University of Alberta

Isolation and Characterization of Arginase Promoters in Loblolly Pine (*Pinus taeda* L.) and *Arabidopsis thaliana*

by

Disa Lynn Brownfield

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the degree requirements for the degree of Doctor of Philosophy

in

Plant Biology

Department of Biological Sciences

Edmonton, Alberta

Fall, 2007

Your file Votre référence
ISBN: 978-0-494-32926-9
Our file Notre référence
ISBN: 978-0-494-32926-9

ASTRACT

Arginase hydrolyzes the amino acid arginine and produces the non-protein amino acid ornithine and urea. In plants, arginase has been associated with the mobilization of seed storage reserves to provide the seedling with nutrients. To understand how arginase is regulated in loblolly pine we isolated two upstream regions that both contain an intron in the 5' UTR and several sets of repeated sequences. PCR using individual megagametophytes indicated that these upstream regions either represent two genes that are closely linked or are alleles of a single gene.

In angiosperms, arginase, in addition to its role in early seedling growth, has also been found in non-seed related tissues. In the model angiosperm, *Arabidopsis thaliana,* there are two arginase genes (A*RGAH1* and *ARGAH2*). We isolated the upstream regions of each Arabidopsis arginase and placed each in front of a GUS reporter gene and generated transgenic plants. We observed staining in immature leaves of seedlings and rosette stage plants and in several floral organs. Plants containing the largest *ARGAH1* fragment stained the pollen and this expression of *ARGAH1* in pollen was confirmed by quantitative PCR. We also identified two types of cis-acting elements that may have a role in the expression of *ARGAH1* in pollen.

In tomato, one of the arginase genes has been shown to be induced in response to wounding and MeJA treatment. We have shown that Arabidopsis is similar in this respect as *ARGAH2* is responsive to wounding and MeJA. We also completed

microarrays comparing *argah2* plants to wild-type plants in both control and MeJA treated plants to identify genes that were differentially expressed. We identified several genes including a pathogenesis related gene (PR-1). PR genes are frequently used as markers for the induction of pathogenesis in plants.

Together these data show that arginase may have roles in defense, early leaf, floral and pollen development. These are in addition to arginase's already described role in seed storage reserve mobilization. Thus, arginase is an enzyme involved in several important and diverse physiological processes in plants.

Table of Contents

List of Tables

List of Figures:

Chapter 3: Arginase in loblolly pine

Chapter 4: Analysis of Arginase in *Arabidopsis thaliana*

List of Abbreviations

ABA	abscisic acid
ADC	arginine decarboxylase
Arg	arginine
Asn	asparagine
AS	asparagine synthetase
β-ME	β-Mercaptoethanol
bp	base pairs
DAI_{30}	days after imbibiton at 30°C
DIC_{22}	days in culture at 22°C
dSAM	S-adenosylmethionine
EST	expressed sequence tag
EtOH	ethanol
Fwd	forward
Fd	ferredoxin
g	gravity
Gln	glutamine
Glu	glutamate
GOGAT	glutamate synthase
GS	glutamine synthetase
GUS	β-glucuronidase
HR	hypersensitive response
kb	kilobases
kD	kilodalton
LAT	late anther tomato
mBAC	mitochondrial carrier family
MCF	mitochondrial carrier family
MeJA	methyl jasmonate
Met	methionine
MS	Murashige and Skoog

N	nitrogen
NAD	nicotinamide adenine dinucleotide
NO	nitrogen oxide
NOS	nitric oxide synthase
OAT	ornithine aminotransferase
ODC	ornithine decarboxylase
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PEG	polyethylene glycol
pI	isoelectric point
PR	pathogenesis related
Pro	proline
RACE	rapid amplification of cDNA ends
Rev	reverse
rpm	revolutions per minute
RT-PCR	reverse trascriptase polymerase chain reaction
qPCR	quantitative PCR
SA	salicyclic acid
SAR	systemic acquired resistance
SDS	sodium dodecyl sulfate
T_2	second generation transformed plant
T_3	third generation transformed plant
TAG	triacylglycerol
TAIR	the Arabidopsis information resource
TIGR	the Institute for Genomic Research
TMV	tobacco mosaic virus
UAS	upstream activation sequence
URS	upstream repression sequence
UTR	untranslated region

Chapter 1: Literature review

1.1 Arginine

The amino acid arginine (Arg) is a structural component of proteins and a medium for the transport and storage of nitrogen (Micallef and Shelp 1989). Arg is also an intermediate of the urea cycle and a precursor of other molecules such as nitric oxide (NO) via nitric oxide synthase (NOS, EC 1.14.13.39), or polyamines via Arg decarboxylase (ADC, EC 4.1.1.19) (Jenkison et al. 1996). Arg can be broken down through the action of arginase (L-Arg amidinohydrolyase, EC 3.5.3.1), which hydrolyzes Arg to produce ornithine and urea. In plants, urea is degraded (Zonia et al. 1995) while ornithine can be utilized in the urea cycle, or as a precursor of polyamines (via ornithine decarboxylase, EC 4.1.1.17) and proline (Pro) (Hanfey et al. 2001).

1.2 Enzymatic activity of arginase

The arginase family includes four groups of enzymes: arginases, agmatinases, formiminoglutamases, and proclavaminate amidino hydrolyases (Perozich et al. 1998). Figure 1.1 shows the reactions catalyzed by this family of enzymes. Arginase and agmatinase catalyze relatively similar reactions; however, arginase utilizes the amino acid Arg as its substrate while the polyamine agmatine is the substrate for agmatinase. Both the substrate and the product of the agmatinase reaction

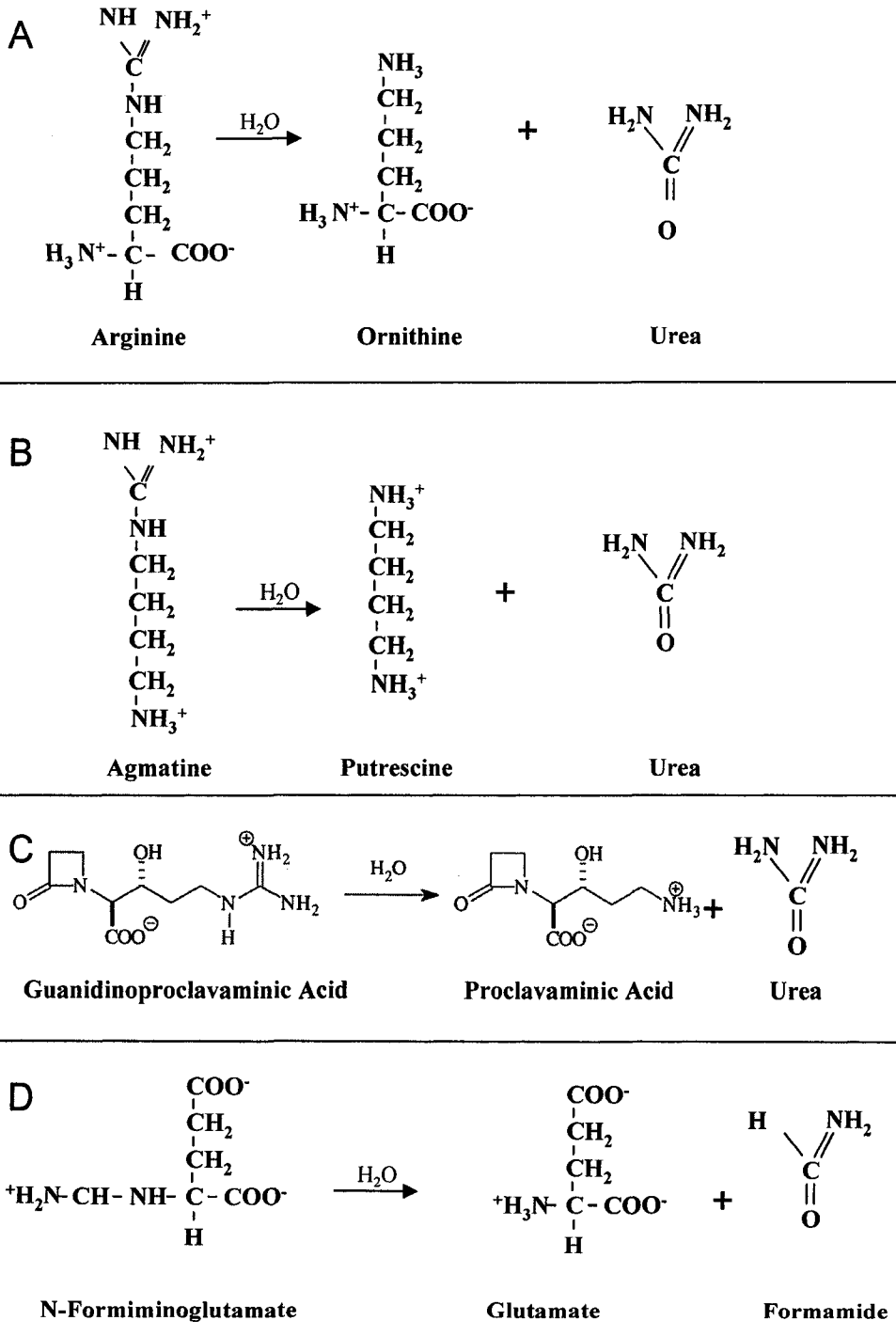


Figure 1.1: The reactions of the arginase family of enzymes. Adapted from Perozich et al. (1998). The arginase reaction. B. The agmatinase reaction. C. The proclavaminate amidino hydrolyase reaction. D. The formiminoglutamase reaction.

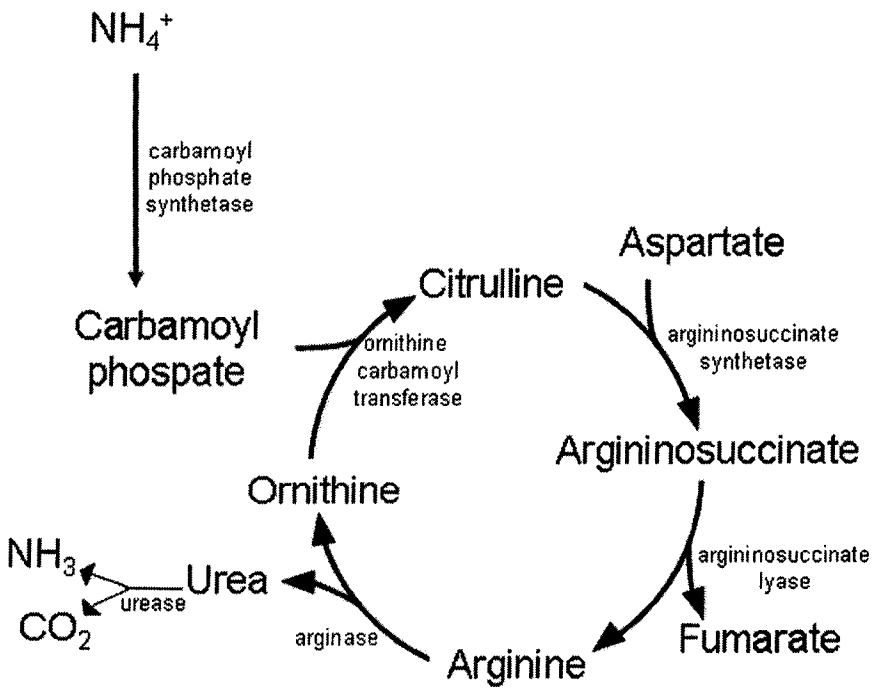


Figure 1.2 The Urea Cycle. Adapted from Jenkinson et al. 1996. In mammals the urea is excreted but in plants the urease enzyme degrades urea into ammonia and carbon dioxide.

agmatinase reaction are polyamines, a class of molecules which will be discussed later.

Arginases are important in organisms from all primary kingdoms (Jenkison et al. 1996). Arginase is one of the 5 key enzymes of the urea cycle (Figure 1.2), which in mammals functions to remove toxic waste products (NH_4^+). In addition, it has been suggested that the urea cycle also has a role in maintaining a homeostatic pH by removing bicarbonate and ammonium ions (Atkinson and Bourke 1984). There are two isozymes (and genes) of arginase in mammals; type I arginase is expressed almost exclusively in the liver of healthy adult mammals and functions in the urea cycle (Morris 2002). Type II arginase is expressed at low levels and almost ubiquitously in healthy human adults and contributes to arginine homeostasis (Morris 2002). Possible roles for type II arginase in the body include the synthesis of polyamines, Pro, and glutamate (Glu). In addition, arginase may play an important role in the regulation of Arg availability for NO synthesis (Morris 2002).

Based on phylogenetic data, plant arginases appear to be more closely related to the agmanitase group of enzymes than to arginases from vertebrates and bacteria (Figure 1.3) (Perozich et al. 1998, Chen et al. 2004); it has therefore been suggested that plant arginases are not authentic arginases (Perozich et al. 1998). However, an antibody produced from a loblolly pine (*Pinus taeda* L.) arginase was able to immunotitrate arginase activity (Todd et al. 2001a), indicating that this enzyme was indeed a functional arginase. More recently, Chen et al. (2004) showed that the tomato (*Lycopersicon esculentum*) arginases *LeARG1* and *LeARG2* have high activity and exhibit strong substrate specificity for L-Arg over other compounds including

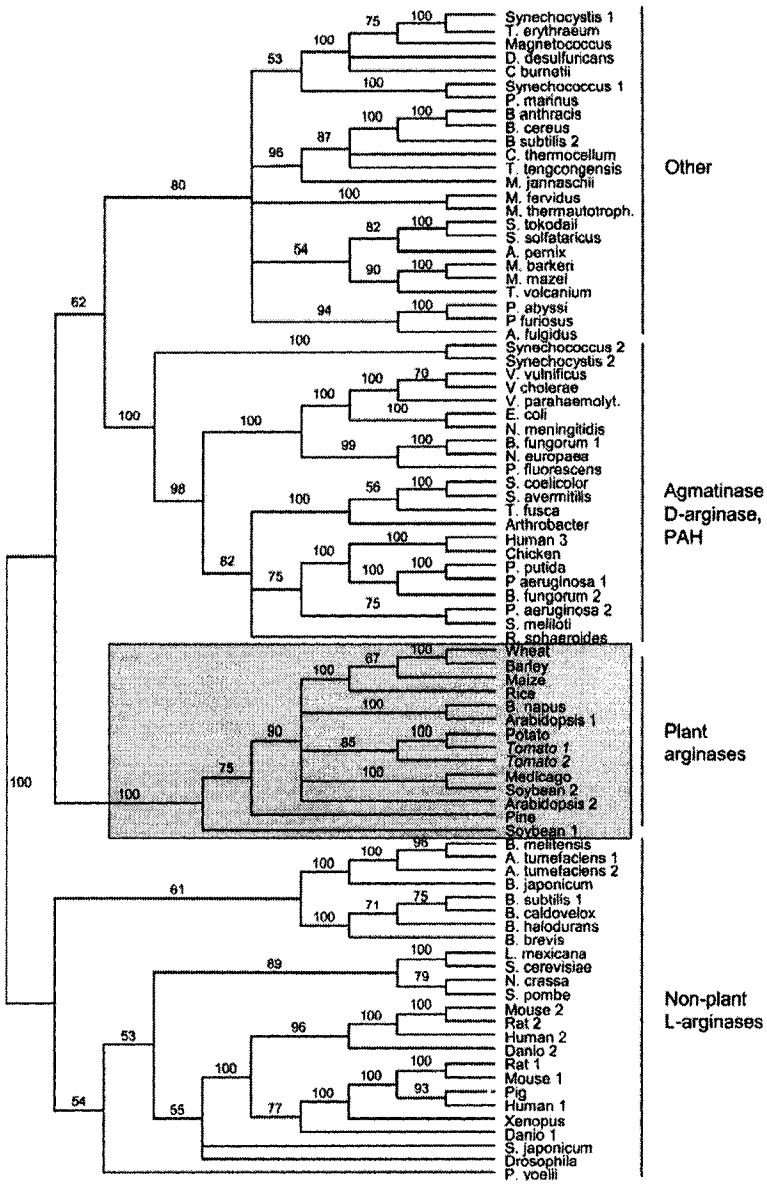


Figure 1.3: The arginase family phylogenetic tree. Adapted from Chen et al. 2004 Tree members were identified from the literature and from BLAST searches of EST databases.

agmatine, indicating again that plant arginases are indeed arginases and not agmanitases. This information along with the close phylogenetic relationship plant arginases have to each other (Chen et al. 2004) suggests that plant arginases are a distinct group of functional arginases.

