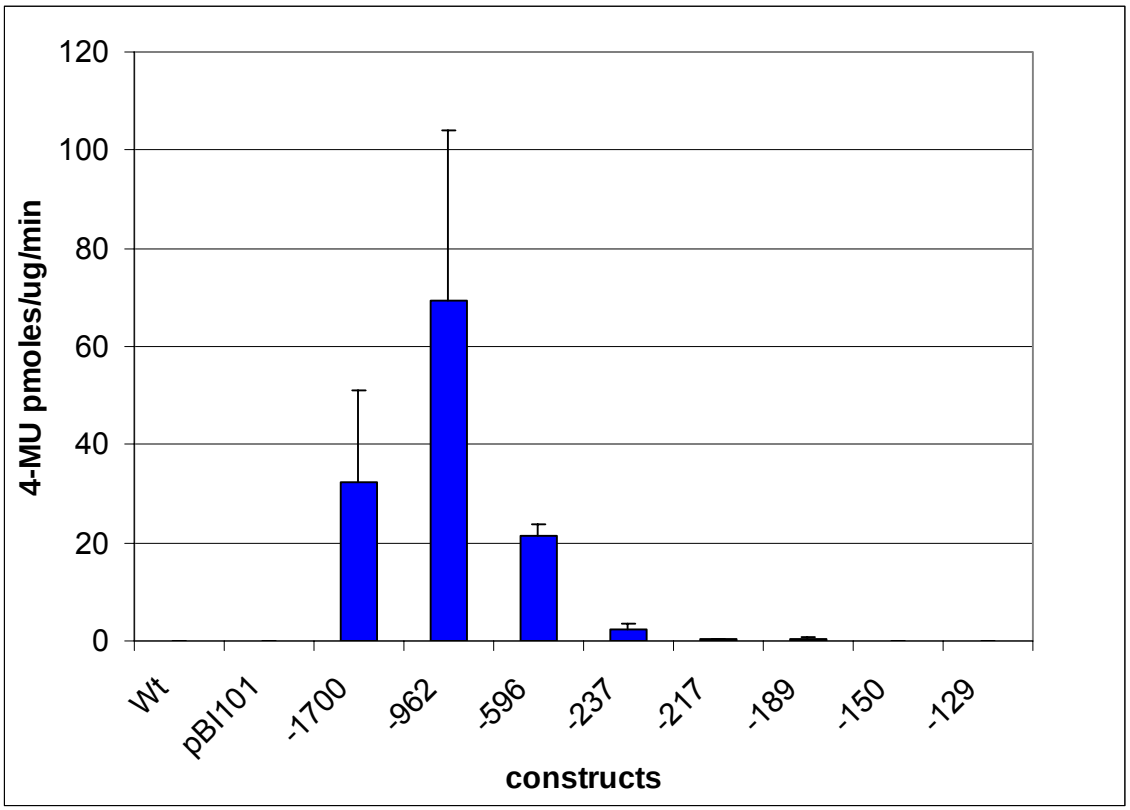


Figure 4.5. Fluorometric GUS assays on flowers containing the different Pro$_{TAPNAC}$:GUS constructs at stage 11.

Different promoter deletion constructs were assayed for GUS enzymatic activity using 4-MUG substrate. The enzymatic activities are expressed in units of 4-MU pmole.ug^{-1}.min^{-1}. At least 6 lines were evaluated per construct, the average enzymatic activity is represented in the graph. Error bars indicate standar deviation.

The next reduction in GUS activity, up to nine fold, was identified within the Pro$_{TAPNAC\text{-}237}$:GUS construct (2.17 pmole.ug^{-1}min^{-1}) and the Pro$_{TAPNAC\text{-}217}$:GUS and Pro$_{TAPNAC\text{-}189}$:GUS which both have similar enzymatic activities (0.3 pmole.ug^{-1}min^{-1}). The Pro$_{TAPNAC\text{-}150}$:GUS construct did not show any significant GUS activity, displaying enzymatic activities similar to the activities found in wild type plants with or without the pBI101 empty vector. The high standard deviations observed within the different promoter constructs (Fig. 4.5) is due to positional effect of the inserted T-DNA constructs and T-DNA copy number.

Results of Fig 4.5 suggested that the regions between -962 bp to -237 bp (Enh$_{TAPNAC\ -962}$:GUS) and -596 bp to -237 bp (Enh$_{TAPNAC\ -596}$:GUS) were potentially acting as enhancers of gene expression, since GUS activities for these two fragments were elevated. Therefore, we generated constructs by fusing these two promoter regions to either pBIN19.64 CaMV minimal promoter vector or to Pro$_{TAPNAC\text{-}150}$:GUS pBI101 vector. No significant GUS activity was observed for the constructs in the pBIN19.64 vector.

However, histochemical analysis revealed some GUS activity in the anthers of flowers at stage 12 for promoter fragments fused to the Pro$_{TAPNAC\text{-}150}$:GUS pBI101. However, MUG assays using samples from floral stages 11 and 12 yielded no significant GUS enzymatic activity (data not shown).

These results imply that the putative enhancer region that we have detected requires DNA regulatory sequences present within the -237 bp to -150 bp. Moreover, the main region should be located within Pro$_{TAPNAC\text{-}237}$ and Pro$_{TAPNAC\text{-}217}$, since it is within these two fragments where the GUS enzyme activity drops nine fold. To continue our search for the tapetal *cis* regulatory sequence, we fused this 20 bp region (within -237 to -217) to the Pro$_{TAPNAC\text{-}150}$:GUS vector and got the same results as we did for the enhancer regions. There was no

significant GUS enzymatic activity detected in plants bearing this construct. The results pointed at the importance of the context of the *TAPNAC* promoter contained in the $Pro_{TAPNAC\text{-}217}$ region, for proper transcription.

In silico promoter analysis

The lack of information on *cis*-regulatory sequences that confer tapetal gene expression compelled us to use bioinformatic tools to identify putative motifs in promoter sequences that confer tapetal specific expression (Table 4.3 and Fig. 4.6). We used a general purpose motif discovery program called Weeder (Pavesi et al., 2004) to locate regulatory sequences that are common in the promoter regions of tapetal specific genes.

We obtained the sequences of the promoter regions including the 5' UTR of *TAPNAC* and six other tapetal expressed genes (Fig 4.6.) from the TAIR website. The genes displayed in Fig 4.6 have similar patterns of expression; they are highly expressed during specific floral stages and their expression disappeared at the time the tapetal tissue is degraded. We ran Weeder with this set of sequences and identified four putative consensus motifs. The *Arabidopsis* genome initiative (AGI) numbers of the genes used for this computational analysis are shown in Table 4.2, and the motifs identified are shown in Table 4.3. These four consensus motifs were analyzed with PLACE, a database of plant *cis*-acting regulatory DNA elements (Higo et al., 1999). The first two motifs identified are novel motifs (no similar sequence present in the data base) and they are very similar, motif 1 is included within motif 2. The other consensus motifs have known associated functions: motif 3 includes an ABREATRD22 motif (ABA responsive element) and motif 4 includes an ACGT motif required

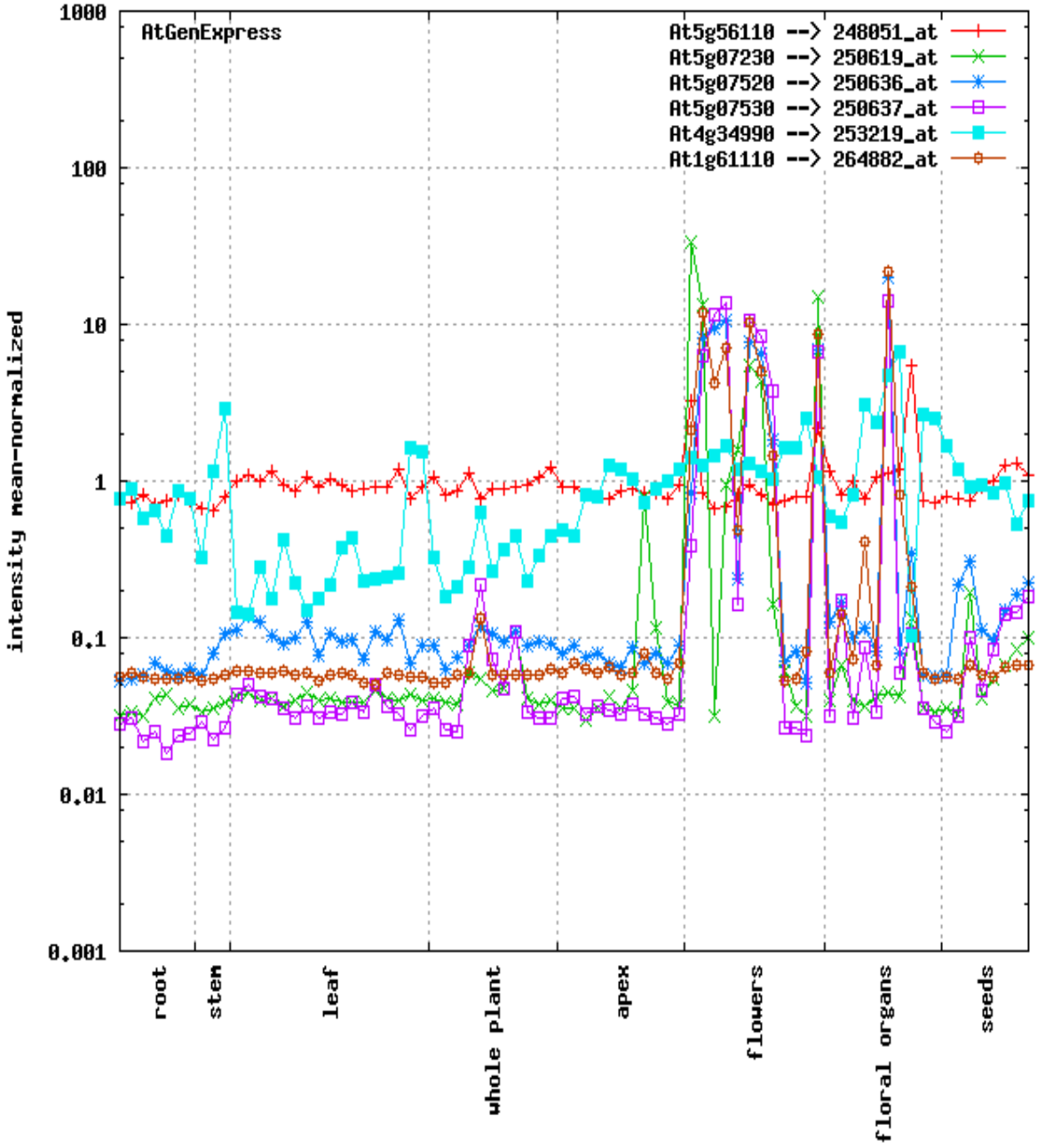


Figure 4.6. Gene expression profile of tapetal specific genes.

The *Arabidopsis* expressed visualization tool (http://jsp.weigel world.org) was used to generate the graph above.

Table 4.2. Tapetal specific genes analyzed for consensus promoter motif.