Arginase activity has been described in several plant species including (but not limited to): bean (Kollöffel and van Duke 1975), pea (Taylor and Stewart 1981), iris (Boutin 1982), artichoke (Wright et al. 1981), pumpkin (Splittstoesser 1969), soybean (Downum et al. 1983), tomato (Alabadí et al. 1996), Arabidopsis (Zonia et al. 1995) and loblolly pine (King and Gifford 1997). Purification of the arginase enzyme has been reported in fewer plant species including, soybean (Kang and Cho 1990) and loblolly pine (Todd et al. 2001); several forms of arginase have also been purified from the lichen *Evernia prunasrti* (Legaz and Vincente 1982, Martín-Falquina and Legaz 1984, Planelles and Legaz 1987). At present, there are no reports describing the purification of arginase from Arabidopsis, but a previously identified Arabidopsis cDNA predicted a 36.5 kD protein with similarity to known arginases (Krumplemen et al. 1995).

Three forms of the arginase enzyme have been purified from the lichen *E. prunastri*: two forms which are only present in the presence of arginine (i.e. inducible) and a constitutive form (Legaz and Vincente 1982, Planelles and Legaz 1987). The constitutive arginase is expressed under most conditions but its activity increases in the presence of Arg. All forms are inhibited by the presence of urea, one of the products of arginase (Legaz 1991).

Two arginases have been purified from *Neurospora crassa* (Marathe et al. 1998). One arginase is present when grown on minimal and Arg supplemented media, while the other form is produced only in the presence of Arg. In *N. crassa* and other organisms that lack a urea cycle, arginase allows the utilization of Arg as a nitrogen source (Marathe et al. 1998).

The kinetics of arginases vary, depending on their sources. The K_m of the constitutive arginase from *N. crassa* is significantly higher than any of the plant arginases isolated to date (Borkovich and Weiss 1987). Among the higher plants, the K_m of arginase from soybean axes is the highest followed by both tomato arginases and the arginase from soybean cotyledons. All three arginases from *E. pruastri* have lower K_m values than any of the isolated plant arginases (Legaz and Vincente 1980, Martin-Falquina and Legaz 1984, Planelles and Legaz 1987).

Most arginases appear to require Mn^{2+} for optimal activity, as demonstrated in soybean (Kang and Cho, 1990). The requirement for Mn^{2+} was also shown for all three arginases in *E. prunastri* (Legaz and Vincente 1982, Planelles and Legaz 1987), and loblolly pine (King 1998, Todd et al. 2001a). Removal of Mn^{2+} from reactions involving the loblolly pine arginase significantly decreased the enzyme activity, but activity was restored to normal levels upon the re-incorporation of Mn^{2+}. Other divalent cations could not restore the activity of the loblolly pine arginase (King 1998, Todd et al. 2001a).

1.3 Arginase regulation and sub-cellular localization

Arginase has been reported to be localized in the mitochondria of several plant species including pea (Taylor and Stewart 1981), bean (Kollöfeel and van Duke 1975), soybean (Goldraij and Polacco 2000), Arabidopsis (Todd et al. unpublished data). In addition, N-terminal mitochondrial localization signals are present in the loblolly pine and tomato cDNA sequences (Todd et al. 2001b, Chen et al. 2004). C-myc tagged arginases of Arabidopsis are localized to the mitochondrial matrix of tobacco cells (Todd et al. unpublished data). The extra-hepatic mammalian arginase (AII) is also localized to the mitochondria; however, arginase I (liver-type) is cytoplasmic (Turner and Weiss 2006). Other cytoplasmic arginases include those from *N. crassa* (Marathe et al. 1998) and yeast (Davis 1986).

Regulation of arginase activity is necessary in part to avoid a futile cycle of concurrent Arg anabolism and catabolism. In soybean (*Glycine max*) seedlings, 30% of ^{14}C labeled Arg taken up by mitochondria is quickly hydrolyzed. On the other hand, in mitochondria of soybean embryos, very little arginase activity is detected even though these embryos contain both a large pool of Arg and a protein that has arginase activity *in vitro* (Goldraij and Polacco 2000; Micallef and Shelp 1989). In addition, Arg uptake into mitochondria is similar in both germinating seedlings and developing embryos, although arginase is active only in the seedlings (Goldraij and Pollaco 1999). Therefore, a mechanism to regulate arginase activity must exist.

One proposed method to regulate the degradation of Arg would be to control its transport into the mitochondria. However, at the present time little is known about

Arg import into the mitochondria. The outer mitochondrial membrane is permeable to molecules of 5000 daltons or less (Alberts et al. 1994). Therefore, specific import into the mitochondrial matrix is accomplished by the inner mitochondrial membrane transporters. Transport of Arg into the mitochondrial matrix in Arabidopsis is accomplished at least in part by 2 mitochondrial basic amino acid carriers (mBAC) which are members of the mitochondrial carrier family (MCF) (Hoyos et al. 2003). AtmBAC1 and AtmBAC2 function to transport Arg, ornithine, and histidine, with Arg being the most efficiently transported substrate (Palmieri et al. 2006). RT-PCR of *AtmBAC1* and *2* in different tissues showed that *AtmBAC1* is expressed at a low level in the stem, slightly higher in the leaf and then higher again in the flower, silique and seedling (Hoyos et al. 2003). *AtmBAC2* had the lowest expression in the stem and the seedling and then higher in the leaf, flower and silique (Hoyos et al. 2003). Promoter fusions of *AtmBAC2* showed GUS staining in rosette leaves as well as in pollen and the siliques of mature plants (Catoni et al. 2003). As these transporters are expressed in a tissue specific manner it may be possible that they have a role in regulating the availability of Arg; however, this has not yet been demonstrated.

The physical separation of Arg biosynthetic and catabolic enzymes (Figure 1.4) has been reported in pea (*Pisum sativum*) where the biosynthetic enzymes were predominately chloroplastic and cytoplasmic while catabolic enzymes were localized to the mitochondria (Taylor and Stewart 1975). A similar situation is present in *N. crassa* where Arg is compartmentalized in the vacuole, while the arginases are

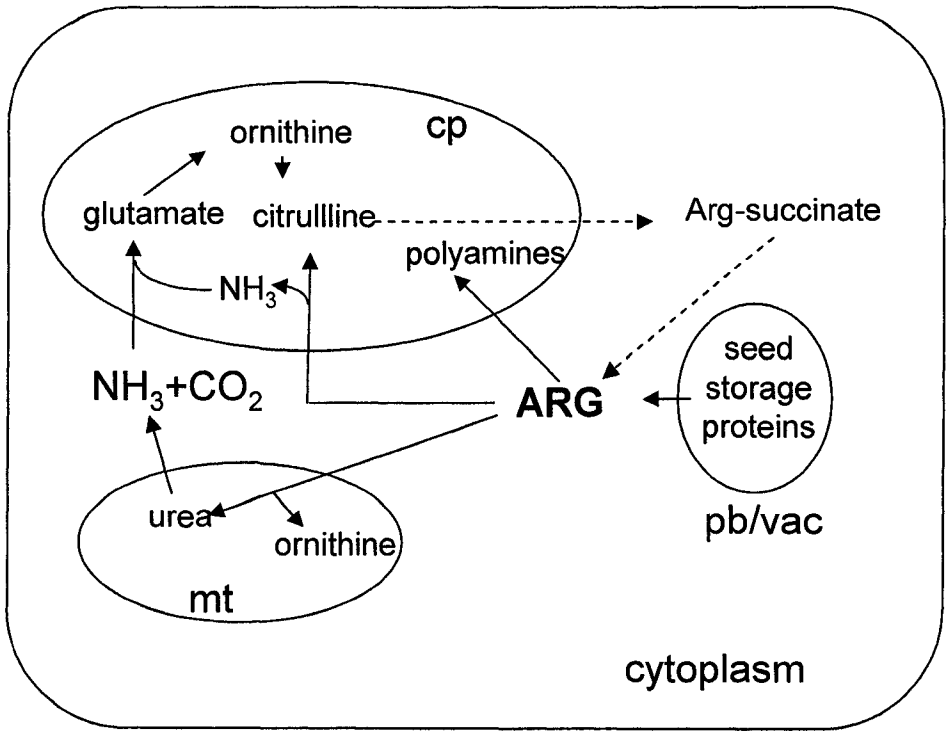


Figure 1.4: The localization of arginine metabolic pathways in germinating soybean (*Glycine max*) cotyledons. Adapted from Goldraij and Polacco (2000). (mt) mitochondria, (cp) chloroplast, (pb/vac) protein body/vacuole

cytoplasmic (Davis 1986). Thus, compartmentalization appears to have a role in regulating Arg anabolism and catabolism.

1.4 Genetic structure of arginase

cDNAs encoding arginase have been cloned from many higher plants, including loblolly pine, soybean, tomato, and Arabidopsis. The 1366bp cDNA isolated from a loblolly pine expression library contained two potential translation initiation codons and a poly A tail. In addition, *in vitro* translation of the cDNA produced a protein that was similar in size to the loblolly pine arginase protein (Todd et al. 2001b). Together, these results indicated that the pine cDNA sequence was likely full length. High-stringency Southern blotting of loblolly pine genomic DNA with arginase cDNA suggested that arginase was present as a single copy gene; however, as low-stringency hybridization conditions were not tested it is possible that more than one arginase is present in loblolly pine (Todd et al. 2001b).

In soybean, Southern analysis detected several arginase genes even under high stringency conditions (Goldraij and Polacco 1999). In both tomato and Arabidopsis, there are 2 genes encoding functional arginases (Krumpelman et al. 1995, Todd et al. unpublished data, Chen et al. 2004). The two tomato arginase cDNAs were expressed in *E.coli* and purified. Enzyme assays showed that both of these arginases catabolized the conversion of Arg into urea and exhibited specificity for L-Arg over other substrates tested (Chen et al. 2004).

The two arginases were also identified in Arabidopsis (*Arabidopsis thaliana).* The first arginase cDNA was identified by complementation of a yeast (*Saccharomyces cerevisiae*) arginase deficient (*car1*) mutant (Krumpelman et al. 1995). The yeast mutant (*car1*, *dur1*, *ura*3-52) was unable to grow on either Arg or urea as a sole N source, but growth on plates with Arg as the sole N source was observed with complementation of an Arabidopsis arginase cDNA. Sequence homology to other arginases, as well as the ability of the complemented mutant to accumulate urea in the presence of Arg confirmed its identification as an arginase. This gene was named *ARGAH1* (At4g08900) in order to differentiate it from the *ALTERED RESPONSE TO GRAVITY* (*ARG1*) gene (www.arabidopsis.org). A second putative arginase (At4g08870), was identified by sequence homology to *ARGAH1* and was named *ARGAH2*. Like *ARGAH1*, the ORF of *ARGAH2* was able to complement the *car1* mutation in yeast, allowing growth on Arg as the sole N source (C. Todd unpublished data).

*ARGAH1*and *ARGAH2* are both located on chromosome 4 and are approximately 55kb apart (Todd et al. unpublished data). Relative to each other, *ARGAH1* and *ARGAH2* are transcribed from opposite strands of the chromosome. Based on the similarity of these gene sequences (88% identity of cDNA sequences, 76% gene identity), it likely that these genes are the result of a duplication event. This may be the result of segmental duplication or transposition. However, the genes immediately surrounding each arginase do not appear to be related, indicating that if the arginases arose through a segmental duplication, this event must have been ancient enough to allow for further significant rearrangements at these loci.

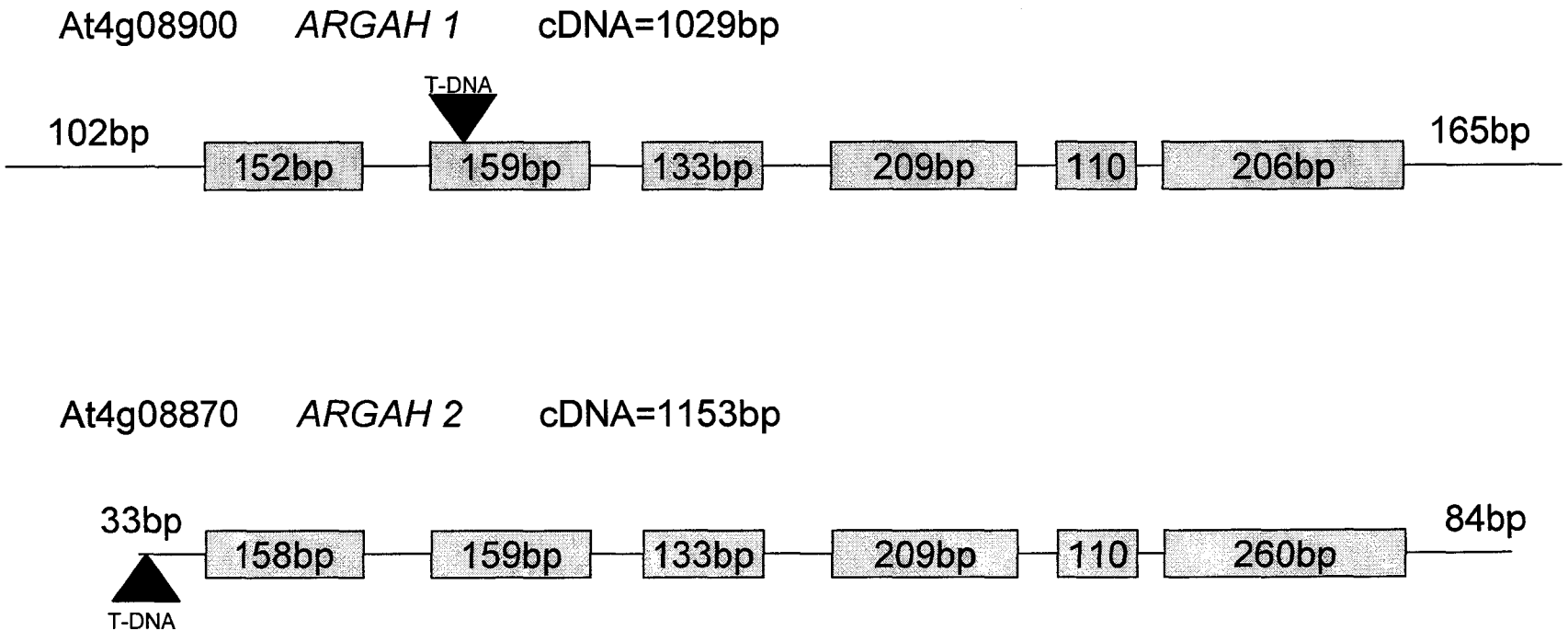


Figure 1.5: The Arabidopsis arginase genes. Predicted exons are indicated by the boxes and the Intron sequences by the lines based on sequence information from TAIR. The predicted size of the 5' and 3' UTRs are labelled by their size at their respective ends. The location of the T-DNA insertion for each line (SALK_057987 and SAIL_181C11) is marked by the black triangle.