Gene AGI number	Function
At5g07230	Protease Inhibitor/Seed Storage/Lipid Transfer Protein (LTP) family protein identical to tapetum-specific protein A9
At1g61110	TAPNAC, NAC gene family protein
At4g34990	Myb Family Transcription Factor (Myb32) similar to Myb DNA-binding protein gi:19052 from [hordeum vulgare],
At5g07520	Glycine-Rich Protein (Grp18) Oleosin
At4g24972	Expressed Protein Tapetum Determinant 1 (also called At4g24973)
At5g56110	Myb Family Transcription Factor.
At5g07530	Glycine-Rich Protein 17 (Grp17) Oleosin

The *Arabidopsis* genome initiative number (AGI) and their associated functions are listed.

Table 4.3. Weeder identified putative motifs in the promoter of tapetal specific genes.

Motif	Sequence
Motif 1	TCGTGT [ACACGA]
Motif 2	TCGTGTGC [GCACACGA]
Motif 3	TCTACCACGT [ACGTGGTAGA]
Motif 4	CTTCGACGTCTA [TAGACGTCGAAG]

for etiolation-induced expression of ERD1 (EARLY RESPONSIVE TO DEHYDRATION) in Arabidopsis.

The putative novel motif is present at specific places throughout the TAPNAC promoter (Fig. 4.7). The abundance of these motifs in the different promoter fragments that were analyzed correlated well with the levels of GUS enzyme activity observed. The higher the number of motifs in the promoter fragment, the higher the GUS activity.

Fluorometric assays (Fig. 4.5) shown that the ProTAPNAC-237 (2.17 4-MU pmoles.ug-1.min-1) displays nine times higher GUS enzymatic activity than the ProTAPNAC-217, (0.33 4-MU pmoles.ug-1.min-1) indicating the presence of an important regulatory sequence within these 20 nucleotides. The only motif identified in this region that is common in other tapetal specific promoters is the consensus motif “TCGTGT” (Fig. 4.8). Therefore, we pursued a more detailed analysis on this consensus sequence.

Identification of other potential regulatory sequences adjacent to the core “TCTTGT” sequence in the *TAPNAC* promoter was performed. For this purpose multiple sequence alignment analysis was employed. Twenty bp DNA sequences flanking the consensus sequence “TCGTGT” were extracted from the *TAPNAC* promoter, and the sequences were submitted to the WebLogo program (http://weblogo.berkeley.edu/ logo.cgi).

WebLogo produces an output that is a graphical representation of a multiple sequence alignment. Each logo consists of stacks of symbols, one stack for each position in the sequence. The overall height of the stack indicates the sequence conservation at that position, while the height of symbols within the stack indicates the relative frequency of each nucleic acid at that position.

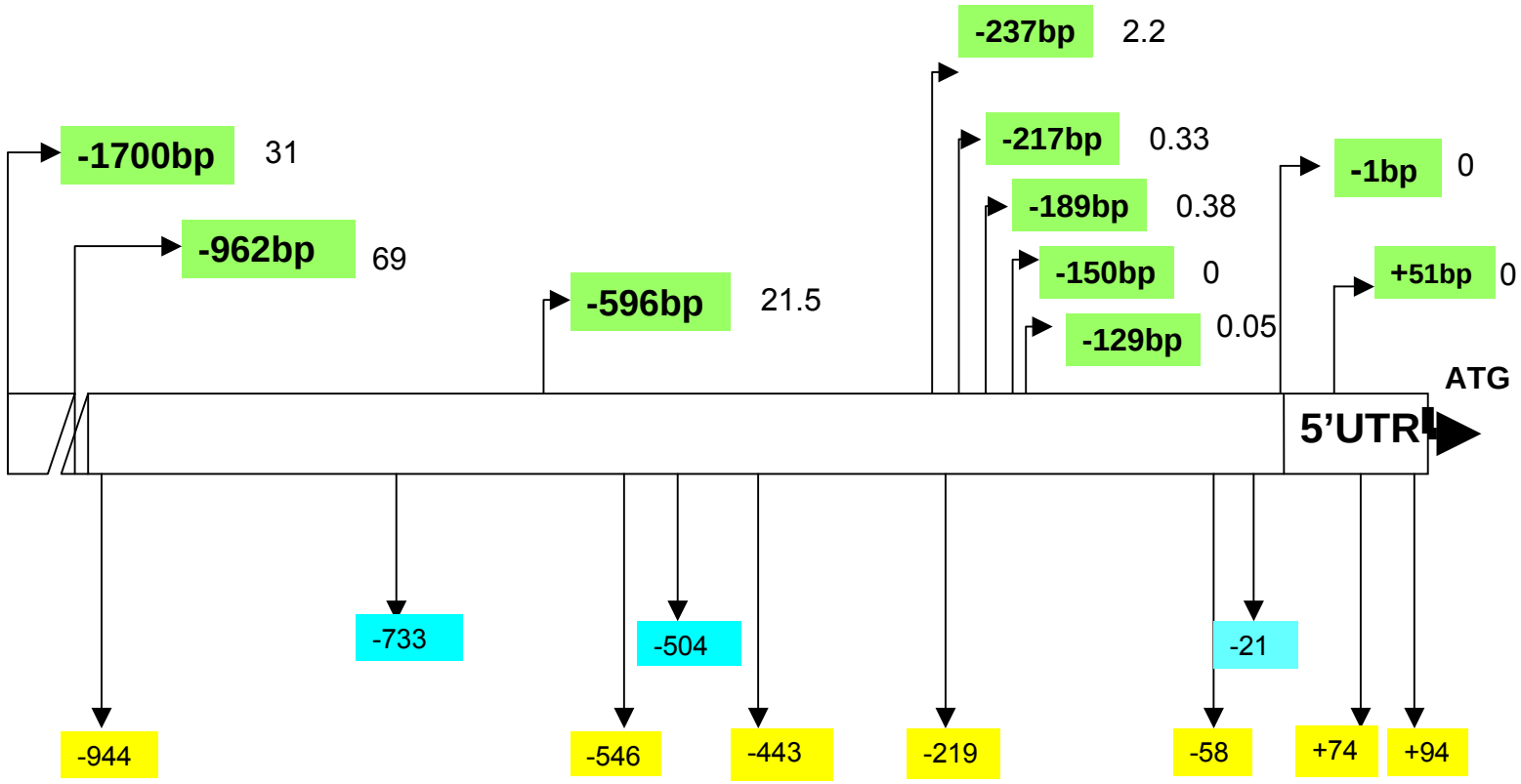


Figure 4.7. Schematic representation of *TAPNAC* promoter:GUS constructs.

Size of the promoter fragments analyzed (green) and their GUS activity in units of 4-MU pmole/μg/min are indicated. Novel consensus motifs identified by Weeder are also highlighted in yellow and light blue.

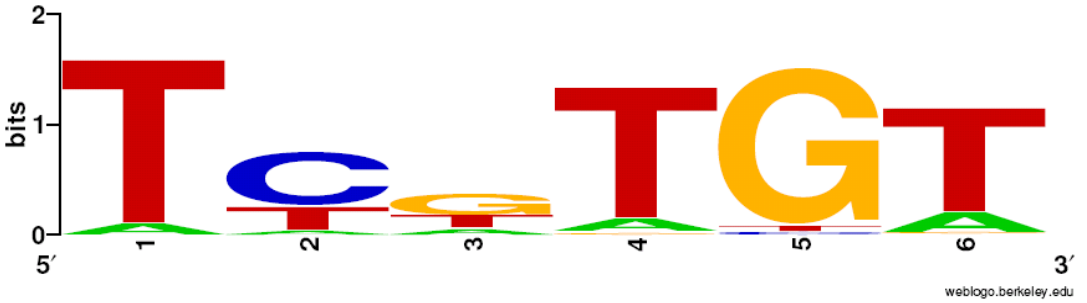


Figure 4.8. Web logo representation of motif 1.

All the sequences identified in all 7 tapetal specific genes were used to generate this logo

Based on this WebLogo analysis we found no additional regulatory sequences adjacent to the consensus motif "TCGTGT" in the *TAPNAC* promoter fragment (Fig. 4.9). However, the consensus motif TCGTGT stands out among the promoter sequences evaluated. Consequently, we wanted to determine if this motif is a real regulatory sequence present among the promoter sequences or if it occurs by chance.

We randomized the promoter sequences of the tapetal specific genes submitted to the Weeder program and resubmitted the sequences to Weeder. As expected, there was no similar motif found in the randomized set, strongly suggesting that the putative consensus sequence must play a role in the regulation of these genes.

Furthermore, Clover software (Frith et al., 2004) was used to assess if these same four consensus motifs were statistically over-represented in the tapetal specific set of sequences. We found that all four motifs displayed p-values of 0, indicating that the sequences were in fact overrepresented in the set of tapetal sequences tested.

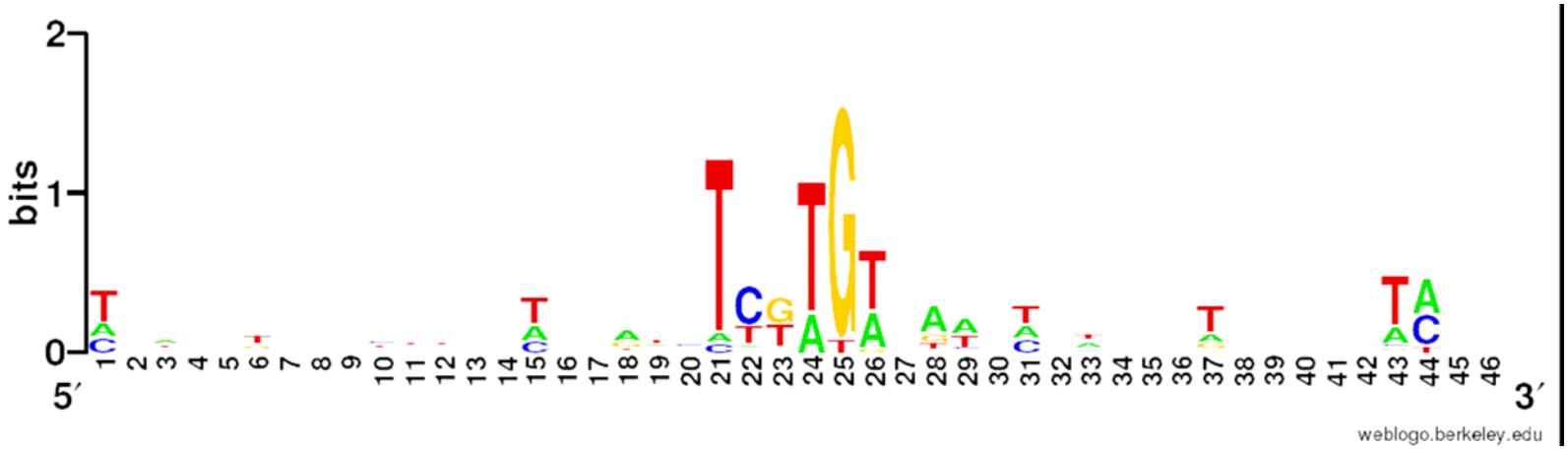


Figure 4.9. Web logo analysis of *TAPNAC* promoter.

The sequences used contain additional 20 bp flanking sequences of the putative consensus sequence for motif 1.

Detailed analysis of the -237 to -217 *TAPNAC* promoter region

We have determined that the GUS enzymatic activity is driven in the tapetum by regulatory sequences contained within the $Pro_{TAPNAC\ -237}$:GUS construct. We also discovered the importance of a 20 bp region (-237 to -217 bp), including a putative novel consensus motif **TCGTGT** (Fig 4.7). Moreover, we know that the 20 bp region requires the native sequences in the -217 to +125 *TAPNAC* promoter region to drive expression of *TAPNAC* in the tapetum. However, we were not certain that the consensus motif identified by Weeder was actually a tapetal specific motif. Therefore, we decided to analyze in more detail the 20 bp promoter region in its natural context.

We created mutations within this 20 bp DNA sequence. Four constructs were generated by PCR mutagenesis, using as a template the initial $Pro_{\ TAPNAC\ -237}$ TOPO 2.1 construct. The mutations were created so that every third of the 20 bp region was mutated in each construct, and also the whole 20 bp region was mutated maintaining the original GC content of -237 to -217 promoter DNA fragment.

GUS enzymatic activities were evaluated for these different mutated constructs $Pro^{m}_{TAPNAC-237}$:GUS once they were introduced into *Arabidopsis* plants via *Agrobacterium* T-DNA insertion. Flowers at stage 11 were collected for subsequent analysis of GUS enzymatic activity. The results of the analysis shown in Table 4.4 indicate that randomized changes of the base composition in the entire 20 bp promoter region (-237 to -217) decreased GUS activity (2.20 $pmole.ug^{-1}min^{-1}$) three fold when compared to the original non mutated 20 bp construct (6.20 $pmole.ug^{-1}min^{-1}$). The mutations in the first and second fractions of the 20 bp sequence displayed similar GUS activities when compare to the whole mutated sequence (3.81 and 2.54 $pmole.ug^{-1}min^{-1}$, respectively).

However, the third mutated containing construct reported GUS enzymatic activities of 6.35 pmole.ug^{-1}min^{-1}, same activities as the activities reported by the original construct (6.20 pmole.ug^{-1}min^{-}1) in this set of experiments. This last result was caused by a fortuitous recreation of a comparable motif once the sequence was randomized, the original sequence contained in the 20 bp promoter region was GAATTCCCACA**TTTCTAGTTT** and was changed to GAATTCCC ACA**ATCTTTTGTT**. Detailed analysis of the third shuffle sequence using Motif locator (Pavesi et al., 2004) revealed the sequence similarity of the sequence TTTTGT in the third construct with the consensus motif sequence TCGTGT shown in Figures 4.8 and 4.9.

Table 4.4. Summary of the MUG assays done for the different randomized *TAPNAC* promoter constructs.

construct	4-MU pmole/ug/min	stdv	No of lines tested
Pro*TAPNAC-217*	0.36	0.18	3
Pro$^{All\ m}$ *TAPNAC-237*	2.24	1.70	27
Pro$^{1st\ m}$ *TAPNAC-237*	3.81	2.02	15
Pro^{2ndm} *TAPNAC-237*	2.54	2.06	12
Pro$^{3rd\ m}$ *TAPNAC-237*	6.35	6.54	13
Pro*TAPNAC-237*	6.20	6.63	8

The number of transgenic lines evaluated is indicated per construct.

Discussion

Finding tapetal specific *cis*-regulatory sequences has not been an easy task as shown in earlier studies done on a tapetal-specific gene from tobacco and two *Brasica napus* tapetal oleosin like genes *OlnB* 4 and *OlnB* 13, (Koltunow et al., 1990; Hong et al., 1997). These studies did not locate tapetal specification motifs. Instead they identified promoter regions of up to 120 bp that confer tapetal gene expression. Regardless of these previous results, we attempted to identify the *cis*-regulatory sequences in the *TAPNAC* promoter that confer tapetal specific gene expression.

We used both a molecular approach and a bioinformatical approach to identify tapetal specific elements. A serial promoter deletion analysis of the *TAPNAC* promoter was conducted. We identified regions -962 to -237 and -596 to -237 as potential enhancer regions that cause GUS enzymatic activities to be 30 and 10 times respectively higher when compared to other promoter regions. We failed in expressing GUS gene when these putative enhancer regions were fused to a vector carrying a minimal promoter, pBIN19.64 CaMV or $Pro_{TAPNAC\text{-}150}$:GUS, indicating that these putative enhancer regions only work in combination with the native *TAPNAC* promoter proximal regions -237 to -150bp.

We also detected promoter fragments where GUS activity was decreased to levels comparable to background levels found in control plants with and without pBI101 empty vector. This low background activity is caused by the presence of an enzyme in the anther tissue that is able to degrade the X-Gluc substrate used in the histochemical assays. There have been reports of GUS like activity in floral tissue (Nishihara et al., 1993; Vitha et al., 1995) specially in the anthers, but the background activities can be overcome by the use of 20% methanol in the substrate solution used for the qualitative histochemical analysis.

Nevertheless, when quantitative fluorometric assays were conducted for the same plants, the GUS like enzymatic activity detected was very low and was similar for both wild type plants with and without pBI101 empty vector. Based on this result, the promoter fragments with background levels of GUS enzymatic activity were considered as fragments that did not drive expression of GUS in the anther tapetum. Several small promoter fragments confer GUS activity barely above background levels (Table 4.2), but a difference in GUS activity of almost nine fold was identified when analyzing finer promoter deletions.

The $Pro_{TAPNAC\ -237}$:GUS promoter region activities were 2.2 $pmole.ug^{-1}min^{-1}$, nine times higher than the GUS activities detected in $Pro_{TAPNAC\ -217}$:GUS and $Pro_{TAPNAC\ -189}$:GUS (0.3 $pmole.ug^{-1}min^{-1}$).

The *in silico* approach done with Weeder and WebLogo, identified four common motifs among the promoter regions of seven tapetal specific *Arabidopsis* genes including *TAPNAC* (Table 4.2). **Motif 1** and **motif 2** included a novel consensus motif **"TCGTGT"** (Fig. 4.8 and Table 4.2). This motif is present only once in the 20 bp region identified through the molecular approach.

Based on these results we decided to analyze in detail the -237 to -217 promoter region. Mutations on the 20 bp DNA promoter sequence were introduced, retaining the rest of the promoter sequence intact $Pro^{m}_{TAPNAC\ -237}$. The purpose of this fine analysis was to clarify if the TCGTGT consensus region identified with Weeder was the motif in control of tapetal specificity and if this motif was sufficient. Four different mutations were designed with X-Blast, one for the entire 20 bp region and three more containing overlapping regions. The mutations were done by randomized changes of the original base composition of the desired region while keeping the GC content of the original 20 bp promoter region. We found that mutations of the entire 20 bp region and in the first and

second overlapping regions reduced GUS activity up to three fold when compared to the original (non randomized) 20 bp region included in the $Pro_{TAPNAC\ -237}$. However, the third overlapping region retained similar GUS enzymatic activities as the ones reported for the original $Pro_{TAPNAC\ -237}$:GUS , 6.35 and 6.20 pmole.ug^{-1}min^{-1} respectively.

The third overlapping region was randomized with X-Blast. In spite of this, the randomized sequence obtained recreated a sequence GAATTCCCACA **ATC<u>TTTTGT</u>T** that was 87% similar to the "TCGTGT" consensus motif (Fig. 4.8 and Fig 4.9).

The information obtained with Weblogo (Fig 4.8) indicates that the key nucleotides in the consensus motif are **<u>T</u>**- -**<u>TGT</u>**, the same bases that are present in the third randomized motif sequence **<u>T</u>**TT**<u>TGT</u>**. We hypothesize that mutations in the first and second overlapping regions will have no effects on GUS enzymatic activities since these sequences did not change the base composition of the putative novel motif. We also hypothesize that the mutation of the entire 20 bp promoter region will result in GUS enzymatic activities similar to the activities found in $Pro_{TAPNAC\ -217}$:GUS construct. Contrary to the expectations, constructs bearing the first and second randomized mutations displayed reductions in GUS enzymatic activities. Moreover, the entire mutated 20 bp region had reduced GUS enzymatic activity (2.24 pmole.ug^{-1}min^{-1}), but it was not comparable to the low GUS activity exhibited by the $Pro_{TAPNAC\ -217}$ (0.36 pmole.ug^{-1}min^{-1}) (Table 4.4).

Additional experiments that are not shown here were conducted to clarify the results obtained with these constructs. First, we compared and analyzed the flanking sequences of the different constructs with the motif locator software. We found that extra plasmid sequences were included only in the $Pro_{TAPNAC\ -237}$:GUS

construct (the sequences were derived from TOPO 2.1 vector), the extra sequences included two sequences similar to the putative novel consensus motif. To disprove that these sequences confer increased GUS enzymatic activity, we designed new constructs where we removed one or both sequences. The results proved that the sequences identified with motif locator were in fact tapetal specification sequences; the presence of these sequences increased the GUS enzymatic activities of the $Pro_{TAPNAC\ -217}$:GUS construct. The increase in GUS enzymatic activity correlated with the number of extra sequences added.

In other words, we confirmed that the **T**CG**TGT** *in silico* identified motif is <u>part</u> of a real *cis*-regulatory region that enhances tapetal specific gene expression <u>only</u> when complemented with the rest of the sequences in the entire $Pro_{TAPNAC\ -237}$ region.

CHAPTER V

CONCLUSIONS AND FUTURE STUDIES

Introduction

NAC genes constitute a large family of plant specific transcription factors. NAC genes were first identified in petunia embryos that carried a mutation, *no apical meristem* (*nam*), that disrupted the formation of the SAM (Souer et al., 1996). The NAC DNA binding domain was named after the three founder genes: *NAM, ATAF1*, and *CUP SHAPED COTYLEDON 1 (CUC1)* (Aida et al., 1997). The NAC gene superfamily contains 109 gene members, and some NAC members display functional redundancy, *i.e*. *CUC1-2* (Aida et al., 1997) and *NST1-2* (Mitsuda et al., 2005). The roles of NAC genes in plant growth and development are quite diverse and the functions of most NAC genes are unknown.

The NAC gene superfamily has been studied in our lab initially focusing on one NAC gene member, *AtNAM* (*At1g52880*). *ATNAM* was isolated from an *Arabidopsis* developing seed library. It was molecularly characterized, using DNA binding and yeast transactivation assays. The activation domain was mapped to the protein C terminal and the DNA binding domain was mapped to the protein N terminal (Duval et al., 2002).

We decided to explore the roles of other NAC genes in plant development. Affymetrix gene chip technology was used to explore the expression of all the NAC genes in different plant tissues. NAC gene expression profiles proved to be very complex. However, one NAC gene member, *At1g61110*, was found to be

specifically and abundantly expressed in floral tissue. This NAC gene was named *TAPNAC* for its tapetal specific expression.