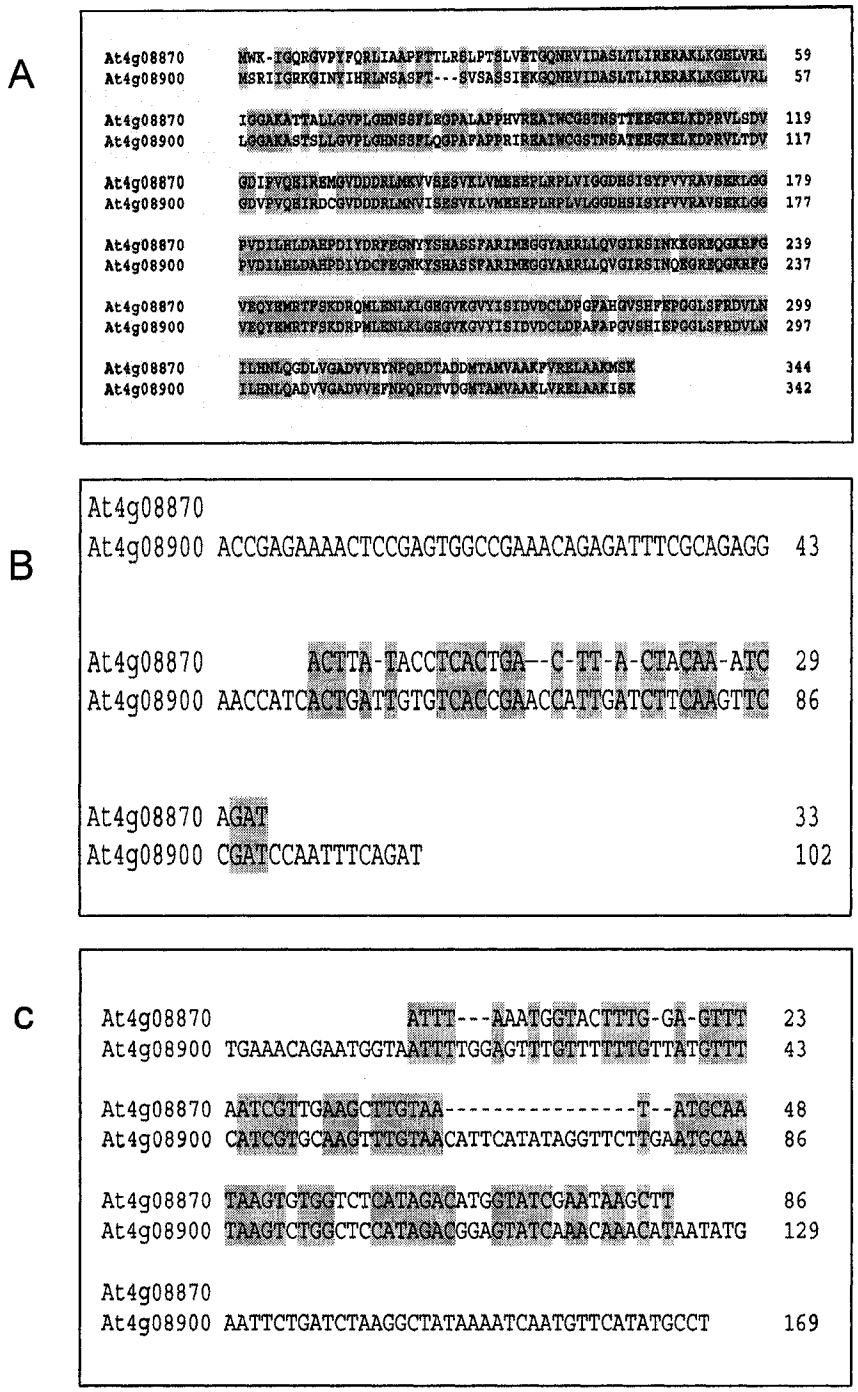


Figure 1.6: Alignment of arginase sequences from Arabidopsis. Regions of sequence similarity are highlighted. A. Alignment of the predicted protein sequences of *ARGAH1* (At4g08900) and *ARGAH2* (At4g08870) genes. B. Alignment of the 5' untranslated regions of the *ARGAH1* (At4g08900) and *ARGAH2* (At4g08870) genes. C. Alignment of the 3' untranslated regions of the *ARGAH1* (At4g08900) and *ARGAH2* (At4g08870) genes

The predicted structure of both *ARGAH1* and *ARGAH2* genes (Figure 1.5) includes a short 5' untranslated (UTR) region followed by 6 exons ranging in size from 110 base pairs to 260 base pairs and a 3' UTR (www.arabidopsis.org). The N termini of the predicted proteins are not highly conserved, with only 13 of 33 amino acids identical (Figure 1.6A). However, among the remaining 311 amino acids, 285 are identical (Todd et al. unpublished data). As expected, the 5' and 3' UTRs are less conserved than the coding regions (Figure 1.6B and C). Thus, the *ARGAH1* and *ARGAH2* genes can be distinguished based on their 5' and 3' UTRs.

1.5 Arginase promoters

Arginase promoter isolation and characterization has been completed in several species including mice (Shi et al. 1998), humans (Takiguchi et al. 1988) and yeast (Dubois and Messenguy 1997). The best characterized arginase promoter is from the yeast *CAR1* gene; however, the yeast arginase is not closely related to plant arginases. A phylogenetic tree of the arginase superfamily generated from EST databases contains four main clusters (Figure 1.3, Chen et al. 2004). The plant arginases are clearly grouped together and away from vertebrate, bacterial, and fungal (including yeast) arginases which are in a separate cluster of the tree. This evolutionary distance between the yeast arginase and plant arginases makes it unlikely that the promoters have fully conserved properties; however, the yeast promoter does provide an example of how arginase is regulated in response to varying environmental conditions.

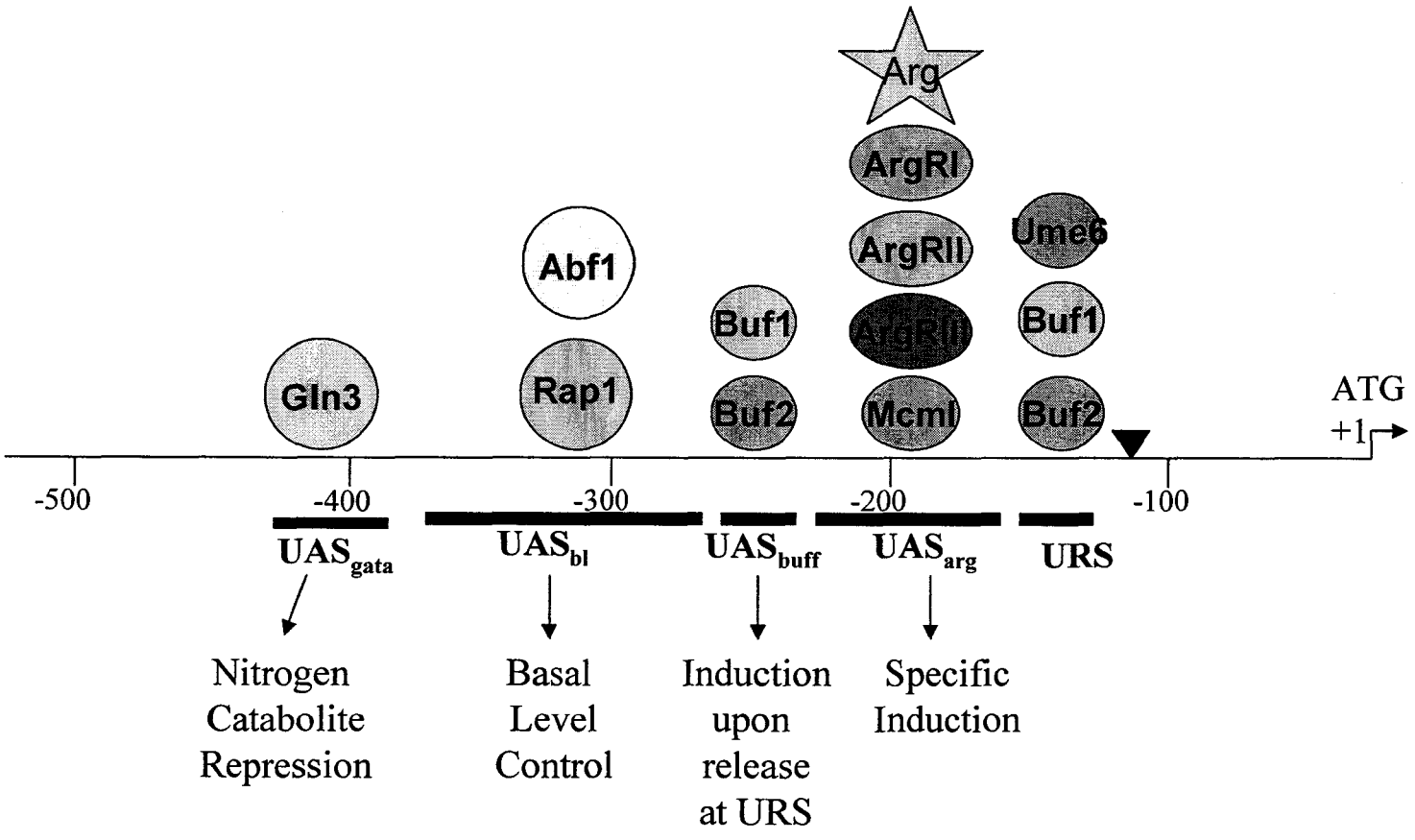


Figure 1.7: The arginase (*CAR1*) promoter from yeast (*Saccharomyces cerevisiae).* Adapted from Dubois and Messenguy 1997. The promoter is shown with the locations of the regulatory sequences are marked by the thick black lines, the protein(s) and amino acids that bind each sequence are shown above.

In yeast, arginase production is induced when cells are grown on Arg as the sole nitrogen source, but repressed in the presence of a preferred nitrogen source such as glutamine or asparagine (Dubois and Messenguy 1997). In the yeast *CAR1* promoter, there are five regions that control arginase expression (Figure 1.7): which include four upstream activation sequences (UAS) and one upstream repression sequence (URS). During basal level transcription of arginase, which occurs in the presence of preferred nitrogen sources, the basal level UAS (UAS_{bl}) is bound by the Abf1 and Rap1 proteins (Dubois and Messenguy 1997). In the presence of an optimal nitrogen source (with or without the presence of Arg), the yeast *CAR1* gene is expressed only at a basal level, and increased expression of arginase due to the presence of Arg is prevented. In order to accomplish this, two events occur. First, the URS is bound by the Buf1, Buf2 and Ume6 proteins, which prevents expression above the basal level; second, the Gln3 protein is prevented from activating transcription through the UAS_{gata} element by a high concentration of optimal nitrogen (Dubois and Messenguy 1997). Thus, in the presence of an optimal nitrogen source, the basal level of control allows arginase to be expressed at a low level and nitrogen catabolite repression (from UAS_{gata}) along with the URS prevents expression above the basal level (Dubois and Messenguy 1997).

In the presence of Arg as the sole nitrogen source, four events occur to allow the yeast *CAR1* gene to be maximally expressed. First, the Buf 1-Buf 2-Ume6 repression at the URS is overcome or eliminated. This frees the Buf1 and Buf2 proteins and allows them to bind to the UAS_{buf} element which induces arginase. However, when the Ume6 protein is released from the URS there is no evidence to

show that it participates in the UAS_{buf} induction of arginase with the Buf1 and Buf2 proteins. The UAS_{buf} does not appear to function as a UAS unless the repression from URS is released. Third, with no high concentration of optimal nitrogen present, there is nothing to prevent Gln3 from binding to UAS_{gata}, thus nitrogen catabolite repression is released. Finally, the presence of Arg induces the binding of the ArgRI-ArgRII-ArgRIII-McmI complex to the UAS_{arg} thus arginase expression is increased by specific induction (Dubois and Messenguy 1997).

Although considerable work has been performed towards understanding the arginase promoter in yeast and the isolation of arginase promoters from several other species, there have been no published reports to date on the isolation or characterization of any plant arginase promoters.

1.6 Products of arginase activity

The amino acid Arg is present in proteins and, as described earlier, is a substrate for the production of nitric oxide (NO), creatine, and the polyamines (Figure 1.8) (Wu et al. 1998). In addition, Arg is hydrolyzed by arginase to produce ornithine and urea. Ornithine can be used to produce polyamines and the amino acids Pro, Arg and Glu. Thus, Arg and the products of the arginase reaction are utilized in many ways in the plant (Catoni et al. 2003).

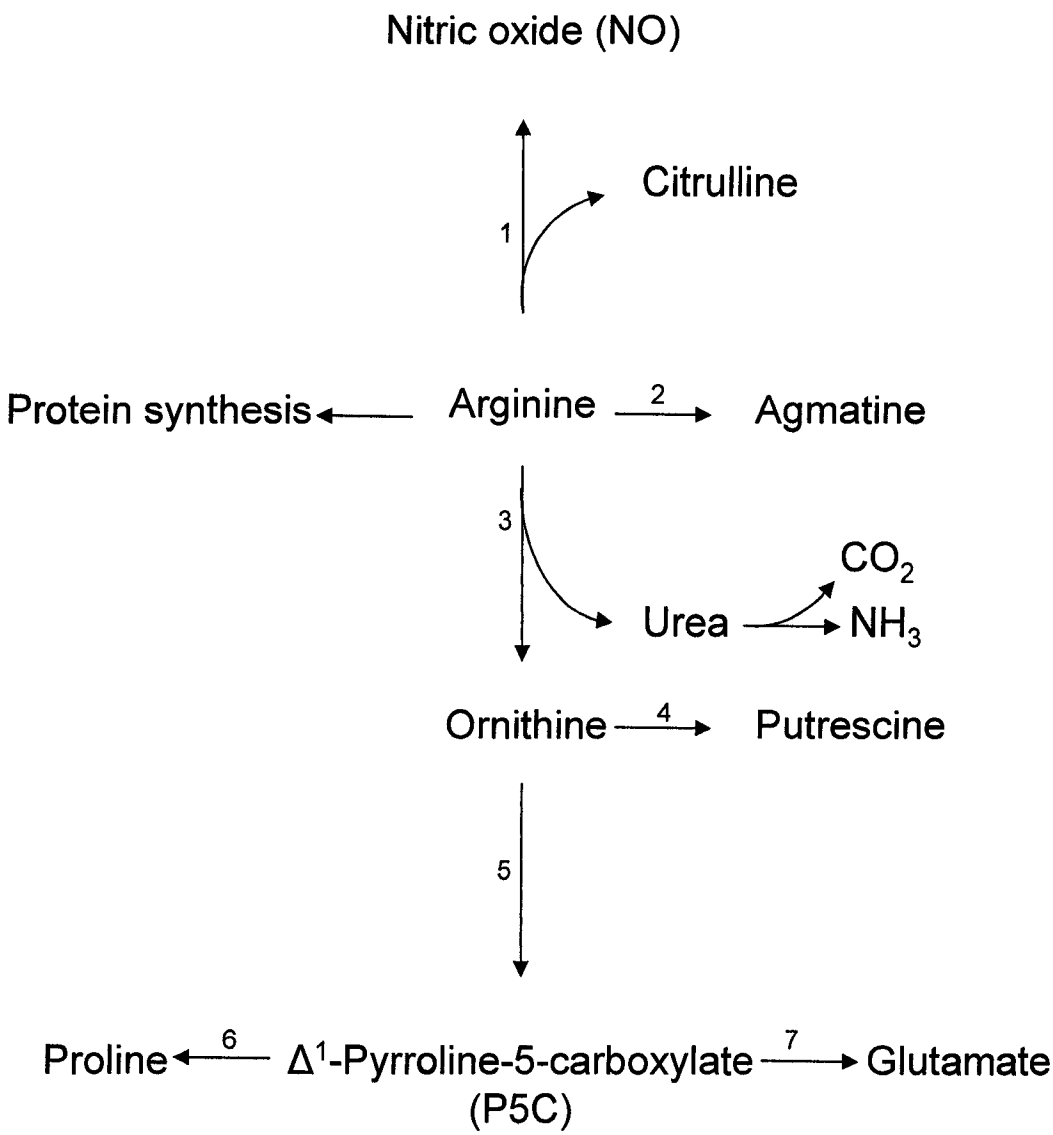


Figure 1.8: The possible fates of arginine. Adapted from Wu and Morris 1998 1. Nitric oxide synthase 2. Arginine decarboxylase 3. Arginase 4. Ornithine decarboxylase 5. δ-ornithine aminotransferase 6. P5C reducatase 7. P5C dehydrogenase.

1.6.1 Urea

Depending on the organism, the urea produced from the hydrolysis of Arg has different fates. In mammals, the cytosolic liver-arginase is active in the urea cycle which eliminates excess nitrogen. In plants, the urea released by arginase is not excreted but first degraded further through the action of urease into CO_2 and NH_3 (Sirko and Brodzik 2000). The nitrogen released as NH_3 is then reincorporated, allowing plants to conserve a potentially limiting nutrient (Cantón et al. 2005)

Other processes also generate nitrogen molecules that are subsequently recycled. Photorespiration generates high levels of ammonium during the regeneration of 3-phosphoglycerate as does the synthesis of phenylpropanoids such as lignin (Cantón et al. 2005). Re-assimilation of nitrogen also occurs in response to pathogen attack and leaf senescence (Cantón et al. 2005). All of these processes release nitrogen-containing compounds which, when recycled, allow the plant to reuse the nitrogen already present in its biomolecules.