Tapetum development is a process that is not well understood at the molecular level. The tapetum is the nutritious tissue that provides energy to support pollen development (meiosis and mitosis), and later pollen maturation. Important compounds are secreted from the tapetum that will form part of the pollen coat. These compounds will confer pollen grains with specific properties required for pollen viability and pollen germination. Production of resistant and viable pollen results in the perpetuation of the specie.

The importance of the tapetal cells in the development of pollen grains has been demonstrated in several studies (Kaul, 1988; Mariani et al., 1990). In general, mutations that interrupt normal tapetum development will affect pollen formation and will affect seed production in self pollinating plants. However, this can be a desirable trait in the production of hybrid seeds. Hybrid seed production is a labor intensive process that usually requires the emasculation of flowers (removal of the anthers) prior to pollination. Male sterile plants will alleviate the seed production labor and reduce the cost of hybrid seeds.

TAPNAC is a NAC transcription factor abundantly expressed in the tapetum. The role of NAC genes in plant developmental programs have been demonstrated previously; *i.e*. SAM formation regulated by *CUC1-2* (Aida et al., 1997). We decided to study the role of TAPNAC in tapetum development and also investigate the *cis*-regulatory regions that confer tapetal specific gene expression. These findings will increase the basic knowledge on the regulation of tapetum development, and later will help in the design of appropriate tools for the production of hybrid seeds. We are also contemplating the idea of using the

tapetum as a storage tissue for the production of other compounds that will be released into the pollen grains and will be easy to pull together and purify later.

The conclusions drawn from the experiments described in this dissertation are outlined below.

Conclusions

The NAC gene transcriptome is very complex

Gene expression profiles of NAC genes were analyzed with Affymetrix ATH1 chips (Chapter II). Tissues of heart embryo, mature embryo, flower, leaves and roots were used in this analysis. The results indicated that in each tissue type highly expressed NAC genes could be identified. In a few cases NAC genes were expressed in a tissue specific manner, *e.i. At1g61110*, which was detected only in floral tissue.

In silico analysis of NAC genes identified clusters of NAC genes with potentially similar functions

Clustal W sequence alignment was used to analyze NAC protein sequences (Chapter II). The alignment was then submitted to phylogenetic analysis using the NJ method and bootstrap analysis. Several clusters containing NAC genes known to be functionally redundant were identified. This was the case for clusters of *CUC1-2* (Aida et al., 1997), *NST1* and *NST2*, genes involved in secondary wall formation (Mitsuda et al., 2005), and for the vascular related NAC domain genes, *VND1-7*, (Kubo et al., 2005). We identified a cluster containing ***ATNAM***, an embryo abundant NAC gene, ***At1g61110***, a floral

specific NAC gene and ***At3g15510***, a putative jasmonic acid NAC responsive gene. We speculated that the functions of these three genes could be similar.

The floral specific *At1g61110* NAC gene displays the simplest pattern of expression

We decided to further characterize *At1g61110*, a floral specific NAC gene. No significant *At1g61110* expression was detected in any other plant tissue studied. Previous studies performed by Wellmer *et.al* (2004) have indicated the anther specificity of *At1g61110;* moreover, *in situ* experiments localized its expression to the anther tapetum (Wellmer et al., 2004).

Expression of *At1g61110* is confined to the anther tapetum. The highest level of expression is detected at floral stage 11

Spatial and temporal expression analysis of *At1g61110* was conducted. Promoter:GUS fusion was used to analyze the spatial expression of *At1g61110*. We confirmed the tapetal specific expression and renamed the *At1g61110* gene, *TAPNAC* (chapter IV). Temporal analysis of *TAPNAC* gene expression was performed by qPCR analysis. The results showed that *TAPNAC* transcript is highly expressed at floral stage 11 (chapter III).

tapnac KO line did not display a morphological phenotype under the environmental conditions studied

Functional redundancy must play a role in compensating for the lack of TAPNAC regulation in the tapetum. The importance of the tapetum as nurturing tissue during pollen development has been demonstrated in several studies. Genetic ablation of tapetal cells results in male sterility (Mariani et al., 1990). For that reason, we expected to find in the *tapnac* KO mutant lines a male sterile phenotype. We did not find the expected morphological phenotype, and decided to characterize the molecular phenotype of the T-DNA insertion line.

Characterization of *tapnac* KO molecular phenotype revealed the role of TAPNAC in regulation of several transporter genes and PLDα1

The disruption of TAPNAC protein in the tapetum of *Arabidopsis* plants caused a change in the transcriptome (chapter III). The function of TAPNAC seems to be related to nutrient remobilization from the tapetum into the anther locule of *Arabidopsis* plants. The transporter genes may play a role in actually mobilizing nutrients in and out of the tapetum. Moreover, PLDα1 enzymatic activity could help in the degradation of phospholipids contained in tapetal cell membranes and in elaioplast and tapetosome lipid bodies. The compounds enclosed in the lipid bodies, such as oleosins and triacylglycerols, will form part of the pollen coat and will confer specific pollen characteristics in relation to water relations. Steryl esters confer the pollen with desiccation tolerance, and the tryacilglycerol and oleosins (amphipathic proteins) are important in water uptake prior to pollen tube germination.

Overexpression of *TAPNAC* in other tissues results in a morphological phenotype probably related to nutrient remobilization

The OE lines were 16% shorter than the wild type plants and yielded 50% less seeds than their wild type counterparts (chapter III). This could indicate a role in nutrient remobilization for TAPNAC, movilizing nutrients from the old leaves to the seeds. Nutrient remobilization is a role that has been already discovered for *TaNAM-B1*, a wheat NAC gene.

TCGTGT consensus motif enhances tapetal expression when flanked by the sequences embedded in $Pro_{TAPNAC\text{-}217}$

Serial promoter deletion analysis was conducted on the *TAPNAC* promoter region. The smallest promoter region that can drive expression of GUS in the anther tissue was $Pro_{TAPNAC\ -217}$. The consensus motif TCGTGT was found to play a significant role in tapetum specification, but it only works when placed in its natural DNA context.

Future studies

Further characterization of *tapnac* KO plants in relation to *PLDα1* and nutrient transporters should be done. Complementation analysis should be conducted to corroborate the molecular phenotype.

The effects of PLDα1 in the degradation of lipid bodies and consequently in pollen coat characteristics should be tested. Pollen germination tests in media containing different water potentials could be used to compare rates of

germination between wild type and mutant plants (KO and dominant suppressor lines).

Pollen nutrient content analysis and pollen lipid profiling will be the best way to analyze morphological differences between wild type and mutant lines. The amount of pollen needed makes this task time and space consuming. Therefore switching to analysis of seeds produced by TAPNAC overexpressing lines can be a better option.

Double and triple mutant knockouts will be produced in order to assess gene function of *TAPNAC* and *ATNAM*. Currently a T-DNA KO line for *ATNAM* in the same Columbia ecotype has been identified in the new Wisconsin D-Lox lines. Double mutants are in the process of being produced. *At3g15510* does not have a T-DNA insertion in the coding region, so it maybe best to use RNAi and antisense technology to knockout these gene and use this mutant line in the production of triple mutants.

Finally we proposed a more fine gene expression analysis of the tapetal cells. The use of laser microcapture techniques to isolate exclusively tapetal cells will allow us to study the tapetum transcriptome. Comparisons of wild type and mutant backgrounds for *TAPNAC* and *ATNAM* and *At3g15510* genes will clarify the role of NAC genes in tapetum development.

REFERENCES

Aida, M., Ishida, T., and Tasaka, M. (1999). Shoot apical meristem and cotyledon formation during *Arabidopsis* embryogenesis: interaction among the *CUP-SHAPED COTYLEDON* and *SHOOT MERISTEMLESS* genes. Development **126,** 1563-1570.

Aida, M., Ishida, T., Fukaki, H., Fujisawa, H., and Tasaka, M. (1997). Genes involved in organ separation in *Arabidopsis*: An analysis of the *cup-shaped cotyledon* mutant. Plant Cell **9,** 841-857.

Ariizumi, T., Hatakeyama, K., Hinata, K., Inatsugi, R., Nishida, I., Sato, S., Kato, T., Tabata, S., and Toriyama, K. (2004). Disruption of the novel plant protein NEF1 affects lipid accumulation in the plastids of the tapetum and exine formation of pollen, resulting in male sterility in *Arabidopsis thaliana*. Plant J. **39,** 170-181.

Ausubel, F.M., Brent, R., Kingston, R.E., Moor, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (1990). Current Protocols in Molecular Biology. (New York: John Wiley and Sons).

Bevan, M. (1984). Binary *Agrobacterium* vectors for plant transformation. Nucl. Acids Res. **25,** 641-647.

Blackmore, S., and Barnes, S.H. (1990). Pollen wall development in angiosperms. In Microspores: Evolution and Ontogeny, S. Blackmore and R.B. Knox, eds**,** (London: Academic Press) 173-192.

Bradford, M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of proteins utilizing the principle of protein dye binding. Anal. Biochem. **72,** 248-254.

Chang , S., Puryear , J., and Cairney , J. (1993). A simple efficient method for isolating RNA from pine trees. Plant Mol. Biol. Rep. **11,** 113-116.

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

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

Collinge, M., and Boller, T. (2001). Differential induction of two potato genes, Stprx2 and StNAC, in response to infection by *Phytophthora infestans* and to wounding. Plant Mol. Biol. **46,** 521-529.

Devaiah, S.P., Pan, X., Hong, Y., Roth, M., Welti, R., and Wang, X. (2007). Enhancing seed quality and viability by suppressing phospholipase D in *Arabidopsis*. Plant J. **50,** 950-957.

Devaiah, S.P., Roth, M.R., Baughman, E., Li, M., Tamura, P., Jeannotte, R., Welti, R., and Wang, X. (2006). Quantitative profiling of polar glycerolipid species from organs of wild-type *Arabidopsis* and a *PHOSPHOLIPASE Dα1* knockout mutant. Phytochemistry **67,** 1907-1924.

Dickinson, H.G., Elleman, C.J., and Doughty, J. (2000). Pollen coatings - chimaeric genetics and new functions. Sex. Plant Reprod. **12,** 302-309.

Domínguez, E., José, A.M., Miguel, A.Q., and Heredia, A. (1999). Pollen sporopollenin: degradation and structural elucidation. Sex. Plant Reprod. **12,** 171-178.

Duval, M., Hsieh, T.-F., Kim Soo, Y., and Thomas, T.L. (2002). Molecular characterization of ATNAM: a member of the *Arabidopsis* NAC domain superfamily. Plant Mol. Biol. **50,** 237-248.

Ernst, H.A., Olsen, A.N., Skriver, K., Larsen, S., and Leggio, L.L. (2004). Structure of the conserved domain of ANAC, a member of the NAC family of transcription factors. EMBO Reports **5,** 1-5.

Fei, H., and Sawhney, V.K. (2001). Ultrastructural characterization of *male sterile 33* (*ms33*) mutant in *Arabidopsis* affected in pollen desiccation and maturation. Can. J. Bot. **79,** 118-129.

Feys, B.J.F., Benedetti, C.E., Penfold, C.N., and Turner, J.G. (1994). *Arabidopsis* mutants selected for resistance to the phytotoxin coronatine are male sterile, insensitive to methyl jasmonate, and resistant to a bacterial pathogen. Plant Cell **6,** 751-759.

Fiebig, A., Kimport, R., and Preuss, D. (2004). Comparisons of pollen coat genes across Brassicaceae species reveal rapid evolution by repeat expansion and diversification. Proc. Natl. Acad. Sci. **101,** 3286-3291.

Frith, M.C., Fu, Y., Yu, L., Chen, J.-F., Hansen, U., and Weng, Z. (2004). Detection of functional DNA motifs via statistical over-representation. Nucl. Acids Res. **32,** 1372-1381.

Fujita, M., Fujita, Y., Maruyama, K., Seki, M., Hiratsu, K., Ohme-Takagi, M., Tran, L.-S.P., Yamaguchi-Shinozaki, K., and Shinozaki, K. (2004). A dehydration-induced NAC protein, RD26, is involved in a novel ABA-dependent stress-signaling pathway. Plant J. **39,** 863-876.

Gallagher, S.R. (1992). GUS protocols. Using the GUS gene as a reporter of gene expression. (San Diego CA: Academic Press, Inc.).

Goldberg, R.B., Beals, T.P., and Sanders, P.M. (1993). Anther development: basic principles and practical applications. Plant Cell **5,** 1217-1229.

Guo, Y., and Gan, S. (2006). ATNAP, a NAC family transcription factor, has an important role in leaf senescence. Plant J. **46,** 601-612.

Hegedus, D., Yu, M., Baldwin, D., Gruber, M., Sharpe, A., Parkin, I., Whitwill, S., and Lydiate, D. (2003). Molecular characterization of *Brassica napus* NAC domain transcriptional activators induced in response to biotic and abiotic stress. Plant Mol. Biol. **53,** 383-397.

Heslop-Harrison, J. (1966). Cytoplasmic connexions between angiosperm meiocytes. Ann. Bot. **30,** 221-222.

Hibara, K.-i., Takada, S., and Tasaka, M. (2003). *CUC1* gene activates the expression of SAM-related genes to induce adventitious shoot formation. Plant J. **36,** 687-696.

Higo, K., Ugawa, Y., Iwamoto, M., and Korenaga, T. (1999). Plant *cis*-acting regulatory DNA elements (PLACE) database: 1999. Nucl. Acids Res. **27,** 297-300.

Hiratsu, K., Matsui, K., Koyama, T., and Ohme-Takagi, M. (2003). Dominant repression of target genes by chimeric repressors that include the EAR motif, a repression domain, in *Arabidopsis*. Plant J. **34,** 733-739.

Hoffman, C.S., and Winston, F. (1987). A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. Gene **57,** 267-272.

Hong, H.P., Ross, J., H. E., Gerster, J., L., Rigas, S., Datla, R., S. S., Hatzopoulos, P., Scoles, G., Keller, W., Murphy, D., J., and Robert, L., S. (1997). Promoter sequences from two different *Brassica napus* tapetal oleosin-like genes direct tapetal expression of β-glucuronidase in transgenic *Brassica* plants. Plant Mol. Biol. **34,** 549-555.

Hsieh, K., and Huang, A.H.C. (2004). Endoplasmic reticulum, oleosins, and oils in seeds and tapetum cells. Plant Physiol. **136,** 3427-3434.

Hsieh, K., and Huang, A.H.C. (2005). Lipid-rich tapetosomes in *Brassica* tapetum are composed of oleosin-coated oil droplets and vesicles, both assembled in and then detached from the endoplasmic reticulum. Plant J. **43,** 889-899.

Hsieh, K., and Huang, A.H.C. (2007). Tapetosomes in *Brassica* tapetum accumulate endoplasmic reticulum-derived flavonoids and alkanes for delivery to the pollen surface. Plant Cell **19,** 582-596.

Ishiguro, S., Kawai-Oda, A., Ueda, J., Nishida, I., and Okada, K. (2001). The *DEFECTIVE IN ANTHER DEHISCENCE1* gene encodes a novel phospholipase A1 catalyzing the initial step of jasmonic acid biosynthesis, which synchronizes pollen maturation, anther dehiscence, and flower opening in *Arabidopsis*. Plant Cell **13,** 2191-2209.

Jefferson, R., Kavanagh, T., and Bevan, M. (1987). GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants EMBO J **6,** 3901-3907.

Kaul, M.L.H. (1988). Male sterility in higher plants. In Monographs on Theoretical and Applied Genetics R. Frankel, ed (Berlin: Springer-Verlag), pp. 15 - 32.

Kikuchi, K., Ueguchi-Tanaka, M., Yoshida, K.T., Nagato, Y., Matsusoka, M., and Hirano, H.Y. (2000). Molecular analysis of the NAC gene family in rice. Mol. Gen. Genet. **262,** 1047-1051.

Koltunow, A.M., Truettner, J., Cox, K.H., Wallroth, M., and Goldberg, R.B. (1990). Different temporal and spatial gene expression patterns occur during anther development. Plant Cell **2,** 1201-1224.

Krizek, B.A., and Fletcher, J.C. (2005). Molecular mechanisms of flower development: an armchair guide. Nat. Rev. Genet. **6,** 688-698.

Kubo, M., Udagawa, M., Nishikubo, N., Horiguchi, G., Yamaguchi, M., Ito, J., Mimura, T., Fukuda, H., and Demura, T. (2005). Transcription switches for protoxylem and metaxylem vessel formation. Genes Dev. **19,** 1855-1860.

Kusano, H., Asano, T., Shimada, H., and Kadowaki, K.-i. (2005). Molecular characterization of ONAC300, a novel NAC gene specifically expressed at early stages in various developing tissues of rice. Mol. Genet. Genomics **272,** 616-626.

Leitch, A.R. (2000). Higher levels of organization in the interphase nucleus of cycling and differentiated cells. Microbiol. Mol. Biol. Rev. **64,** 138-152.

Lin, J.-F., and Wu, S.-H. (2004). Molecular events in senescing *Arabidopsis* leaves. Plant J. **39,** 612-628.

Liu, X.C., and Dickinson, H.G. (1989). Cellular energy levels and their effect on male cell abortion in cytoplasmically male sterile lines of *Petunia hybrida*. Sex. Plant Reprod. **2,** 167-172.

Luscombe, N.M., Austin, S.E., Berman, H.M., and Thornton, J.M. (2000). An overview of the structures of protein-DNA complexes. Genome Biol. **1,** 1 - 37.

Ma, H. (2005). Molecular genetic analyses of microsporogenesis and microgametogenesis in flowering plants. Annu. Rev. Plant Physiol. **56,** 393-434.

Mariani, C., Beuckeleer, M.D., Truettner, J., Leemans, J., and Goldberg, R.B. (1990). Induction of male sterility in plants by a chimaeric ribonuclease gene. Nature **347,** 737-741.

Mayfield, J.A., Fiebig, A., Johnstone, S.E., and Preuss, D. (2001). Gene families from the *Arabidopsis thaliana* pollen coat proteome. Science **292,** 2482-2485.

Mitsuda, N., Seki, M., Shinozaki, K., and Ohme-Takagi, M. (2005). The NAC transcription factors NST1 and NST2 of *Arabidopsis* regulate secondary wall thickenings and are required for anther dehiscence. Plant Cell **17,** 2993-3006.

Nishihara, M., Ito, M., Tanaka, I., Kyo, M., Ono, K., Irifune, K., and Morikawa, H. (1993). Expression of the β-glucuronidase gene in pollen of lily (*Lilium longiflorum*), tobacco (*Nicotiana tabacum*), nicotiana rustica, and peony (*Paeonia lactiflora*) by particle bombardment. Plant Physiol. **102,** 357-361.

Olsen, A.N., Ernst, H.A., Leggio, L.L., and Skriver, K. (2005). NAC transcription factors: structurally distinct, functionally diverse. Trends Plant Sci. **10,** 79-87.

Olsen, A.N., Ernst, H.A., Leggio, L., Johansson, E., Larsen, S., and Skriver, K. (2004). Preliminary crystallographic analysis of the NAC domain of ANAC, a member of the plant-specific NAC transcription factor family. Acta Crystallogr., Sect D: Biol. Crystallogr. **60,** 112-115.

Ooka, H., Satoh, K., Doi, K., Nagata, T., Otomo, Y., Murakami, K., Matsubara, K., Osato, N., Kawai, J., Carninci, P., Hayashizaki, Y., Suzuki, K., Kojima, K., Takahara, Y., Yamamoto, K., and Kikuchi, S. (2003). Comprehensive analysis of NAC family genes in *Oryza sativa* and *Arabidopsis thaliana*. DNA Res. **10,** 239-247.

Pabo, C.O., and Sauer, R.T. (1992). Transcription factors: structural families and principles of DNA recognition. Annu. Rev. Biochem. **61,** 1053-1095.

Pacini, E. (1990). Tapetum and microspore function. In Microspores: Evolution and Ontogeny, S. Blackmore and R.B. Knox, eds (San Diego: Academic Press), pp. 213–237.

Papini, A., Mosti, S., and Brighigna, L. (1999). Programmed-cell-death events during tapetum development of angiosperms. Protoplasma **207,** 213-221.

Pavesi, G., Mereghetti, P., Mauri, G., and Pesole, G. (2004). Weeder Web: discovery of transcription factor binding sites in a set of sequences from co-regulated genes. Nucl. Acids Res. **32,** 199-203.

Paxson-Sowders, D.M., Dodrill, C.H., Owen, H.A., and Makaroff, C.A. (2001). DEX1, a novel plant protein, is required for exine pattern formation during pollen development in *Arabidopsis*. Plant Physiol. **127,** 1739-1749.

Piffanelli, P., Ross, J.H.E., and Murphy, D.J. (1998). Biogenesis and function of the lipidic structures of pollen grains. Sex. Plant Reprod. **11,** 65-80.

Ping-Li, L., Nai-Zhi, C., Rui, A., Zhao, S., Bi-Shu, Q., Fei, R., Jia, C., and Xue-Chen, W. (2007). A novel drought-inducible gene, ATAF1, encodes a NAC family protein that negatively regulates the expression of stress-responsive genes in *Arabidopsis*. Plant Mol. Biol. **63,** 289-305.

Preuss, D., Lemieux, B., Yen, G., and Davis, R.W. (1993). A conditional sterile mutation eliminates surface components from *Arabidopsis* pollen and disrupts cell signaling during fertilization. Genes Dev. **7,** 974-985.

Quinn, J.M., and Merchant, S. (1995). Two copper-responsive elements associated with the *Chlamydomonas* Cyc6 gene function as targets for transcriptional activators. Plant Cell **7,** 623-638.

Quinn, J.M., Barraco, P., Eriksson, M., and Merchant, S. (2000). Coordinate copper- and oxygen-responsive Cyc6 and Cpx1 expression in *Chlamydomonas* is mediated by the same element. J. Biol. Chem. **275,** 6080-6089.

Reznickova, S.A., and Dickinson, H.G. (1982). Ultrastructural aspects of storage lipid mobilization in the tapetum of *Lilium hybrida* var. enchantment. Planta **155,** 400-408.