Nitrogen recycling in plants is accomplished through the GS/GOGAT cycle (Figure 1.9), which is the same cycle involved in primary nitrogen assimilation (Gómez-Maldonado et al. 2004). GS/GOGAT is the collaboration of two enzymes: glutamine synthetase (GS, EC 6.3.1.2) and glutamate synthase (GOGAT, EC1.4.7.1 or 1.4.1.14). GS, using ATP as an energy source, adds an amino group from ammonium to glutamate and produces glutamine. GOGAT transfers the new amino group from glutamine on to 2-oxoglutarate and, using reduced ferredoxin (Fd) or

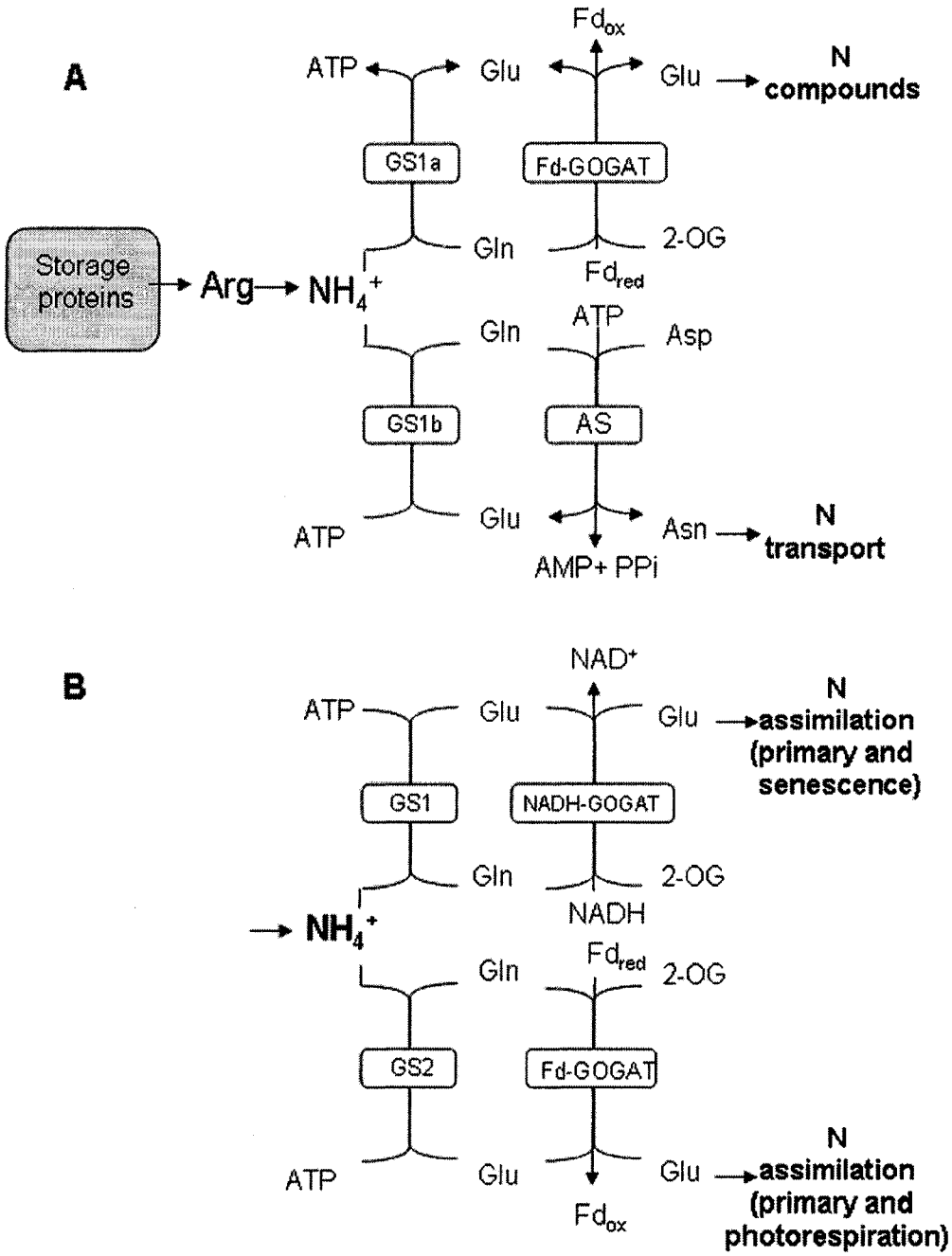


Figure 1.9. Nitrogen recycling in gymnosperms and angiosperms. Adapted from Avila et al. 2001. A. Nitrogen movement from storage protein degredation in the early stages of seedling development in conifers. B. Nitrogen movement in angiosperms.

NADH, produces 2 molecules of glutamate, one of which is re-used by GS and the second is available for metabolism, storage or transport (Tobin et al. 2001). Thus, ammonium is recycled into an organic molecule and is available for use in other metabolic functions.

In angiosperms, there are two classes of GS: GS1 and GS2. GS1 is a cytosolic enzyme encoded by a small gene family with a number of isoforms that are developmentally expressed and respond to various stimuli (Lam et al. 1996). GS1 isoforms are expressed in leaves and are especially abundant in the vascular tissue of roots, nodules, flowers and fruit (Cantón et al. 2005). Angiosperm *GS2* is usually encoded as a single gene whose product is localized to the chloroplast and is present predominately in leaves. There are also two forms of GOGAT in angiosperm species: Fd-GOGAT and NADH-GOGAT. NADH-GOGAT is present in the vascular tissues and, along with GS1, has been proposed to be involved in nitrogen assimilation from the soil and the re-assimilation of nitrogen mobilized during senescence (Cantón et al. 2005). Fd-GOGAT co-localizes with GS2 and is most common in photosynthetic tissues. Together their role involves the re-assimilation of nitrogen from photorespiration and nitrate (NO_3^-) reduction (Cantón et al. 2005).

The GS/GOGAT cycle in conifers is similar; however, there are some differences that exist. At present, no *GS2* has been identified, however two distinct forms of *GS1* have been identified in Scots pine: *GS1a and GS1b* (de la Torre et al. 2002). *GS1a* is expressed in seedling photosynthetic tissues and its amino acid sequence is similar to the angiosperm GS2, thus GS1a in conifers appears to be analogous to GS2 in angiosperms. The amino acid sequence of pine GS1b exhibits

more similarity to the angiosperm GS1 than the pine GS1a. Another difference that exists between GS in angiosperms and conifers is its localization. Both forms of GS in pine have been localized to the cytosol; thus, conifer GS1a and angiosperm GS2 appear to be localized in different compartments of the cell, despite their sequence similarity.

GS, GOGAT, and urease together allow the nitrogen released from Arg in the form of urea to be utilized. In addition, the GS/GOGAT cycle also allows the re-incorporation of nitrogen from senescence, photorespiration and lignin synthesis making GS/GOGAT important for nitrogen usage throughout plant development.

1.6.2 Ornithine

1.6.2.1 Proline biosynthesis

Proline synthesis is one possible use for the ornithine produced after arginase hydrolyzes Arg. In plants, there are two routes for synthesis of Pro (Figure 1.10). In the first, Pro is synthesized from Glu; this pathway has been well documented in plants and is similar to that of *E. coli* (Buchanan et al. 2000). Two enzymes P5CS (Δ^1-pyrroline-5-carboxylate synthetase, EC 2.6.1.13) and P5CR (Δ^1-pyrroline-5-carboxylate reductase, EC 1.5.1.2) are required to reduce Glu to Pro via the intermediate of P5C (Δ^1-pyrroline-5-carboxylate) (Roosens et al. 1998). In the second, less well characterized pathway, ornithine can also be used to make Pro; this is believed to involve δ-OAT (δ-ornithine aminotransferase, EC 2.6.1.13) and the

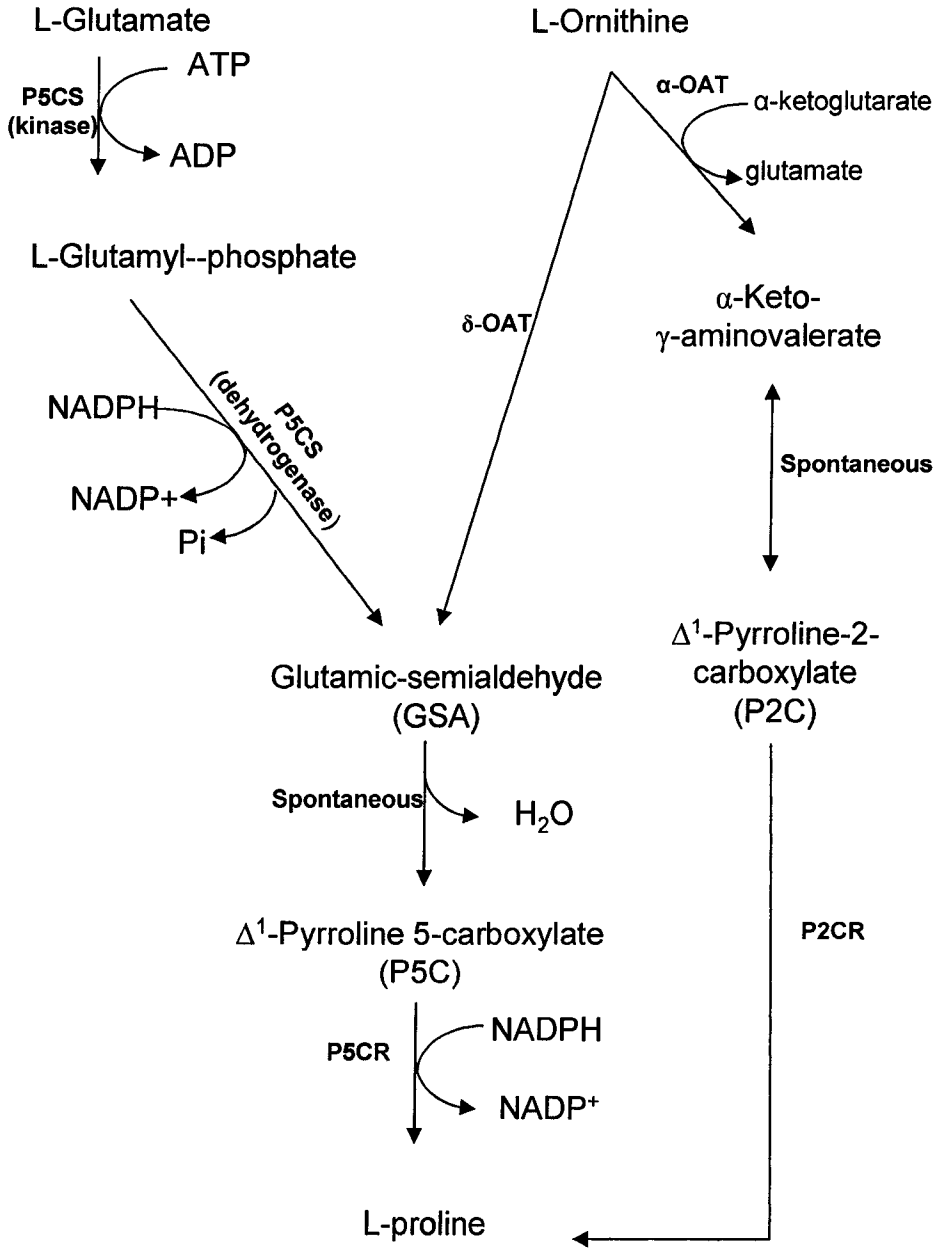


Figure 1.10: Proposed routes of proline synthesis. Adapted from Buchanan et al. 2000 and TAIR AraCyc (www.arabidopsis.org). Proline can be produced from glutamate or ornithine as a substrate. In the pathway from glutamate P5CS (Δ^1-Pyrroline-5-carboxylate synthetase) has both kinase and dehydrogenase activities and is able to conduct both steps of the reaction from gluatmate to GSA. P5CR (Δ^1-Pyrroline-5-carboxylate reductase) converts the intermediate P5C to proline. From ornithine there are 2 possible routes to produce proline via α–OAT (α-ornithine aminotransferase) which then utilizes the enzyme P2CR (Pyrroline-2-carboxylate reductase) or γ-OAT (γ-ornithine aminotransferase) to produce GSA as an intermediate.

intermediate P5C. Supporting the idea that δ-OAT is used to produce Pro from ornithine, a δ-OAT cDNA was isolated by complementation of Pro biosynthesis in *E. coli* mutants deficient in Pro biosynthesis with a moth bean (*Vigna acontifolia*) cDNA library (Delauney et al. 1993). One class of cDNAs recovered from the complementation screen showed high homology to other *δ-OATs*, but lacked homology to α-aminotransferases. This indicated that it was likely not α-OAT that is used to make Pro from ornithine via P2C (Δ^1-pyrroline-2-carboxylate) as a precursor.

In plants, both the Glu and ornithine pathways contribute to Pro synthesis under normal conditions (Delauney and Verma 1993). The synthesis of Pro is controlled by feedback inhibition of its end product in unstressed tobacco as well as barley leaves and the loss of this feedback inhibition during water stress results in increased Pro biosynthesis (Boggess et al. 1976, Noguchi et al. 1968). Early reports indicated that the Glu pathway was relatively more important during stress while the ornithine pathway became more important under excess nitrogen conditions (Rhodes et al. 1986). However, over-expression of an Arabidopsis *δ-OAT* in tobacco (*Nicotiana plumagnifolia*) resulted in transgenic plants with increased Pro biosynthesis, higher biomass, and higher germination rates under osmotic stress conditions (Roosens et al. 2002). This indicates that *OAT* is likely involved in stress-related Pro biosynthesis; thus, the production of ornithine via arginase may also have a role in stress responses. Pro accumulation under stress functions not only in a protective osmolytic role but in other potential roles as well (Delauney and Verma 1993). In bacteria, Pro increased the stability of proteins and membranes in water-stressed environments (Delauney and Verma 1993).

1.6.2.2 Polyamine biosynthesis

The ornithine produced by the hydrolysis of Arg by arginase can also be used in the synthesis of polyamines. Polyamines are organic cationic compounds containing two or more amino groups. The major forms of polyamines in plants include putrescine, spermidine, and spermine (Martin-Tanguy 2001) (Heldt and Heldt 1998). Polyamines are widespread, being present in both prokaryotic and eukaryotic organisms. In general, prokaryotic cells contain more putrescine with less spermidine and no spermine, whereas eukaryotic cells have some putrescine, with higher levels of spermidine and even higher levels of spermine (Bais and Ravishankar 2002).

In plants there are two pathways for polyamine synthesis (Figure 1.11). In the first, ornithine is decarboxylated by ornithine decarboxylase (ODC, EC 4.1.1.17) releasing carbon dioxide (CO_2) and the polyamine putrescine. In the second, Arg is decarboxylated by Arg decarboxylase (ADC, EC 4.1.1.19), and via the intermediates agmatine and N-carbamoylputrescine, putrescine is produced (Buchanan et al. 2000). Putrescine is a precursor for the production of the other two polyamines, spermidine and spermine, commonly found in plants (Hanfey et al. 2001). Spermidine synthase (EC 2.5.1.16) adds an amino-propyl group from decarboxylated S-adenosylmethionine (dSAM) to putrescine
to produce spermidine. Spermidine can then be used to produce spermine by the addition of another amino-propyl group from dSAM to spermidine by spermine synthase (EC, 2.5.1.22)

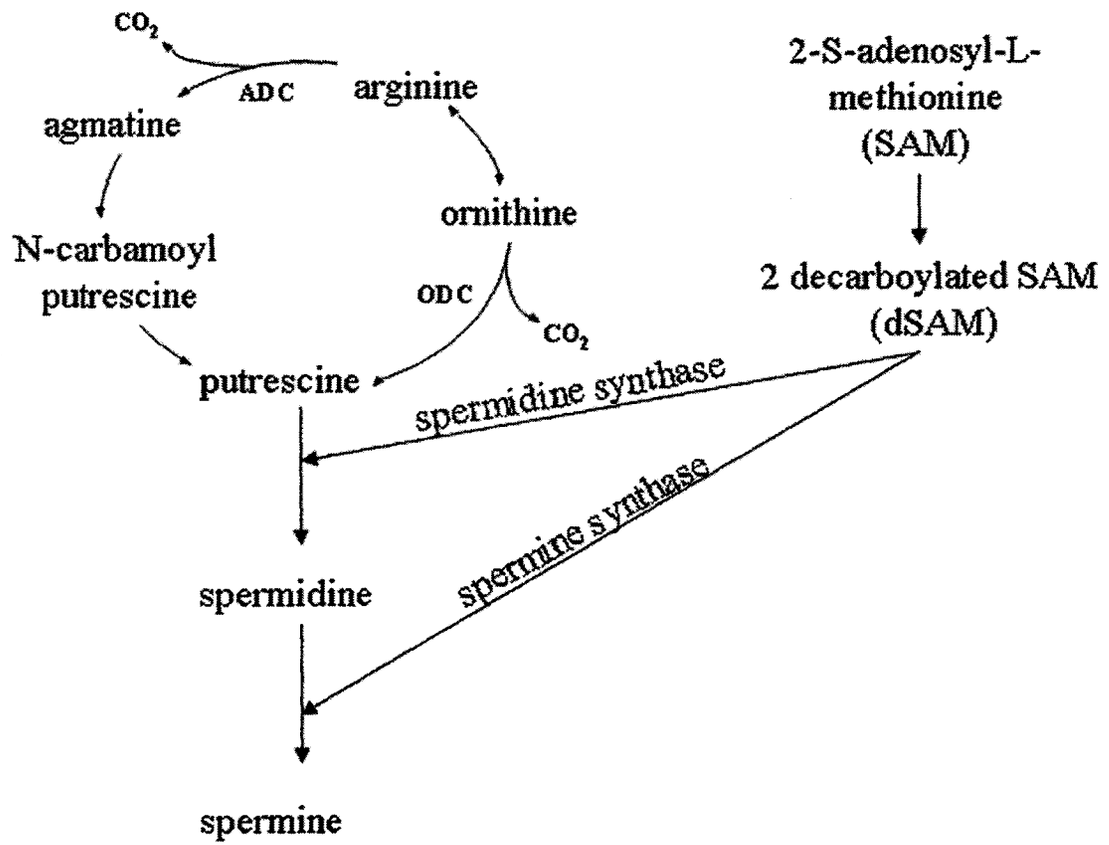


Figure 1.11. The synthesis of the polyamines putrescine, spermidine and spermine. Adapted from Crozier A, Kamiya Y, Bishop G, Yakota. In. Biochemistry & Molecular Biology of Plants. 2000. Buchanan B, Gruissem W, Jones R. American Society of Plant Biology. Rockville MD.