Rhee, S.Y., and Somerville, C.R. (1998). Tetrad pollen formation in quartet mutants of *Arabidopsis thaliana* is associated with persistence of pectic polysaccharides of the pollen mother cell wall. Plant J. **15,** 79-88.

Riechmann, J.L., Heard, J., Martin, G., Reuber, L., Z, C., Jiang, Keddie, J., Adam, L., Pineda, O., Ratcliffe, O.J., Samaha, R.R., Creelman, R., Pilgrim, M., Broun, P., Zhang, J.Z., Ghandehari, D., Sherman, B.K., and L. Yu, G. (2000). *Arabidopsis* transcription factors: genome-wide comparative analysis among eukaryotes. Science **290,** 2105-2110.

Ruiz-Medrano, R., Xoconostle-Cazares, B., and Lucas, W.J. (1999). Phloem long-distance transport of CmNACP mRNA: implications for supracellular regulation in plants. Development **126,** 4405-4419.

Sanders, P.M., Lee, P.Y., Biesgen, C., Boone, J.D., Beals, T.P., Weiler, E.W., and Goldberg, R.B. (2000). The *Arabidopsis DELAYED DEHISCENCE1* gene encodes an enzyme in the jasmonic acid synthesis pathway. Plant Cell **12,** 1041-1062.

Sanders, P.M., AnhthuQ., B., Weterings, K., McIntire, K.N., Hsu, Y.-C., Lee, P.Y., Truong, M.T., Beals, T.P., and Goldberg, R.B. (1999). Anther developmental defects in *Arabidopsis thaliana* male-sterile mutants. Sex. Plant Reprod. **11,** 297-322.

Sang, Y., Zheng, S., Li, W., Huang, B., and Wang, X. (2001). Regulation of plant water loss by manipulating the expression of phospholipase Dα1. Plant J. **28,** 135-144.

Schwarz-Sommer, Z., Huijser, P., Nacken, W., Saedler, H., and Sommer, H. (1990). Genetic control of flower development by homeotic genes in *Antirrhinum majus*. Science **250,** 931-936.

Scott, R.J., Spielman, M., and Dickinson, H.G. (2004). Stamen structure and function. Plant Cell **16,** S46-60.

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

Sorensen, A.-M., Krober, S., Unte, U.S., Huijser, P., Dekker, K., and Saedler, H. (2003). The *Arabidopsis ABORTED MICROSPORES* (*AMS*) gene encodes a MYC class transcription factor. Plant J. **33,** 413-423.

Souer, E., van Houwelingen, A., Kloos, D., Mol, J., and Koes, R. (1996). The no apical meristem gene of Petunia is required for pattern formation in embryos and flowers and is expressed at meristem and primordia boundaries. Cell **85,** 159-170.

Stadler, R., Truernit, E., Gahrtz, M., and Sauer, N. (1999). The AtSUC1 sucrose carrier may represent the osmotic driving force for anther dehiscence and pollen tube growth in *Arabidopsis*. Plant J. **19,** 269-278.

Steedman, H.F. (1957). A new ribboning embedding medium for histology. Nature **179,** 1345.

Steiner-Lange, S., Unte, U.S., Eckstein, L., Yang, C., Wilson, Z.A., Schmelzer, E., Dekker, K., and Saedler, H. (2003). Disruption of *Arabidopsis thaliana* MYB26 results in male sterility due to non-dehiscent anthers. Plant J. **34,** 519-528.

Stintzi, A., and Browse, J. (2000). The *Arabidopsis* male-sterile mutant, *opr3*, lacks the 12-oxophytodienoic acid reductase required for jasmonate synthesis. Proc. Natl. Acad. Sci. **97,** 10625-10630.

Taji, T., Seki, M., Yamaguchi-Shinozaki, K., Kamada, H., Giraudat, J., and Shinozaki, K. (1999). Mapping of 25 drought-inducible genes, *RD* and *ERD*, in *Arabidopsis thaliana*. Plant Cell Physiol. **40,** 119-123.

Takada, S., Hibara, K., Ishida, T., and Tasaka, M. (2001). The *CUP-SHAPED COTYLEDON 1* gene of *Arabidopsis* regulates shoot apical meristem formation. Development **128,** 1127-1135.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007). MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. **24,** 1596-1599.

Taoka, K.-i., Yanagimoto, Y., Daimon, Y., Hibara, K.-i., Aida, M., and Tasaka, M. (2004). The NAC domain mediates functional specificity of CUP-SHAPED COTYLEDON proteins. Plant J. **40,** 462-473.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). ClustalW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl. Acids Res. **22,** 4673- 4680.

Tohru, A., Katsunori, H., Kokichi, H., Shusei, S., Tomohiko, K., Satoshi, T., and Kinya, T. (2003). A novel male-sterile mutant of *Arabidopsis thaliana*, *faceless pollen-1*, produces pollen with a smooth surface and an acetolysis-sensitive exine. Plant Mol. Biol. **53,** 107-116.

Tompa, M., Li, N., Bailey, T.L., Church, G.M., Moor, B.D., Eskin, E., Favorov, A.V., Frith, M.C., Fu, Y., Kent, W.J., Makeev, V.J., Mironov, A.A., Noble, W.S., Pavesi, G., Pesole, G., Régnier, M., Simonis, N., Sinha, S., Thijs, G., van

Helden, J., Vandenbogaert, M., Weng, Z., Workman, C., Ye, C., and Zhu, Z. (2005). Assessing computational tools for the discovery of transcription factor binding sites. Nature Biotechnology **23,** 137-144.

Uauy, C., Distelfeld, A., Fahima, T., Blechl, A., and Dubcovsky, J. (2006). A NAC gene regulating senescence improves grain protein, zinc, and iron content in wheat. Science **314,** 1298-1301.

Vitha, S., Balu_ka, F., Jasik, J., Volkmann, D., and Barlow, P. (2000). Steedman's wax for F-actin visualization. In Actin: A Dynamic Framework for Multiple Plant Cell Functions, Volkmann and P. Barlow, eds (The Netherlands: Kluwer: Dordrecht), pp. 619-636.

Vitha, S., K, Benes, J.P., Phillips, and Gartland, K.M.A. (1995). Histochemical GUS Analysis. In: *Agrobacterium* protocols. (Totowa, NJ: Humana Press).

Vroemen, C.W., Mordhorst, A.P., Albrecht, C., Kwaaitaal, M.A.C.J., and de Vries, S.C. (2003). The *CUP-SHAPED COTYLEDON 3* gene is required for boundary and shoot meristem formation in *Arabidopsis*. Plant Cell **15,** 1563-1577.

Weiss, H., and Maluszynska, J. (2001). Molecular cytogenetic analysis of polyploidization in the anther tapetum of diploid and autotetraploid *Arabidopsis thaliana* plants. Ann. Bot. **87,** 729-735.

Wellmer, F., Riechmann, J.L., Alves-Ferreira, M., and Meyerowitz, E.M. (2004). Genome-wide analysis of spatial gene expression in *Arabidopsis* flowers. Plant Cell **16,** 1314-1326.

Wellmer, F., Alves-Ferreira, M., Dubois, A., Riechmann, J.L., and Meyerowitz, E.M. (2006). Genome-wide analysis of gene expression during early *Arabidopsis* flower development. PLoS Genetics **2,** e117.

Welti, R., Li, W., Li, M., Sang, Y., Biesiada, H., Zhou, H.-E., Rajashekar, C.B., Williams, T.D., and Wang, X. (2002). Profiling membrane lipids in plant stress responses. Role of phospholipase Dα in freezing-induced lipid changes in *Arabidopsis*. J. Biol. Chem. **277,** 31994-32002.

Wu, H.-m., and Cheung, A.Y. (2000). Programmed cell death in plant reproduction. Plant Mol. Biol. **44,** 267-281.

Xie, D.-X., Feys, B.F., James, S., Nieto-Rostro, M., and Turner, J.G. (1998). *COI1*: An *Arabidopsis* gene required for jasmonate-regulated defense and fertility. Science **280,** 1091-1094.

Yang, C., Xu, Z., Song, J., Conner, K., Barrena, G.V., and Wilson, Z.A. (2007). *Arabidopsis* MYB26/MALE STERILE35 regulates secondary thickening in the endothecium and is essential for anther dehiscence. Plant Cell **19,** 534-548.

Zhang, W., Qin, C., Zhao, J., and Wang, X. (2004). Phospholipase Dα1-derived phosphatidic acid interacts with ABI1 phosphatase 2C and regulates abscisic acid signaling. Proc. Natl. Acad. Sci. **101,** 9508-9513.

Zhao, D., and Ma, H. (2000). Male fertility: a case of enzyme identity. Curr. Biol. **10,** R904-R907.

Zheng, Z., Xia, Q., Dauk, M., Shen, W., Selvaraj, G., and Zou, J. (2003). Arabidopsis AtGPAT1, a member of the membrane-bound glycerol-3-phosphate acyltransferase gene family, is essential for tapetum differentiation and male fertility. Plant Cell **15,** 1872-1887.

APPENDIX

Figure A.1. Phylogenetic analysis of NAC proteins

Bootstrap numbers supporting this neighbor-joining tree are indicated for every cluster.

Figure A.1. Continued

Table A.1. Primers used in qPCR to verify NAC gene profiles

Gene	Forward Primer	Reverse Primer
At1g61110	CGGTATGGTCGTTTCTTCAGACA	TCGTGGCCCATTCGGTAT
At1g79580	GATCATGCACGAATATCGCTTG	AACCCACCCATCTTCCTGAATT
At3g15500	CGTTCCGAGTATTCGTTACGGT	CGAGGAATCCCCTCAGTTTG
At1g52890	TCCACGGTCTTGCGGATAC	TGCTCAACATTACCGGCAAA
At4g27410	CCGACCAATGGATTACCGAGT	CTTCCGCCGCTCGAAA
At5g07680	GCAACCCTTCCCTTAGCTCC	GGTTGGTAGAGTGGAATCATCTGG

Table A.2. Genes expressed only in flower
From affymetrix experiment comparing different plant organ tissues (t- test p- value cut off=0.1)

AGI number	Gene name	Function aasigned	No of genes
At3g52810		acid phosphatase activity	
At4g36350		acid phosphatase activity	
		acid phosphatase activity	2
At3g28790		ATP binding	
At1g14460		ATP binding	
		ATP binding	2
At4g25950		vacuolar ATP synthase, putative	
At1g80660	AHA9	hydrogen-exporting ATPase activity	
At3g13220		ATPase activity,	
		ATPase activity	3
At4g27580		vacuolar calcium related	
At5g17480		calcium ion binding	
At3g22910		calcium-transporting ATPase activity/ calmodulin binding	
At2g26410		calmodulin binding	
At5g35670		calmodulin binding	
		Calcium related	5
At1g12570		aldehyde-lyase activity	
At3g04290		carboxylic ester hydrolase activity	
At5g33370		carboxylic ester hydrolase activity	
At5g42170		carboxylic ester hydrolase activity	
At1g75940		hydrolase activity, hydrolyzing O-glycosyl compounds	
At2g44560		hydrolase activity, hydrolyzing O-glycosyl compounds	
At3g43860		hydrolase activity, hydrolyzing O-glycosyl compounds	
At3g52600		hydrolase activity, hydrolyzing O-glycosyl compounds	
At4g35010		beta-galactosidase activity	
At3g01700		arabinogalactan-protein (agp11)	
At3g20865		arabinogalactan-protein, putative (agp)	
At2g24450		fasciclin-like arabinogalactan family protein	
		Carbohydrate metabolism	12
At1g14420		pectate lyase activity	
At1g30350		pectate lyase activity	
At4g22080		pectate lyase activity	
At5g55720		pectate lyase activity	
At2g26450		pectinesterase activity	
At3g17060		pectinesterase activity	
At3g62170		pectinesterase activity	
At5g07420		pectinesterase activity	
At5g27870		pectinesterase activity	
At2g47050		pectinesterase inhibitor activity	
At3g62180		pectinesterase inhibitor activity	

Table A.2. Continued

AGI number	Gene name	Function aasigned	No of genes
At5g46940		pectinesterase inhibitor activity	
At1g02790	PGA4	polygalacturonase activity	
At3g07850		polygalacturonase activity	
At3g06260		polygalacturonate activity	
At3g25050	XTH3	xyloglucosyl transferase activity	
At3g62710		xylan 1,4-beta-xylosidase activity	
At5g47000		xylan 1,4-beta-xylosidase activity	
		Cell wall activity	18
At3g28780		contains a signal peptide	
At5g39880		contains a signal peptide	
At5g48210		contains a signal peptide	
At5g50830		contains a signal peptide	
At5g22430		contains a signal peptide	
		Contains a signal peptide	5
At1g01280		oxygen binding	
At1g28430		oxygen binding	
At1g57750		oxygen binding	
At1g69500		oxygen binding	
At1g74540		oxygen binding	
At5g44620		oxygen binding	
		Cytochrome P450	6
At2g19770		actin binding	
At4g25590		actin binding	
At5g52360		actin binding	
At3g28630		actin crosslinking proteins	
At5g42490		microtubule motor activity	
		Cytoeskeleton	5
At4g18395		MYB DNA binding domain	
At2g26490		nucleotide binding	
		DNA binding	2
At1g01980		electron carrier activity	
At5g56490		electron transport	
		Electron trasport	2
At1g49290		expressed protein,	
At1g02813		expressed protein	
At1g45190		expressed protein	
At3g28980		expressed protein	
At5g46770		expressed protein	
At5g54095		expressed protein	
At5g61605		expressed protein	
At1g23650		expressed protein	
At5g61720		expressed protein	
		Expressed protein	9
At4g37900		glycine-rich protein	
At2g19000		glycine rich protein	
At1g03170		hydroxiproline rich glycoprotein	
At1g30795		hydroxyproline-rich glycoprotein	
		Glycine/hydroxiproline rich protein	4

Table A.2. Continued

AGI number	Gene name	Function aasigned	No of genes
At2g36020		abscisic acid-responsive hva22	
At5g13380		response to auxin stimulus	
At1g54070		dormancy/auxin associated protein-related	
At5g38760		similar to ABA-inducible protein	
At5g53820		similar to aba-inducible protein (fagus sylvatica)	
At1g28270	RALFL4	signal transducer activity	
At3g25165	RALFL25	signal transducer activity	
		Hormone responsive	7
At1g55560		copper ion binding oxidoreductase activity	
At1g55570		copper ion binding oxidoreductase activity	
At2g45800		zinc ion binding	
At5g04180		zinc ion binding	
At1g61680		lyase activity magnesium ion binding	
At3g05930		manganese ion binding	
		Ionic homeostasis	6
At1g01460		1-phosphatidylinositol-4-phosphate 5-kinase activity	
At1g01130		cipk9 (cbl-interacting protein kinase 9)	
At1g76370		kinase activity	
At2g41860		kinase activity	
At2g43230		kinase activity	
At3g59420		kinase activity	
At3g59830		kinase activity	
At4g10260		kinase activity	
		Kinase activity	8
At1g52680		Late embryiogenesis abundant protein-related	
At4g13560		LEA containing domain	
At5g27980		seed maturation family protein, similar to lea d-34	
At2g03740		lea domain-containing protein	
		LEA protein related/ storage	4
At1g75910	EXL4	lipase activity	
At1g75920		lipase activity	
At1g75930	EXL6	lipase activity	
At5g45960		lipase activity	
At1g66850		lipid binding	
At3g51590		lipid binding	
At3g52130		lipid binding	
At4g08670		lipid binding	
At4g14815		lipid binding	
At5g07230		lipid binding	

Table A.2. Continued

AGI number	Gene name	Function aasigned	No of genes
At5g07540	GRP16	lipid binding	
At5g07560	GRP20	lipid binding	
		Lipid binding	8
At4g34850		acyltransferase activity	
At1g63710		fatty acid (omega-1)-hydroxylase	
	CYP86A7	activity	
		Lipid metabolism	2
At4g28395	ATA7	lipid transporter activity	
		Lipid transporter	1
At4g28680		tyrosine decarboxylase activity	
At2g29790		contains a signal peptide and transmembrane regions	
At3g01250		stress related ozone induced protein	
At3g52620		Phosphate actyltransferase	
At3g28830		endomembrane system	
At5g19580		endomembrane system	
At5g47470		express mainly in mature pollen	
At3g62230		F-box family protein	
At1g44222		collagen alpha chain domain (panther classification system)	
At1g23580		similar to OBP32pep protein	
AT2G28180	ATCHX8	sodium:hydrogen antiporter activity	
At3g12000		sugar binding	
At1g04540		C2 domain-containing protein, low similarity to cold-regulated gene src2 (glycine max)	
		Other functions	13
At3g11980	MS2	oxidoreductase activity, acting on the CH-CH group of donors, NAD or NADP as acceptor	
At5g16960		oxidoreductase activity/ zinc ion binding	
At5g60020		oxidoreductase activity/copper ion binding	
		Oxidoreductase activity	3
At1g06260		cysteine-type peptidase activity	
At1g72290		endopeptidase inhibitor activity	
		Peptidase related	2
At3g18220		phosphatidate phosphatase activity	
At5g66020		phosphoinositide 5-phosphatase activity	
		Phosphatase activity	2
At3g26110		BCP1 anther specific protein	
		Pollen and tapetum specific protein	1
At3g09530		AtEXO70H3 exocyst subunit- pollen grain specific	
At1g49490		extensin cell-wall protein pollen specific	

Table A.2. Continued

AGI number	Gene name	Function aasigned	No of genes
At1g29140		Pollen ole e1 allergen and extensin family protein	
At2g28355		Lcr5 | expressed protein, contains similarity to anther-specific protein	
At4g11760		Encodes a member of a family of small,secreted, cysteine rich protein	
At5g38330		Encodes a member of a family of small,secreted, cysteine rich protein	
		Pollen specific proteins	6
At1g62940		4-coumarate-CoA ligase activity	
At2g23800	GGPS2	farnesyltranstransferase activity	
		Secondary metabolism	2
At4g36490		sec14 cytosolic factor, putative / phosphoglyceride transfer protein	
At1g11250		t-SNARE activity/ vesicle trafficking	
At3g03800		t-SNARE activity	
At1g03050		cell secretion epsin sorting of multiubiquitinated cargo	
		Secretion	4
At5g26060		S1 self-incompatibility protein-related	
At5g43510		Encodes a defensin-like (defl) family protein.	
At3g26860		Self-incompatibility protein-related	
		Self incompatibility	3
At2g05850		serine carboxypeptidase activity/ pollen	
At3g52000		serine carboxypeptidase activity/pollen	
		Serine carboxypeptidase activity	2
At1g15360		transcription factor activity	
At1g25330		transcription factor activity	
At1g61110		transcription factor activity	
At1g69120	AP1	transcription factor activity	
At1g69180	CRC	transcription factor activity	
At2g16210		transcription factor activity	
At3g06160		transcription factor activity	
At5g25390		transcription factor activity	
At5g40350		transcription factor activity	
		Transcription factor activity	9
At5g23970		transferase activity	
At5g60500		transferase activity	
At5g54010		transferase activity, transferring glycosyl groups	
At1g22015		transferase activity, transferring hexosyl groups	
		Transferase activity	4

Table A.3. Down-regulated genes in the *tapnac* KO background

AGI No.	Gene name	Affy - ID	Norm T-DNA	Norm Wt	Function assigned	No. of genes
At2g38750	ANNAT4	266418_at	0.32	1.05	calcium-dependent phospholipid binding	
At2g38760	ANNAT3	266419_at	0.53	0.95	calcium-dependent phospholipid binding	
					Calcium-dependent phospholipid binding	2
At1g67070	DIN9	255881_at	0.33	0.96	mannose-6-phosphate isomerase activity	
					Carbohydrate metabolism	1
At1g06570	PDS1	262635_at	0.56	0.98	4-hydroxyphenylpyruvate dioxygenase activity/ carotenoid biosynthetic process	
					Carotenoid biosynthetic process	1
At3g62860		251235_at	0.38	0.95	catalytic activity/ aromatic compound metabolic process	
At2g23970		266562_at	0.45	1.07	catalytic activity/ defense response	
					Catalytic activity	2
At5g55590	QRT1	248066_at	0.58	1.00	pectinesterase activity.	
					Cell wall metabolism	1
At1g53490		260973_at	0.42	1.09	DNA binding	
					DNA binding	1
At5g44400		249046_at	0.45	1.01	electron carrier activity	
					Electron carrier activity	1
At1g68290		260438_at	0.54	0.97	endonuclease activity/ DNA catabolic process	
					Endonuclease activity	1
At1g20180		261243_at	0.39	1.12	expressed protein	
At1g30515		261807_at	0.47	1.12	expressed protein	
At1g54860		264187_at	0.45	1.02	expressed protein	
At1g67920		260005_at	0.42	1.00	expressed protein	
At3g03680		259222_at	0.55	1.02	expressed protein	
At3g49300		252253_at	0.43	1.00	expressed protein	
At5g54150		248196_at	0.38	0.92	expressed protein	
At5g03670		250907_at	0.57	0.99	expressed protein	
At5g49900		248581_at	0.50	0.99	expressed protein	
					Expressed protein	9
At2g27150	AAO3	263570_at	0.5	0.90	abscisic aldehyde oxidase activity	
At5g45340	CYP707A3	248964_at	0.44	0.96	oxygen binding/ abscisic acid catabolic process	
At5g24860	FPF1	246967_at	0.54	1.01	positive regulation of flower development/ response to GA stimulus	
					Hormone responsive	3