This synthesis of polyamines directly from Arg via arginine decarboxylase (*ADC*) is found only in plants and some bacteria (Hanfey et al. 2001). Polyamines can also be produced from ornithine via ODC. *ODC* genes have been identified in several species including humans and mice (Hanfey et al. 2001). An *ODC* cDNA was identified in the plant *Datura stramonium* (Michael et al. 1996); a partial *ODC* cDNA has also been identified in apple but its expression was reported to be quite low and a full length transcript could not be isolated (Hao et al. 2005). The occurrence of polyamine synthesis via *ODC* in Arabidopsis is a matter of some controversy. An *ODC* gene has not been identified in Arabidopsis, even though ODC enzyme activity was previously reported in leaves of this species (Feirer et al.1997). Hanfey et al. (2001) have suggested that *ODC* may not be present in Arabidopsis and discounted the presence of ODC reported by Ferier at al. (1997) because of the low level of ODC activity reported and because Hanfey et al. (2001) found no *ODC* activity in Arabidopsis cell cultures. More recently, ODC activity was again reported in leaves of Arabidopsis (Tassoni et al. 2003); however, as a putative *ODC* gene still has not been identified in Arabidopsis, the existence of *ODC* in *Arabidopsis* remains unclear. If Arabidopsis did lack *ODC* to our knowledge it would be the second of only two eukaryotes to lack *ODC*; the other being the protozoan *Trypanosoma cruzi* (Hanfey et al. 2001).

Polyamines have been implicated in a wide variety of plant processes: cell division, root formation and growth, bud formation and floral development, pollen development, fruit development, vascular development, and senescence (Bais and Ravishankar 2002). In addition, the activities of hormones that regulate plant growth

and development appear to be correlated with changes in polyamines (Bais and Ravishankar 2002). When stimulated by gibberellins, elongation of dwarf pea seedlings in pea is accompanied by an increase in polyamine abundance. Polyamines inhibit ethylene formation in a number of plants including apple, bean, and tobacco. This inhibition is of particular commercial interest, as polyamines show anti-senescence qualities. Polyamines, specifically putrescine, accumulate in response to various stresses in plants including potassium deficiency, acid stress, osmotic stress, ozone, and hydrogen fluoride; however, the exact role of polyamines in response to stress is not well understood (Bais and Ravishankar 2002). It has been shown that polyamines have a role in both the protection and attenuation of several types of plant stress including osmotic stress, heat and cold, oxidative, and nutrient related stresses (Perez-Amador et al. 2002, Zhao and Qin et al. 2004).

Polyamine production can also increase in response to methyl jasmonate (MeJA) treatment. Jasmonates act as natural regulators of a large number of processes in plant growth and development including the promotion of senescence and abscission, stomatal closure, inhibition of root growth and the inhibition of seed and pollen germination (Buchanan et al. 2000, Walters et al. 2002). In addition, jasmonates have been implicated in defense response signaling by inducing the expression of many defense related genes (Walters et al. 2002). Tobacco explants treated with MeJA accumulated polyamines through increased expression of *ADC* and *ODC* as well as increases in the specific activities of both enzymes (Biondi et al. 2001). The relationship between plant stress, MeJA and polyamines is unclear at present. Treatment of barley seedlings with MeJA systemically decreased the

infection of powdery mildew and increased the levels of several polyamines (Walters et al. 2002). Earlier experiments demonstrated that the pre-treatment of potato and tomato with MeJA causes a protective response when the plants are subsequently exposed to the pathogen *Phytophthora infestans* (Cohen et al. 1993). However, the polyamine content was not tested in this experiment. It is unclear whether the increase in polyamine content observed in barley had a role in conferring the systemic resistance to powdery mildew or if the MeJA treatment induced the systemic resistance and increased the production polyamines.

The exact role of polyamines in stress is unclear, as is their role in the other biological processes in which they appear to be involved. The increased accumulation of polyamines requires Arg or ornithine as a substrate for either ADC or ODC. Thus, as arginase produces ornithine through the hydrolysis of Arg, arginase may have an indirect role in these processes. Treatment of tomato plants with MeJA caused an increased arginase expression (Chen et al. 2004); however, as polyamine levels were not quantified we currently do not know if the increased arginase expression is related to increased polyamine synthesis.

1.7 The physiological roles of arginase

In plants, Arg degradation through arginase is often associated with post-germinative growth in both angiosperms and gymnosperms (Goldraij et al. 2000, King and Gifford 1997). Arg is a major source of stored nitrogen in seeds (Van Etten et al. 1963), and the hydrolysis of Arg via arginase allows nitrogen to be available for

seedling growth and development. Arginase activity has also been reported in tomato ovaries (Alabadí et al. 1996) and the arginase genes identified in tomato and Arabidopsis are expressed in non-seed related tissues (Chen et al. 2004, C Todd et al. unpublished data). In tomato, arginase is also induced in response to wounding, jasmonic acid, and the pathogen *Pseudomonas syringae* (Chen et al. 2004). This work collectively shows that arginase is involved in diverse physiological processes in plants.

1.7.1 Physiological roles of arginase in pine

In loblolly pine (*Pinus taeda* L.) the role of arginase in mobilizing seed storage reserves has been investigated. Several different tissues are present in a mature seed. The diploid embryo is enclosed by a maternally derived haploid megagametophyte (Figure 1.12). A nucellar cap is present outside of the megagametophyte at the micropylar end of the seed. These tissues are first surrounded by a thin papery layer and then by a hard outer seed coat (Stone and Gifford 1997). The megagametophyte is a maternally derived haploid tissue that holds the majority of both main types of storage reserves: triacylglycerols (TAGs) and storage proteins. TAGs comprise almost 60% of the total storage reserves in the mature seed, with 80% of those present in the megagametophyte (Stone and Gifford 1999). Over 90% of storage proteins, the other major reserve, are also located in the megagametophyte (Stone and Gifford 1997). These storage reserves are hydrolyzed

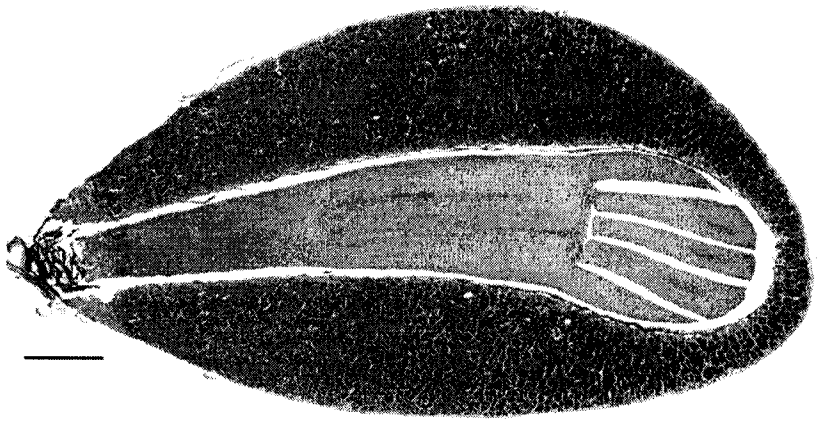


Figure 1.12: The loblolly pine seed. Adapted from Stone and Gifford 1997. A longitudinal section of the loblolly pine seed with the seed coat and outer integuments removed. The diploid dmbryo is enclosed in the maternally derived haploid megagametphyte. While there are no connective tissues between the megagametpyte in a hydrated mature seed the embryo is directly pressed against the megagamtophyte so they are in direct conact. The space between the two tissues observed above is a result of the processing.

after the completion of germination and supply the seedling with the nutrients it requires until it is autotrophic.

While the megagametophyte is the major source of TAGs in the seed, they are present in the embryo as well. TAGs, stored in the cell's lipid bodies, begin to be depleted after germination is complete, which is marked by the emergence of the radicle through the micropyle and the split in the seed coat (Tillman-Sutela and Kauppi 2000). During germination and early seedling growth TAGs are degraded and free fatty acids are transported to the glyoxosome. There, fatty acids are degraded in the β–oxidation pathway, followed by the glyoxalate cycle and gluconeogenesis to produce carbohydrates (Buchanan et al. 2000). The glycerol released from the fatty acids is converted to carbohydrates as well. Thus, the seedling is supplied with carbohydrates during early seedling growth.

Seed storage proteins are also degraded during early germination, and their products have several uses. Amino acids released can be incorporated into new proteins, or may be further metabolized. One of the major amino acid components of the storage proteins in loblolly pine is Arg (King and Gifford 1997). The majority of the seed storage proteins in the megagametophyte (75%) are insoluble in buffer (Groome et al. 1991) and almost 25% of the amino acids present in this fraction are Arg (King and Gifford 1997). This led the authors to conservatively estimate that 46% of total nitrogen present in the insoluble fraction was present as Arg (King and Gifford 1997).

Arginase activity is present in both the embryo and the megagametophyte of the mature seed; however, higher levels of arginase specific activity are localized to

the shoot pole (consisting of the cotyledons and the shoot apex, King (1998)) of the seedling, with maximum activity in $8DAI_{30}$ (days after imbibiton at 30°C) seedlings when the radicle is approximately 20-26mm in length (King and Gifford 1997). Arginase transcript and protein levels increase until slightly after the maximum specific activity and reach their maximum at $10DAI_{30}$ when the radicle is 41-47mm in length (Todd et al. 2001b). Arginase is most active in the shoot pole of the seedling where Arg accumulates (King and Gifford 1997). The majority of the Arg is likely exported from the megagametophyte to the seedling as Arg accumulates to a much higher level in the seedling when compared to the megagametophyte (King and Gifford 1997).

The megagametophyte appears to have a role in regulating arginase activity. Seedlings cultured with and without their megagametophytes have similar levels of arginase transcripts, protein, and enzyme activity during the first stages of development, but after 2 DIC_{30} (days in culture at 30°C) differences become apparent (King and Gifford 1997, Todd and Gifford 2002). Seedlings cultured without their megagametophytes have lower transcript abundance and arginase activity than their counterparts with a megagametophyte. Similarly, if the megagametophyte is removed from a seedling grown in culture, the arginase activity and transcript levels are lower when compared to seedlings that did not have their megagametophytes removed (King and Gifford 1997, Todd and Gifford 2002). The exogenous application of Arg to seedlings cultured without their megagametophytes can increase arginase transcript abundance, protein level and activity (Todd and Gifford 2003). However, the effect on transcript abundance is temporary; 24 hours after the

application of Arg the transcript level was lower when compared to transcript level 12 hours after spraying (Todd and Gifford 2003). Thus, arginase transcription appears to be very responsive to Arg; this allows the breakdown and incorporation of nutrients from storage proteins only during desirable times in seedling growth.

1.7.2 Physiological roles of arginase in tomato

Two arginases were identified from the tomato (*Lycopersicon esculentum*) EST database at The Institute for Genomic Research (TIGR) (Chen et al. 2004). The arginases were determined to be functional when the enzyme, which was expressed in *E. coli,* hydrolyzed L-Arg preferentially over other substrates tested. Southern analysis using gene specific probes from each gene's UTR showed that they are both single copy genes. The tomato arginases exhibited tissue specificity: *LeARG1* is expressed in roots, flower buds, mature unopened flowers, mature open flowers, and immature green fruit. *LeARG2* is expressed also expressed in buds, mature unopened flowers, mature open flowers, and immature green fruit but absent in roots. Neither *LeARG1* nor *LeARG2* is expressed at significant levels in stem or leaves of healthy tomato plants. Arginase activity was previously detected in unpollinated ovaries and developing fruit (Alabadí et al. 1996). Levels of arginase activity coincided with levels of polyamines present and Alabadí et al. (1996) proposed that in these tissues the role of arginase is to produce a substrate for polyamine synthesis.

Chen et al. (2004) also showed that *LeARG2* but not *LeARG1* was induced in leaves in response to wounding and jasmonic acid treatment. Jasmonates have been

implicated in defense response signaling by inducing the expression of many defense related genes (Walters et al. 2002). JA induces the expression of proteinase inhibitors involved in protecting the plant from insect damage and genes involved in synthesis of and antimicrobial secondary metabolites including alkaloids, terpenoids and flavonoids (Pauw and Memelink 2005). Pre-treatment of some plant species with jasmonates has been shown to have a protective effect when the plants are later exposed to pathogen attack (Cohen et al. 1993). In addition to the induction of *LeARG2* by wounding and jasmonic acid, it is also induced by the pathogen *P. syringae*, which causes bacterial speck disease (Chen et al. 2004).

1.7.3 Physiological roles of arginase in Arabidopsis

In comparison to loblolly pine, less work has been conducted on arginase in Arabidopsis. However, because the complete sequence of the Arabidopsis genome is available, much more is known about the structures of the arginase genes. Furthermore, resources such as T-DNA insertion lines are available for Arabidopsis, making Arabidopsis an excellent organism to use for the study of arginase in angiosperms.

T-DNA insertion lines are available for each of the arginases through the SALK (Alonso et al. 2003) and SAIL (Syngenta Arabidopsis Insertion Library) collections (Sessions et al. 2002). The SALK line (SALK_057987) has a T-DNA inserted into the second exon of *ARGAH1*, while the SAIL line (SAIL 181_C11) has a T-DNA inserted into the promoter of *ARGAH2*. Todd et al. (unpublished data) have

shown that no full length transcript is detectable in the homozygous T-DNA insertion lines. Using these T-DNA insertion lines Palmieri et al. (2006) have shown each arginase is differentially expressed and has different specific activity in germinating and developing seedlings. Analysis of arginase enzyme activity in either *argah*1 or *argah*2 T-DNA insertion lines indicates that early in seedling development, Argah 1 is active at low levels while Argah 2 is has much higher activity. These authors have also shown spatial differences in expression of *ARGAH1* and *ARGAH2* using semi-quantitative RT-PCR. According to this analysis, *ARGAH1* transcripts are expressed in seedlings, as well as in leaves, roots, stems, flowers and siliques. *ARGAH2* transcripts are expressed in many of the same tissues and stages including seedlings, leaves, flowers and siliques. However, *ARGAH2* transcripts are not detected in either the roots or stem.

The presence of arginase in these various tissues and stages of Arabidopsis shows that arginase is involved in several processes in Arabidopsis development. However, the specific roles of each arginase are not well understood.

1.7.4 Other physiological roles of arginase

In mammalian systems, pathogens and wounding up-regulate arginase expression (Duleu et al. 2004). The up-regulation of arginases expression in both plant and animal systems may be involved in increased polyamine production, which in animals is involved in tissue repair (Hesse et al. 2001). Tobacco explants treated with MeJA had increased expression of *ADC* and *ODC* and increased levels of

polyamines (Biondi et al. 2001). While arginase levels were not tested in this experiment, increased arginase activity would provide the explants with a source of ornithine as a substrate for ODC.