Table A.3. Continued

AGI No.	Gene name	Affy - ID	Norm T-DNA	Norm Wt	Function assigned	No. of genes
At5g24090		249767_at	0.45	0.95	hydrolyzing O-glycosyl compounds	
At5g47330		248812_at	0.19	1.01	palmitoyl-(protein) hydrolase activity	
					Hydrolase activity	2
At2g43230		266453_at	0.51	1.00	kinase activity	
At5g65710		247154_at	0.42	0.99	kinase activity	
At1g76040		262671_at	0.20	0.94	kinase activity calcium- and calmodulin-dependent	
					Kinase activity	3
At4g10850		254956_at	0.55	1.01	membrane	
At1g52870		260155_at	0.49	0.98	chloroplast/ integral to membrane	
					Membrane protein	2
At1g19230	RbohE	256011_at	0.08	1.17	iron ion binding	
					Metal binding protein	1
At1g14530		261482_at	0.57	0.95	virion binding	
At5g37670		249575_at	0.57	0.97	response to heat	
At2g03740		264039_at	0.58	1.00	embryonic development ending in seed dormancy	
At1g69610		260420_at	0.46	0.95	N-terminal protein myristoylation	
At3g28960		257543_at	0.53	0.98	aa permease activity	
At2g43080		266449_at	0.53	0.96	oxidoreductase activity	
At1g76800		259871_at	0.13	1.03	molecular_function	
					Other functions	7
At3g22370	AOX1A	258452_at	0.50	0.98	alternative oxidase activity/ response to cold	
At2g19500	CKX2	265945_at	0.22	1.03	amine oxidase activity / cytokinin catabolic process	
					Oxidase activity	2
At3g15730	PLDα1	258226_at	0.06	1.02	phospholipase D activity	
					Phospholipase activity	1
At2g38400	AGT3	267035_at	0.57	1.02	alanine-glyoxylate transaminase activity	
					Transaminase activity	1
At1g61110	NAM	264882_at	0.00	0.99	transcription factor activity	
At3g21270		258044_at	0.41	1.05	transcription factor activity	
At5g61430	NAM	247519_at	0.49	0.96	transcription factor activity	
At5g61620		247535_at	0.32	0.98	transcription factor activity	
At1g02340	HFR1	259417_at	0.53	1.01	transcription factor activity / blue light signaling pathway	
At1g15580	IAA5	261766_at	0.14	0.97	transcription factor activity/auxin stimulus	
					Transcription factor activity	6

Table A.3. Continued

AGI No.	Gene name	Affy - ID	Norm T-DNA	Norm Wt	Function assigned	No. of genes
At5g37180		249633_at	0.20	0.99	transferase activity UDP-glycosyl transferase activity	
At1g02950	ATGSTF4	262120_at	0.36	0.97	glutathione transferase activity/ toxin catabolic process	
					Transferase activity	2
At5g63850	AAP4	247304_at	0.36	0.94	transporter activity acidic amino acid	
At5g53550		248276_at	0.45	1.01	transporter activity oligopeptide transporter activity/ response to nematode	
At1g59740		262912_at	0.25	1.04	transporter activity/ oligopeptide transport	
					Transporter activity	3

Table A.4. Up-regulated genes in the *tapnac* KO background

AGI No.	Gene name	Affy - ID	Norm T-DNA	Norm Wt	Function assigned	No. of genes
At5g24420		249732_at	2.11	1.06	6-phosphoglucono-lactonase activity	
At4g15210	ATBETA-AMY	245275_at	2.03	1.07	beta-amylase activity/ starch catabolic process	
					Carbohydrate metabolism	2
At5g62340		247477_at	1.78	1.09	pectinesterase inhibitor	
At1g58370	RXF12	256025_at	2.42	1.01	endo-1,4-beta-xylanase activity/ cell wall	
					Cell wall biogenesis	2
At3g25760	AOC1	257641_s_at	1.94	0.97	chloroplast thylakoid membrane allene-oxide cyclase activity	
At1g13990		262607_at	2.99	1.02	chloroplast	
At1g52410		259609_at	2.38	1.01	chloroplast thylakoid membrane calcium ion binding	
					Chloroplast specific	3
At3g13662		257570_at	2.08	1.15	molecular_function/ lignan biosynthetic process/ defense response	
At1g72260	THI2.1	259802_at	2.28	1.01	toxin receptor binding/ defense response/ JA mediated signaling pathway	
					Defense response	2
At1g30135		256159_at	1.83	1.09	expressed protein	
At2g39330		266989_at	2.22	0.97	expressed protein	
At3g11930		258727_at	1.92	1.01	expressed protein/response to stress	
At3g30720		256940_at	12.81	1.12	expressed protein	
At3g61930		251293_at	6.01	1.04	expressed protein/ N-terminal protein myristoylation	
At4g02360		255527_at	1.78	0.98	expressed protein	
At5g16920		246416_at	6.71	1.05	expressed protein/cell adhesion	
At5g24660		249752_at	2.04	0.99	expressed protein	
					Expressed protein	8
At1g15460		262585_at	2.99	1.06	anion exchanger activity	
At1g18140		256128_at	1.71	1.11	oxidoreductase activity /copper ion binding/	
At1g22990		264729_at	1.83	1.02	metal ion binding	
At4g39950		252827_at	1.77	1.02	oxygen binding /camalexin biosynthetic process	
At5g02330		251043_s_at	3.74	0.89	zinc ion binding	
					Ion homeostasis	5

Table A.4. Continued

AGI No.	Gene name	Affy - ID	Norm T-DNA	Norm Wt	Function assigned	No. of genes
At2g32660		267548_at	3.39	1.04	kinase activity/ defense response	
					kinase activity	1
At4g34210	ASK11	253271_s_at	1.88	1.05	ubiquitin-protein ligase activity	
					Ligase activity	1
At4g21990	APR3 gene	254343_at	1.77	1.13	Encodes a protein disulfide isomerase-like (PDIL)	
At1g68500		260264_at	2.74	1.00	cellular_component	
At3g05180		259308_at	3.95	1.09	carboxylic ester hydrolase activity/lipid metabolic process	
At3g24503		258140_at	2.27	1.08	aldehyde dehydrogenase (NAD) activity/ phenylpropanoid biosynthetic process	
At4g29700		253697_at	1.76	1.08	type I phosphodiesterase/nucleotide pyrophosphatase family protein	
At5g23820		249817_at	1.72	1.11	MD-2-related lipid recognition domain-containing protein	
At5g48850		248676_at	3.07	0.91	binding/MS5 family protein	
					Other functions	7
At3g48350		252365_at	1.79	1.11	cysteine-type peptidase activity/ proteolysis	
					Proteolysis	1
At3g16470	JR1	259383_at	1.72	1.01	response to wounding/ response to JA stimulus	
					Stress response	1
At5g67300		246987_at	1.79	1.06	transcription factor activity	
At3g26790	FUS3	258258_at	2.71	0.91	transcriptional activator activity/ response to auxin and ABA	
					Transcription factor activity	2
At3g08860		258983_at	1.76	1.05	alanine-glyoxylate transaminase activity/ mitochondria	
At4g23600	CORI3	254232_at	2.22	0.93	transaminase activity	
					Transferase activity	2
At5g10180	AST68	250475_at	1.91	1.03	sulfate transporter activity	
At1g16310		262751_at	1.94	1.02	cation transporter activity	
At3g25260		257841_at	5.53	0.95	oligopeptide transport	
At4g04760		255295_at	3.28	1.04	carbohydrate transport	
					Transporter activity	4

Table A.5. Primers used in the qPCR analysis of up- or down-regulated genes in *tapnac* KO background

Gene	Forward Primer	Reverse Primer
18S rRNA	GTCATCAGCTCGCGTTGACTAC	GAGCGACGGGCGGTG
AMSexon2-3	CGGGATATATGCGGAGACTCTT	CCTAGTACAGATTGTTTCCTGCATAAA
OPR3exon1-2	TTCGATCTCTCTCATCGAGTGG	CCGTTCAACGCCCTGC
At1g13990	TCTCTTGAGGACTTGAGGGATAATAAATAG	TACACAGTATCACCCACGGTTAATTACT
At1g15580	GTCCATCTCCGGGAAGAAGAG	CGCCGGTTCACATTTCAAAT
At1g19230	CAGTCACGCCGATGGAATC	GCCGGCAAATGGAGATGAT
At1g22990	GCATGAGCCGATGTTATAATTTGT	TGCGTTGCTTCTCACATAATACAG
At1g52410	TTCTTGTACTGGCTTTGTCTTTGTG	CATCATCCTGACAAGAAACCTCAT
At1g59740	CCACGTCGAAGCATCTTCACT	GACGCAAGCTTTCTTTTTAATATGG
At2g19500	CGC GAG CATGTCAACATTTC	TCCTCTTGTTAGTAACGTCCTTGAATT
At3g15730	CACGGGACTTTACATGCTACCA	CTTGCCTAACACCACCACCAT
At3g25260	GGTAAGATGCTATGCATGTAGTGGAA	ACCACCACACAAAAAAAACCATT
At3g30720	GGTTCATTTTGCCTCACACTTCT	CCCATGATATGACCCTCATTTTG
At3g61930	CGC GATTAAAGTGAGGATGAGAA	CGACGCGCCAAGTACAAA
At4g04760	GCTGGAATGTTCTTGGGTTGTC	TGATTCTCCATGCCTTCAAGAAG
At5g10180	CATGGCCAACTCAGGTTCATC	ACCGGTCAGGCTGGTCTTG
At5g47330	CCTATGTGTGGTTCTGGTTTATTCTG	TGAAGTCGCTATAGATATCTCCCTTGAT
At5g48850	GAGACTCCCAATCTTCGAACAGA	GAAAACACAGAATGAATTAACTAGCAAAC
At5g61620	GGTGACGGTGGCGAAAAC	AGACAAGTCCGTGCGTTATGG
At5g63850	AGATGATGACCCAAAACACTCAAA	AATGAGGAGATGGGTGGTGAGA

Primers listed were designed by Primer Express version 3.1 (ABI)

VITA

Veria Alvarado Chavez
Biology Department
BSBE 201 Texas A&M University
E-mail: valvarado@mail.bio.tamu.edu

EDUCATION

M.S. University of California, Davis, Vegetable Crops 2000.
B.S., Universidad Nacional Agraria, La Molina, Peru, Biology, 1994.

WORK EXPERIENCE

Graduate research assistant, Texas A&M University, 2000-2007
Graduate teaching assistant introductory biology and cell molecular biology, Texas A&M University, Fall 2003, Fall 2004.
Graduate teaching assistant, Seed production and quality, plant genetics and biotechnology laboratory, UC. Davis, 1999-2000
Internship Seed company, Nunhen, The Netherlands, Sept-Dec 1998.
Research Assistant, International Potato Center, Lima, Peru 1994-1997.
Research Assistant, Laboratory of Plant Phys UNALM, Peru 1992-1994.

PEER-REVIEWED PUBLICATIONS

Veria Alvarado and Kent.J. Bradford (2002). A hydrothermal time model explains the cardinal temperatures for seed germination. Plant Cell and Environment **25**, 1061-1069.

Kent J. Bradford, A. Bruce Downie, Oliver H. Gee, Veria Alvarado, Hong Yang, and Peetambar Dahal (2003). Abscisic acid and giberellin differentially regulate expression of genes of the SNF1-related kinase complex in tomato seeds. Plant Physiol. **132**, 1560-1576.

Bradford, K. J. and Alvarado, V. (2005). Hydrothermal time analysis of seed dormancy in true (botanical) potato seeds. Seed Science Research **15** (2), 77-88

AWARDS AND HONORS

Good Neighbor Scholarship: 2002-2006
MEPS Fellowship for First Year Students 2001
Research Award from the Vegetable Crops Program, UC Davis 1998