In addition to serving as a substrate for arginase, Arg can also be precursor of NO via NOS in animals and this may also occur in plants; thus arginase and NOS may compete for Arg as a substrate. NO is a signaling molecule and has been implicated in plant growth, fertility, stomatal movements, and hormone signaling and defense in plants (Guo et al. 2003). Several lines of evidence show that NO is involved in plant defense signaling. Tobacco plants resistant to the tobacco mosaic virus had increased NO production when exposed to the virus but susceptible tobacco plants did not (Durner et al. 1998). Exposure of either tobacco plants or cell suspensions to mammalian NOS triggers the expression of defense related genes and finally inhibitors of NOS in Arabidopsis compromised the hypersensitive response of plants exposed to *P. syringae* (Delledonne et al. 1998). In animals NO has roles in vasodilation, neurotransmission, and immune responses (Wu and Morris 1998). Inhibition of arginase resulted in increased NO production and increased parasite killing when mice macrophages were infected with trypanosomes (Duleu et al. 2004). Duleu et al. (2004) also demonstrated that trypanosomes directly induced both arginases possibly as a strategy to escape the NO response of the immune system. Thus, arginase may play several roles in defense: competing with NO production and in the production of polyamines.

The existence of a higher plant NOS gene has been a matter of controversy and at the present time there is no consensus about the role of a previously reported

AtNOS1 gene from Arabidopsis (Guo et al. 2003) or iNOS protein from tobacco (Chandok et al. 2003). NO can be produced in plants using other pathways, from nitrite via nitrate reductase (Yamasaki et al. 1999) but some researchers are still searching for Arg dependent NOS genes, which are present in mammalian systems. NOS activity was detected in several tissues and stages of pea seedlings during development (Corpas et al. 2006) and in Arabidopsis suspension cultures after exposure to the lipopolysaccarides (LPS) of several Gram negative bacteria (Zeidler et al. 2004). In addition, Guo and Crawford (2005) reported that NO production could be inhibited by NOS inhibitors in mitochondria from wild-type Arabidopsis plants; thus, NO production via NOS in plants may occur. AtNOS1 mutants showed increased susceptibility to the attack of the pathogen *Pseudomonas syringae* (Zeidler et al. 2004); however, the exact role of AtNOS1, now named At NOA1 (Crawford et al. 2006), is still unclear.

1.8 The present study

To our knowledge, no reports of plant arginase promoter isolation or characterization exist. In order to further our understanding of arginase and how it is regulated in both Arabidopsis and pine, we isolated the presumptive promoters (including all major cis-regulatory elements) from each species and produced a series of deletions for each of the presumptive promoters isolated. To indicate the patterns conferred by the different promoter fragments we utilized stable or transient expression of the promoter:reporter fusions. This analysis, in conjunction with the

identification of the cis-acting elements present in the upstream regions was used to identify regions that may be important in the regulation of arginase. In addition, we wanted to determine if the Arabidopsis arginase genes responded to wounding and MeJA treatment in a manner similar to those of tomato.

We also investigated the accumulation patterns of the major storage reserves of a pine somatic embryo line throughout maturation. Part of the economic relevance of understanding arginase function in pine is to improve the nutrition of tissues that do not have sufficient storage reserves, such as somatic embryos. Although storage reserves are present in somatic embryos, the levels may not be sufficient to allow the complete conversion of embryo to seedling to occur (Attree et al. 1992, Mhaske et al. 1998). In order to determine if the storage reserves present in this somatic line were quantitatively or qualitatively different from those of a zygotic embryo we quantified the major storage reserves and compared them to those found in mature zygotic embryos.

1.9 References:

Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD. 1994. Molecular Biology of the Cell. 4th edition. New York: Garland Publishing Inc.

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, Gadrinab C, Heller C, Jeske A, Koesema E, Meyers CC, Parker H, Prednis L, Ansari Y, Choy N, Deen H, Geralt M, Hazari N, Hom E, Karnes M, Mulholland C, Ndubaku R, Schmidt I, Guzman P, Aguilar-Henonin L, Schmid M, Weigel D, Carter DE, Marchand T, Risseeuw E, Brogden D, Zeko A, Crosby WL, Berry CC, Ecker JR. 2003. Genome Wide Insertional Mutagenesis of Arabidopsis thaliana. Science 310: 653-656.

Alabadí D, Agüero MS, Pérez-Amador MA, Carbonell J. 1996. Arginase, arginine decarboxylase, ornithine decarboxylase, and polyamines in tomato ovaries. Plant Physiology 112: 1237-1244.

Atkinson DE and Bourke E. 1984. The role of ureagenesis in pH homeostasis. Trends in Biochemical Sciences 9: 297-300.

Attree SM, Pomeroy MK, Fowke LC. 1992. Manipulation of conditions for the culture of somatic embryos of white spruce for improved triacylglycerol biosynthesis and desiccation tolerance. Planta 187: 395-404.

Ávila C, Suárez MF, Gómez-Maldonado J, Cánovas FM. 2001. Spatial and temporal expression of two cytosolic glutamine synthetase genes in Scots pine: functional implications on nitrogen metabolism during early stages of conifer development. The Plant Journal 25: 93-102.

Bais HP, Ravishankar GA. 2002. Role of polyamines in the ontogeny of plants and their biotechnological applications. Plant Cell, Tissue and Organ Culture 69: 1-34.

Biondi S, Scaramanli S, Capitani F, Altamur MM, Torrigiani P. 2001. Methyl jasmonate upregulates biosynthetic gene expression, oxidation and conjugation of polyamines, and inhibits shoot formation in tobacco thin layers. Journal of Experimental Botany 52: 231-242.

Boutin JP. 1982. Purification, properties, and subunit structure of the arginase from iris bulbs. European Journal Biochemistry 127: 237-243.

Boggess SF, Aspinall D, Paleg LG. 1976. Stress metabolism. IX. The significance of end-product inhibition of proline biosythesis and of compartmentation in relation to stress-induced proline accumulation. Australian Journal of Plant Physiology 3: 513-525.

Borkovich KA and Weiss RL. 1987. Purification and characterization of arginase from *Neurospora crassa*. The Journal of Biological Chemistry 262: 7081-7086.

Buchanan BB, Gruissem W, Jones RS. 2000. Biochemistry & Molecular Biology of Plants. American Society of Plant Biology. Rockville MD.

Cantón FR, Suárez FM, Cánovas FM. 2005. Molecular aspects of nitrogen mobilization and recycling in trees. Photosynthesis Research 83: 265-278.

Catoni E, Desimone M, Hilpert M, Wipf D, Kunze R, Schneider A, Flügge UI, Schumacher K, Frommer WB. 2003. Expression pattern of a nuclear encoded mitochondrial arginine-ornithine translocator gene from Arabidopsis. BMC Plant Biology 3: 1-10

Chandok MR, Ytterberg AJ, van Wijk KJ, Klessig. 2003. The pathogen-inducible nitric oxide synthase (iNOS) in plants is a variant of the P protein of the glycine decarboxylase complex. Cell 113: 469-482.

Chen H, McCaig BC, Melotto M, He SY, Howe GA. 2004. Regulation of plant arginase by wounding, jasmonate, and the phytotoxin coronatine. Journal of Biological Chemistry 279: 45998-46007.

Cohen Y, Gysi U, Niderman T. 1993. Local and systemic protection against *Phytopthora infestans* induced in potato and tomato plants by jasmonic acid and jasmonic methyl ester. Phytopathology 83: 1054-1062.

Corpas FJ, Barroso JB, Carreras A, Valderrama R, Palma JM, León AM, Sandalio LM, del Rio LA. 2006. Constitutive arginine-dependent nitric oxide synthase activity in different organs of pea seedlings during plant development. Planta 224: 246-254.

Crawford NM, Galli M, Tischner R, Heimer YM, Okamoto M, Mack A. 2006. Response to Zemojtel et al: Plant nitric oxide synthase: back to square one. Trends in Plant Science 11: 526-527.

Davis RH. Compartmental and regulatory mechanisms in the arginine pathway of *Neurospora crassa* and *Saccharomyces cerevisiae*. Microbiological Reviews 50: 280-313.

de la Torre F, García-Gutiérrez A, Crepillo R, Cantón FR, Ávila C, Cánovas FM. 2002. Functional expression of two pine glutamine synthetase genes in bacteria reveals that they encode cytosolic holoenzymes with different molecular and catalytic properties. Plant Cell Physiology 43: 802-809.

Delauney AJ, Hu CAA, Kishor PBK and Verma DPS. 1993. Cloning of ornithine δ–aminotransferase cDNA from *Vigna aconitifolia* by trans-complementation in

Escherichia coli and regulation of proline biosynthesis. Journal of Biological Chemistry 268: 18673-18678.

Delauney AJ, Verma DPS. 1993. Proline biosynthesis and osmoregulation in plants. The Plant Journal 4: 215-223.

Delledonne M, Xia Y, Dixon RA, Lamn C. 1998. Nitric oxide functions as a signal in plant disease resistance. Nature 394: 585-588.

Downum KR, Rosenthal GA, Cohen WS. 1983. L-arginine and L-canavanine metabolism in jack bean, *Canavalia ensiformis* (L.) DC. and soybean, *Glycine max* (L.) Merr. Plant Physiology 73: 965-968.

Durner J, Wendehenne D, Klessig DF. 1998. Defense gene induction in tobacco by nitric oxide, cyclic GMP and cyclic ADP-ribose. Proceedings of the National Academy of Sciences 95: 10328-10333.

Dubois E and Messenguy F. 1997. Integration of the multiple controls regulating the expression of the arginase gene CAR1 of *Saccharomyces cerevisiae* in response to different nitrogen signals: role of Gln3p, ArgRp-Mcm1p, and Ume6p. Molecular and General Genetics 253: 568-580.

Duleu S, Vincendeau P, Courtois P, Semballa S, Lagroye I, Daulouède S, Boucher JL, Wilson KT, Veyret B, Gobert AP. 2004. Mouse strain susceptibility to trypanosome infection: and arginase-dependent effect. The Journal of Immunology. 172: 6298-6303.

Feirer RP, Hocking KL, Woods PJ. 1997. Involvement of arginine decarboxylase in the response of *Arabidopsis thaliana* to osmotic stress. Journal of Plant Physiology 153: 733-738.

Gómez-Maldonado J, Ávila C, Barnestein P, Crespillo R, Cánovas FC. 2004. Interaction of cis-acting elements in the expression of a gene encoding cytosolic glutamine synthetase in pine seedlings. Physiologia Plantarum 121: 537-545.

Groome MC, Axler SR and Gifford DJ. 1991. Hydrolysis of lipid and protein reserves in loblolly pine seeds in relation to protein electrophoretic patterns following imbibition. Physiologia Plantarum 83: 99-106.

Goldraij A, and Pollaco JC. 1999. Arginase is inoperative in developing soybean embryos. Plant Physiology 119: 297-303.

Goldraij A, and Pollaco JC. 2000. Arginine degredation by arginase in mitochondria of soybean seedling cotyledons. Planta 210: 652-658.

Guo FQ, Okamoto M, Crawford NM. 2003. Identification of a plant nitric oxide synthase gene involved in hormonal signaling. Science 302: 100-103.

Guo FQ and Crawford NM. 2005. *Arabidopsis* nitric oxide synthase 1 is targeted to mitochondria and protects against oxidative damage and dark-induced senescence. The Plant Cell 17: 3436-3450.

Gupta PK and Durzan DJ. 1987. Biotechnology of somatic polyembryogenesis and plantlet regeneration in loblolly pine. Bio/technology 5: 147-151.

Hanfrey C, Sommer S, Mayer MJ, Burtin D, Michael AJ. 2001. *Arabidopsis* polyamine biosynthesis: absence of ornithine decarboxylase and the mechanism of arginine decarboxylase activity. Plant Journal 27: 551-560.

Hao YJ, Kitashiba H, Honda C, Nada K, Moriguchi T. 2005. Expression of arginine decarboxylase and ornithine decarboxylase genes in apple cells and stressed shoots. Journal of Experimental Botany 56: 1105-1115.

Heldt HW and Heldt F. 1998. Plant Biochemistry and Molecular Biology. Oxford University Press.

Hesse M, Modolell M, LaFlamme AC, Schito M, Fuentes JM, Cheever AW, Pearce EJ, Wynn TA. 2001. Differential regulation of nitric oxide synthase-2 and arginase-1 by type 1/type2 cytokines in vivo: gramulomatous pathology is shaped by the pattern of L-arginine metabolism. Journal of Immunology 167: 6533-6544.

Hoyos ME, Palmieri L, Wertin T, Arrigoni R, Polacco JC, Palmieri F. 2003. Identification of a mitochondrial transported for basic amino acids in *Arabidopsis thaliana* by functional reconstitution into liposomes and complementation in yeast. The Plant Journal 33: 1027-1035.

Jenkinson CP, Grody WW, Cederbaum SD. 1996. Comparative Properties of Arginases. Comparative Biochemistry and Physiology 114B: 107-132.

Kang JH and Cho YD. 1990. Purification and properties of arginase from soybean, *Glycine max*, axes. Plant Physiology 93:1230-1234.

King JE. 1998. PhD thesis. Arginine and Arginase in *Pinus taeda* seedlings. University of Alberta, Edmonton Alberta.

King JE and Gifford DJ. 1997. Amino acid utilization in seeds of loblolly pine during germination and early seedling growth. I. Arginine and arginase activity. Plant Physiology 113: 1125-1135.

Kollöffel C and van Duke HD. 1975. Mitochondrial arginase activity from cotyledons of developing and germinating seeds of *Vicia faba* L. Plant Physiology 55: 507-510.

Krumpelman PM, Freyermuth SK, Cannon JF, Fink GR, Polacco JC. 1995. Nucleotide Sequence of *Arabidopsis thaliana* Arginase Expressed in Yeast. Plant Physiology 107: 1479-1480.

Lam HM, Koshigano KT, Oliveira IC, Melo-Oliveira R, Coruzzi GM. 1996. The molecular-genetics of nitrogen assimilation into amino acids in higher plants Annual Review of Plant Physiology and Plant Molecular Biology 47: 569-593.

Legaz ME. 1991. Regulation of different arginase forms in *Evernia prunastri* thallus. Symbiosis 11: 263-277.

Legaz ME and Vincente C. 1980. Arginase regulation in *Evernia prunastri* (L.) ach. Cryptogamie, Bryologie et Lichenologie 1: 407-414.

Legaz ME and Vincente C. 1982. Two forms of arginase in *Evernia prunastri* thallus. Biochemical and Biophysical Research Communications 104: 1441-1446.

Marathe S, Yu YG, Turner GE, Palmier C, Weiss RL. Multiple forms of arginase are differentially expressed from a single locus in *Neurospora crassa*. The Journal of Biological Chemistry 273: 29776-29785.

Martin-Falquina A and Legaz MA. 1984. Purification and properties of the constitutive arginase of *Evernia prunastri*. Plant Physiology 76: 1065-1069.

Martin-Tanguy J. 2001. Metabolism and function of polyamines in plants: recent development (new approaches). Plant Growth Regulation 34: 135-148.

Mhaske VB, Chengalrayan K, Hazra S. 1998. Influence of osmotica and abscisic acid on triglyceride accumulation in peanut somatic embryos. Plant Cell Reports 17:742-746.

Micallef BJ and Shelp BJ. 1989. Arginine metabolism in developing soybean cotyledons. I. Relationship to nitrogen nutrition. Plant Physiology 90: 624-630.

Michael AJ, Furze JM, Rhodes MJ, Burtin D. 1996. Molecular cloning and functional identification of a plant ornithine decarboxylase cDNA. Biochemistry Journal 314: 241-248.

Morris SM. 2002. Regulation of enzymes of the urea cycle and arginine metabolism. Annual Review of Nutrition 22: 87-105.

Noguchi M, Koiwai A, Yokoyama M, Tamaki E. 1968. Studies on nitrogen metabolism in tobacco plants IX. Effect of various compounds on proline biosynthesis in the green leaves. Plant and Cell Physiology 9: 35-47.

Palmieri L, Todd CD, Arrigoni R, Hoyos ME, Santoro A, Polacco JC, Palmieri F. 2006. Arabidopsis mitochondria have two basic amino acid transporters with partially overlapping specificities and differential expression in seedling development. Biochimica et Biophysica Acta 1757: 1277-1283.

Pauw B and Memelink J. 2005. Jasmonate-responsive gene expression. Journal of Plant Growth and Regulation 23: 200-210.

Perez-Amador MA, Leon J, Green PJ, Carbonell J. 2002. Induction of the arginine decarboxylase *ADC*2 gene provides evidence for the involvement of polyamines in the wound response in Arabidopsis. Plant Physiology 130:1454-1463.

Perozich J, Hempel J, Morris SM Jr. 1998. Roles of conserved residues in the arginase family. Biochimica et Biophysica Acta 1382: 23-37.

Planelles V and Legaz ME. 1987. Purification and some properties of the secreted arginase of the lichen *Evernia prunastri* and its regulation by usnic acid. Plant Science 51: 9-16.

Pullman GS, Montello P, Cairney J, Xu N, Feng X. 2003. Loblolly pine (*Pinus taeda* L.) somatic embryogenesis: maturation improvements by metal analysis of zygotic and somatic embryos. Plant Science 164: 955-969.

Rhodes D, Handa S, Bressan RA. 1986. Metabolic changes associated with adaptation of plant cells to water stress. Plant Physiology 82: 890-903.

Roosens NHCJ, Thu TT, Iskandar HM, Jacobs M. 1998. Isolation of the ornithine-δ-aminotransferase cDNA and effect of salt stress on its expression in *Arabidopsis thaliana.* Plant Physiology 117: 263-271.

Roosens NH, Bitar FA, Loenders K, Angenon G, Jacobs M. 2002. Overexpression of ornithine- δ -aminotransferase increases proline biosynthesis and confers osmotolerance in transgenic plants. Molecular Breeding 9: 73-80.

Sessions A, Burkea E, Prestinga G, Auxb G, McElverb J, Patton D, Dietrich B, Hoa P, Bacwadena J, Koa C, Clarkea JD, Cottona D, Bullisa D, Snella J, Miguela T, Hutchisona D, Kimmerly B, Mitzelc T, Katagiria F, Glazebrooka J, Lawb M, Goffa SA. 2002. A High-Throughput Arabidopsis Reverse Genetics System. The Plant Cell 14: 2985-2994.

Shi O, Kepka-Lenhart D, Morris SM, O'Brien WE. 1998. Structure of the murine arginase II gene. Mammalian Genome 9: 822-824.

Sirko A and Brodzik R. 2000. Plant ureases: Roles and regulation. Acta Biochimica Polonica 47: 1189-1195.

Splittoesser WE. 1969. The appearance of arginine and arginase in pumpkin cotyledons. Characterization of arginase. Phytochemistry 8: 753-758.

Stone SL and Gifford DJ. 1997. Structural and biochemical changes in loblolly pine (*Pinus taeda* L.) seeds during germination and early-seedling growth. I. Storage protein reserves. International Journal of Plant Science 158: 727-737.

Stone SL and Gifford DJ. 1999. Structural and biochemical changes in loblolly pine (*Pinus taeda* L.) seeds during germination and early-seedling growth. II. Storage triacylglycerols and carbohydrates. International Journal of Plant Science 160: 663-671.

Takiguchi M, Haraguchi Y, Mori M. Human liver-type arginase gene: structure of the gene and analysis of the promoter region. Nucleic Acids Research 16: 8789-8802.

Tassoni A, Fornale S, Bagni N. 2003. Putative ornithine decarboxylase activity in Arabidopsis thaliana: inhibition and intracellular localization. Plant Physiology and Biochemistry 41: 871-875.

Taylor AA and Stewart GR. 1981. Tissue and subcellular localization of enzymes of arginine metabolism in *Pisum sativum*. Biochemical and Biophysical Research Communications 101:1281-1289.

Tillman-Sutela E and Kauppi A. 2000. Structures contributing to the completion of conifer seed germination. Trees 14: 191-197.

Tobin AK, Tamaya T. 2001. Cellular compartmentation of ammonium assimilation in rice and barley. Journal of Experimental Botany 52: 591-604.

Todd CD, Cooke JEK, Gifford DJ. 2001a.Purification and properties of *Pinus taeda* arginase from germinated seedlings. Plant Physiology and Biochemistry 39: 1037-1045.

Todd CD, Cooke JEK, Mullen RT, Gifford DJ. 2001b. Regulation of loblolly pine (*Pinus taeda L.*) arginase in developing seedling tissue during germination and post-germinative growth. Plant Molecular Biology 45: 555-565.

Todd CD and Gifford DJ. 2002. The role of the megagametophyte in maintaining loblolly pine (*Pinus taeda* L.) seedling arginase gene expression in vitro. Planta 215: 110-118.

Todd CD and Gifford DJ. 2003. Loblolly pine arginase responds to arginine in vitro. Planta 217: 610-615.

Turner GE and Weiss RL. Developmental expression of two forms of arginase in *Neurospora crassa*. Biochimica et Biophysica Acta 1760: 848-857.

Van Etten CH, Miller RW, Wolff IA. 1963. Amino acid composition of seeds from 200 angiospermous plant species. Agricultural and Food Chemistry 11:399-410.

Walters D, Cowley T, Mitchell A. 2002. Methyl jasmonate alters polyamine metabolism and induces systemic protection against powdery mildew infection in barley seedlings. Journal of Experimental Botany 53: 747-756.

Wright LC, Brady CJ, Hinde RW. 1981. Purification and properties of the arginase from Jerusalem artichoke tubers. Phytochemistry 20: 2641-2645.

Wu G and Morris SM. 1998. Arginine metabolism: nitric oxide and beyond. Biochemstry Journal 336: 1-17.

Yamasaki H, Sakihama Y, Takahashi S. 1999. An alternative pathway for nitric oxide production in plants: new features of an old enzyme. Trends in Plant Sscience 4: 128-129.

Zeidler D, Zähringer U, Gerber I, Dubery I, Hartung T, Bors W, Hutzler P, Durner J. 2004. Innate immunity in *Arabidopsis thaliana*: lipopolysaccharides activate nitric oxide synthase (NOS) and induce defense genes. Proceedings of the National Academy of Science 101: 15811-15816.

Zhao FG and Qin P. 2004. Protective effect of exogenous polyamines on root tonoplast function against salt stress in barley seedlings. Plant Growth and Regulation 42: 97-103.

Zonia LE, Stebbins NM, Polacco JC. 1995. Essential role of urease in germination of nitrogen-limited *Arabidopsis thaliana* seeds. Plant Physiology 107: 1097-1103.

Chapter 2: Analysis of loblolly pine somatic embryos

2.1 Introduction

Loblolly pine (*Pinus taeda* L.) has many characteristics that make it attractive as a commercial species including its ability to tolerate a range of soils and environments; consequently loblolly pine has been planted in a number of countries including the United States, Argentina, Australia, Brazil, China, New Zealand, and South Africa (Harms et al. 2000). In the southeastern United States, loblolly pine is a major timber species, being planted on greater than 45% of the commercial forest land (Schultz 1999). However, as demand for paper products is increasing (Pullman et al. 2003) there is increasing pressure to make existing forests more productive and one of the proposed methods to achieve this is through somatic embryogenesis.

Somatic embryogenesis is a multi-stage process (Figure 2.1) in which cellular material is first propagated in culture and then induced to form embryo-like structures that have the potential to grow into functional plants. In the first stage, plant material is isolated and placed in a liquid culture system to proliferate. In flowering plants, embryogenic cultures are often initiated using mature plant sections (Stasolla and Yeung 2003). However, in conifers, the initial starting material is usually immature zygotic embryos. This requirement for young tissue is specific to conifers; the embryogenic tissue grows and divides until it is the appropriate size where it is transferred to a solid medium where the somatic embryos undergo maturation. During this time the embryos

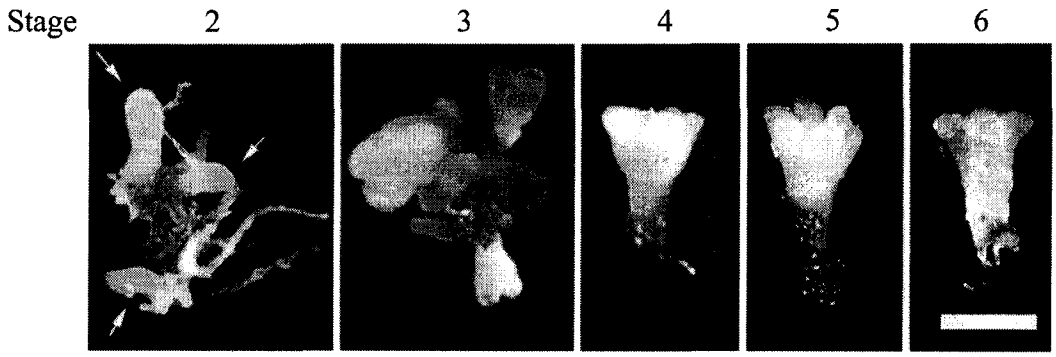


Figure 2.1: Stages 2-6 of somatic embryos maturation in loblolly pine. Adapted from Brownfield et al. 2007. The arrows in Stage 2 indicate developing embryo. Scale bar 1mm.

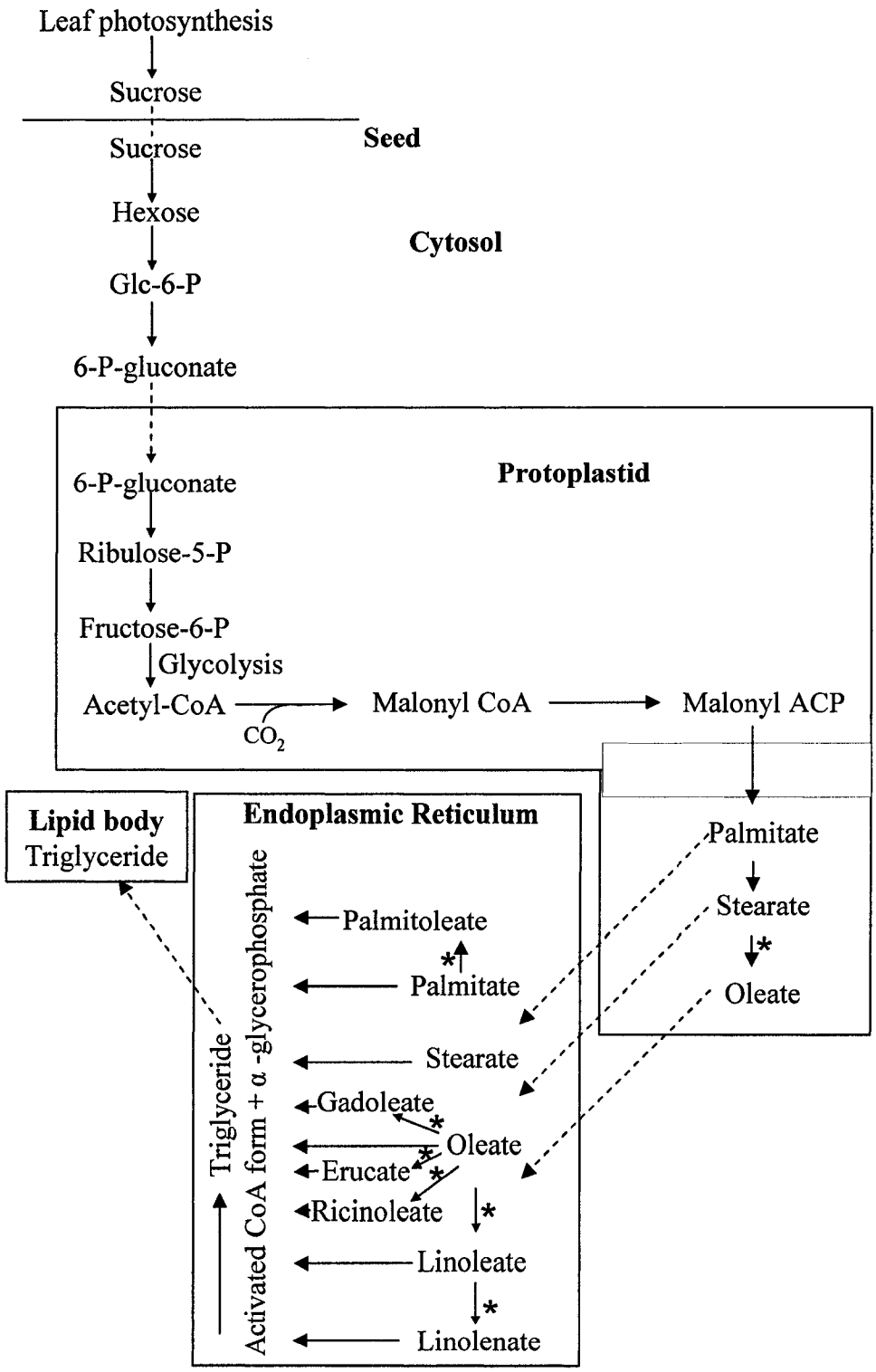


Figure 2.2: An overview of fatty acid synthesis from sucrose as a starting molecule. Adapted from Bewley and Black 1985. The dotted arrows represent a translocation and the asterisk (*) represents a desaturation.

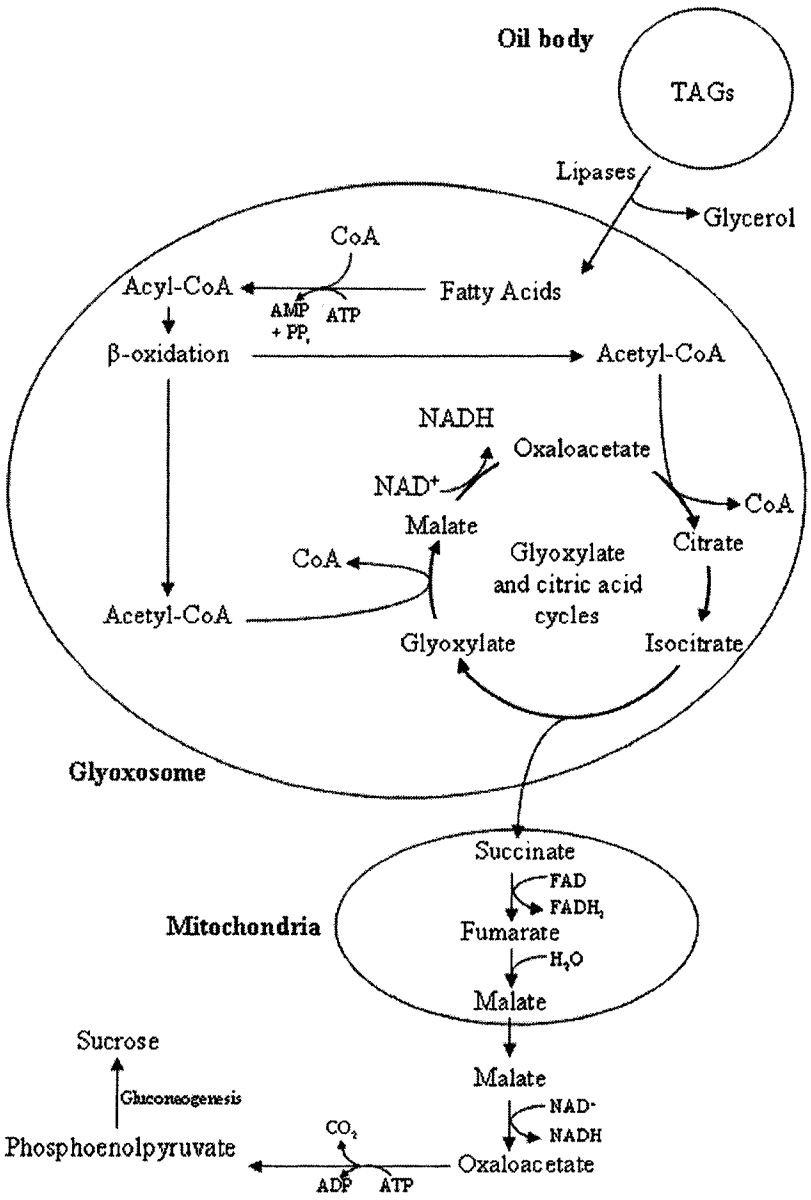


Figure 2.3: An overview of triacylglycerol (TAG) mobilization. Adapted from Sommerville C, Browse J, Jaworski JG, Ohlrogge JB. In: Biochemistry & Molecular Biology of Plants. 2000. Buchanan B, Gruissem W, Jones R. American Society of Plant Biology. Rockville MD. After germination storage TAGs (along with proteins) are degraded to supply the seedling with nutrients until it is autotrophic.

not only grow but also accumulate storage reserves (Figure 2.2) which are used to sustain the embryo's germination and growth.

Production of somatic embryos has been achieved for several conifer species including white spruce (Attree et al. 1995), black spruce (Beardmore and Charest 1995), Douglas fir (Taber et al. 1998) and loblolly pine (Becwar et al. 1990). However, some species are recalcitrant to *in vitro* conditions (Stasolla et al. 2002) and even if germination does occur for these species, subsequent growth may not be sufficiently vigorous.

Conifer somatic embryos are fundamentally different from their zygotic counterparts. Somatic embryos develop in culture without the tissues that normally surround the zygotic embryo in the seed, most importantly the megagametophyte. The absence of a megagametophyte may be problematic given that the majority of the storage reserves are located in this tissue. In loblolly pine, approximately 80% of the total seed reserves reside in the megagametophyte (Stone and Gifford 1997, Stone and Gifford 1999). The storage reserves present in zygotic seedlings are of two main types: storage proteins and triacylglycerols (TAGs). These reserves are broken down during germination and provide the developing seedling with all of the nutrients it requires until it becomes autotrophic (King and Gifford 1997, Stone and Gifford 1999). The TAGs comprise 59% of the total storage reserves in the mature loblolly pine seed with 80% of that being stored in the megagametophyte (Stone and Gifford 1999). These TAGs, once mobilized (Figure 2.3), are believed to be converted into carbohydrates (sucrose) that are exported to the cotyledons of the seedling (Stone and

Gifford 1999). Similarly, the storage proteins are degraded to amino acids by proteases and transported to the seedling. Although TAGs and storage proteins are present in a somatic embryo, they may not be sufficient to allow the conversion to occur, thus it has been suggested that an increase in somatic embryo storage reserves might result in improved conversion (Attree et al. 1992, Mhaske et al. 1998).

There are four classes of seed storage proteins, which have been traditionally classified by their solubility (Bewley and Black 1985). Albumins are soluble in water and dilute buffers at neutral pH. Globulins are insoluble in water but are soluble in salt solutions. Glutelins are soluble in most dilute alkali or acid solutions. Finally, prolamins are soluble in aqueous alcohols. In order to maintain consistency we follow the nomenclature of Groome et al. (1991) and Stone and Gifford (1997) and refer to the extracted proteins as two separate pools, the buffer-soluble and buffer-insoluble, relative to the 50mM sodium phosphate buffer (pH7.5) used in the extraction.

During the production of somatic embryos, specific hormones are added or removed from culture medium to allow the embryos to progress through development. In the initiation of embryogenic cultures, low concentrations of auxins and cytokinins are generally used (Stasolla and Yeung 2003). The removal of these auxins and cytokinins along with the addition of ABA (abscisic acid) allows the generation of embryos from the established culture. Specifically, this medium arrests cell proliferation and induces the cells to develop into somatic embryos (Stasolla and Yeung 2003).

In angiosperms, ABA has been associated with inhibiting early germination, promoting the synthesis of storage reserves, acquiring desiccation tolerance,

preventing reserve mobilization, and inducing dormancy (Dunstan et al. 1998). In alfalfa, if immature embryos are harvested and cultured they germinate early without proceeding through the later stages of embryo development (Xu and Bewley 1995). Thus, the conditions *in planta* prevent early germination and allow proper development of the embryo. Xu and Bewley (1991) also showed that alfalfa embryo germination can be prevented *in vitro* by the application of either ABA or an osmoticum. Osmotica such as polyethylene glycol (PEG) induce a slight desiccation of the embryo, which is an important stage of embryo development both *in vivo* and *in vitro* (Stasolla and Yeung 2003). In conifers, using ABA in combination with PEG has been shown to increase somatic embryo storage reserve deposition as well as produce embryos with low final moisture content (Attree et al. 1992).

Desiccation is important in the transition between seed development and seed germination (Xu and Bewley 1994). In alfalfa, desiccation stops storage protein synthesis, one of the main processes of seed maturation. When germinated alfalfa embryos that did not undergo desiccation were transferred to media containing ABA or osmotica, the embryos were capable of maintaining storage protein synthesis even though they should be breaking down these storage reserves for seedling growth and development. However, if germinated somatic embryos that had been desiccated were transferred to the same conditions, storage protein synthesis was not resumed (Xu and Bewley 1994). This indicates that desiccation causes the embryos to become insensitive to ABA and osmoticum thus switching the developmental program from seed maturation to that of germination. In conifers, the inclusion of a drying period has been shown to increase germination efficiency and the conversion of somatic

embryos to plantlets for many genera (Stasolla et al. 2002). Furthermore, drying may be an efficient method for the long-term storage of somatic embryos (Bomal and Tremblay 1999).

Although loblolly pine somatic embryo production was first reported in 1987 (Gupta and Durzan 1987), germination and seedling conversion rates for these somatic embryos are still relatively low (Pullman et al. 2003b). Consequently, much effort has been spent on manipulating conditions for the production and maturation of somatic embryos and plantlet regeneration. In contrast, comparatively little work has been published on the basic biology of the somatic embryos. Here we present storage protein and triacylglycerol (TAG) accumulation patterns throughout somatic embryo maturation for a single embryogenic line. We have quantified these major storage reserves and compared them to those found in mature zygotic embryos; somatic embryos utilized in this study contained significantly lower levels of TAGs and higher levels of storage proteins. A shift in the ratio of soluble to insoluble protein present was also observed.

2.2 Materials and methods

2.2.1 Loblolly pine seeds

Somatic embryos were provided as a gift from the Pacific Biotech Institute (Vancouver, BC). Somatic embryos were prepared from embryonic tissue in liquid culture and progressed through development and maturation over 13 weeks. During this time the developing embryos were transferred from liquid culture (Stage 1, Table 2.1) into media with increasing concentrations of ABA and PEG. ABA was gradually increased from 100μM to 250μM, while PEG increased from 5% to 16%. The basal media consisted of 925 mg/L $MgSO_4$, 211 mg/L $CaCl_2$-$2H_2O$, 950 mg/L KNO_3, 170 mg/L KH_2PO_4, 4.15 mg/L $ZnSO_4$-$7H_2O$, 31 mg/L H_3BO_3, 21 mg/L $MnSO_4$-H_2O, 4.15 mg/L KI, 1.3mg/L Na_2MoO_4-H2O, 0.5mg/L $CuSO_4$-$5H_2O$, 0.13mg/L $CoCl_2$-$6H_2O$, 27.8 mg/L $FeSO_4$-$7H_2O$, 37.2mg/L Na_2EDTA, 100 mg/L myo-inositol, 0.5 mg/L Nicotinic acid, 0.1 mg/L Pyridoxine-HCl, 0.1 mg/L Thiamine-HCl, 2.5% (w/v) Maltose, 3000 mg/L Casamino acids, 2500 mg/L L-Glutamine, 0.6% (w/v) Phytagel. Thiamine-HCl, 2.5% (w/v) Maltose, and was modified by the addition of ABA and PEG to levels indicated in Table 2.1. Media pH was adjusted to 5.7. Casamino acids, glutamine and ABA were filter sterilized prior to their addition to the cooled media. The length of each maturation stage is also indicated in Table 2.1. Cultures were maintained in the dark at 22°C. After 13 weeks of maturation the somatic embryos were allowed to undergo desiccation by placing them on filter paper and allowing them to air dry slowly until were sufficiently decreased in moisture content.

Table 2.1: Somatic embryo stages and maturation conditions. Adapted from Brownfield et al. 2007. Stage 1 is liquid culture.

Stage	Weeks	[PEG] (%)	[ABA] (μM)
1	0	-	-
2	1-3	5	100
3	4-6	10	150
4	7-9	13	200
5	10-12	16	250
6	13	Air dried	

Somatic embryos were harvested every week, frozen in liquid nitrogen and stored at -80°C for analysis.

2.2.2 TAG quantification

Triacylglycerols were extracted from 100-500mg (fresh weight) of somatic embryos using the method of Stone and Gifford (1999). Samples were extracted in 1.5mL of isopropanol in an Eppendorf tube and mixed gently for 15 minutes before being centrifuged at 14,000 rpm for 15 minutes at room temperature. The supernatant was removed with a pulled pipet; the pellet was re-extracted with 1.25mL of isopropanol and again nutated and centrifuged. The supernatants were then pooled and labeled as pooled crude supernatants. An 800μL aliquot of the pooled crude supernatant was purified using 0.8g of alumina. After addition of the alumina the tubes were mixed gently and then centrifuged at 3000rpm for 10 minutes before removing the supernatant using a pulled pipet. The alumina pellet was washed twice with 2.6mL of isopropanol each time followed by nutation for 15 minutes, centrifugation at 3000rpm for 10 minutes and supernatant removal with a pulled pipet. All supernatants from alumina purification were pooled and labeled as pooled purified supernatant. The pooled purified supernatant was re-centrifuged at 3000rpm for 5 minutes, the supernatant collected with a pulled pipet and used in the TAG quantification assay.

TAG quantification was performed in triplicate using the method of Feirer et al. (1989) using triolein (Sigma) as a standard. A standard curve was prepared by

adding 0, 96, 320, 440 and 880μg of triolein to approximately 6.5mL of isopropanol (a volume similar to that of the pooled purified supernatants) and assayed as below. Experimental samples were diluted appropriately to ensure the assayed sample A_{410} values were within the range of the triolein standard curve. An 800μL aliquot of the diluted purified sample was added to an acid washed pyrex test tube before the addition of 200μL of 1N KOH. The sample was vortexed and the test tube opening covered with a marble to prevent water from entering the tube. The test tubes were incubated in a 60°C water bath for 5 minutes and then cooled in a basin of room temperature water for 2 minutes and the marbles removed. The solution was incubated with 200μL of sodium-m-periodate solution (12mM) for 10 minutes after which 1.2mL of 4°C colour reagent (0.6M NH_4OAc, 12.5μM acetylacetone in isopropanol) was added to each tube and the tubes were covered in new marbles. The tubes again were placed in a 60°C water bath for 30 minutes before being cooled in a basin of room temperature water for 2 minutes and the marbles removed. The absorbance of the solution at 410nm was recorded and compared to that of the standard curve.

2.2.3 Protein isolation

Proteins were isolated from 100-500mg (fresh weight) of somatic embryos using the method of Stone and Gifford (1997). Samples were extracted in 1mL of 0.05M sodium phosphate buffer (pH 7.5) with PVPP (20mg per sample) and centrifuged at 14,000rpm at 4°C. The supernatant was collected using a pulled pipet

the pellet was re-extracted three times each with 1ml of phosphate buffer and the supernatants pooled to generate the soluble protein fraction. The remaining pellet was then extracted with 1mL of Laemmli buffer and boiled for 5 minutes. After centrifuging (14,000rpm, 20 minutes, room temperature) the supernatant was isolated and the pellet re-extracted three times (1mL each) as above. The supernatants were pooled and together are the phosphate buffer insoluble protein fraction. Protein concentration was determined using the method of Lowry et al. (1951) using bovine serum albumin as the standard. Protein was reduced using β–mercaptoethanol and 8μg of protein was applied to 12% SDS-PAGE gels and electrophoresis was conducted as described in Stone and Gifford (1997).

2.2.4 Tree Construction

Amino acid sequences used in tree construction were identified from the literature (Shutov et al. 2003, Tai et al. 1999, Hager and Wind 1997, Wind and Hager 1996, Hager et al. 1995, Shutov et al. 1998, Klimaszewska et al. 2004, Newton et al. 1992, Leal and Misra 1993; Chattai and Misra 1998, Krebbers et al 1988, Baszczynski and Fallis 1990, Tai et al. 2001, Gander et al 1991, Kortt et al. 1991), In addition we also utilized consensus EST sequences identified in the Gene Index (www.compbio.dfci.harvard.edu). The open reading frame (ORF) of each consensus EST was predicted in the Gene Index and an amino acid sequence was generated. Amino acid sequences were aligned using Clustal W (Higgins et al. 1994) and the

neighbour joining trees were constructed using the Mega version 4 software (Tamura et al. 2007).

2.3 Results and discussion

To characterize the deposition of TAGs in developing somatic embryos, we quantified the TAG levels at one week intervals, which we associated with the developmental stages shown in Figure 2.1. We observed an overall weekly increase in TAG levels throughout development after early maturation phase (Figure 2.4A), with the largest increase in TAG accumulation occurring in the transition between early and mid-maturation (stages 3-4). After desiccation, the final TAG content of the somatic embryos was approximately 60μg per mg dry weight (Figure 2.4B). Compared to zygotic embryos, the somatic embryos contained only 24% of the TAGs on a dry weight basis. Due to the potential differences in the hydration state of mature somatic and zygotic embryos, comparisons between the two embryo types were based on dry weight. Somatic embryos containing significantly more TAGs have been achieved using similar maturation protocols in other species; white spruce (*Picea glauca*) somatic embryos containing five to nine fold more TAGs than their zygotic counterparts have been produced (Attree et al.1992). However, using this maturation protocol, loblolly pine somatic embryos did not achieve this dramatic increase in TAG levels. A comparison of the developmental changes in the fatty acid composition of the somatic embryos would be useful to illustrate differences between somatic and zygotic embryos and to determine if there are any specific fatty acids that are missing from the somatic embryo (Janick et al. 1991, Attree et al. 1992).

We isolated total protein from maturing somatic embryos and we observed an

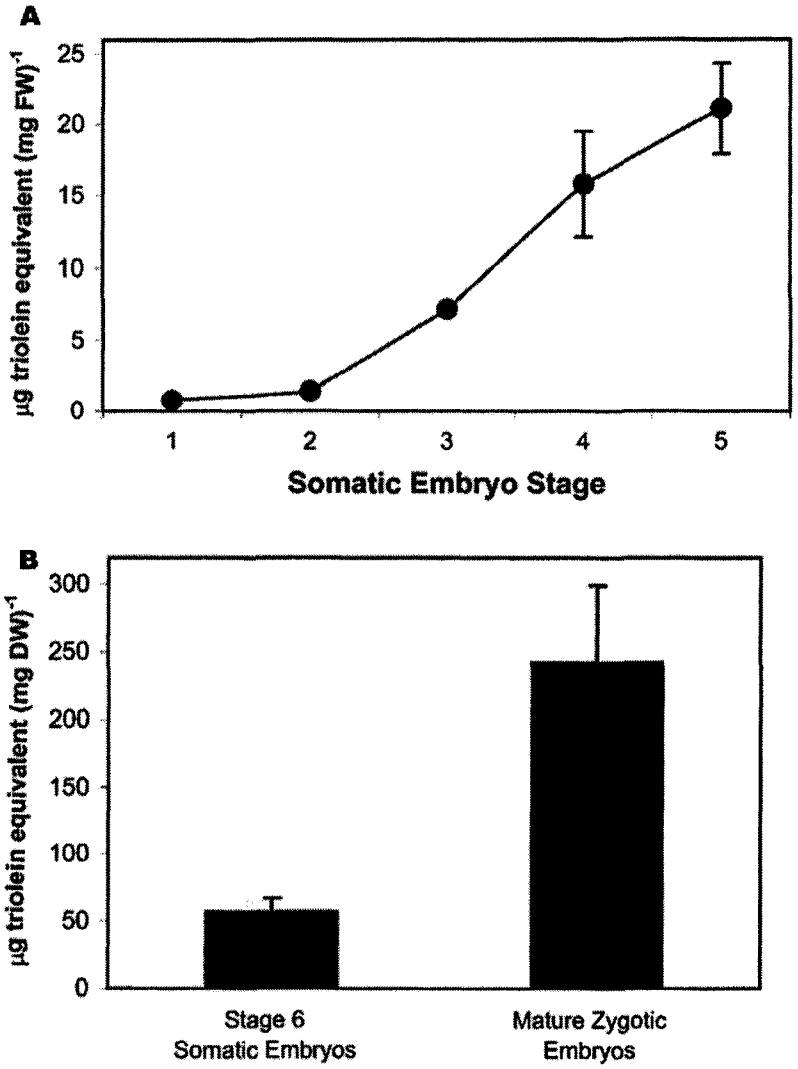


Figure 2.4: Triacylglycerols in loblolly pine somatic embryos. Adapted from Brownfield et al 2007. Error bars represent the standard error of at least three independent measurements. A. Triacylglycerol accumlation during somatic embryo maturation. B. Triacylglycerol content of mature, desiccated somatic embryos and mature zygotic embryos on a dry weight basis.

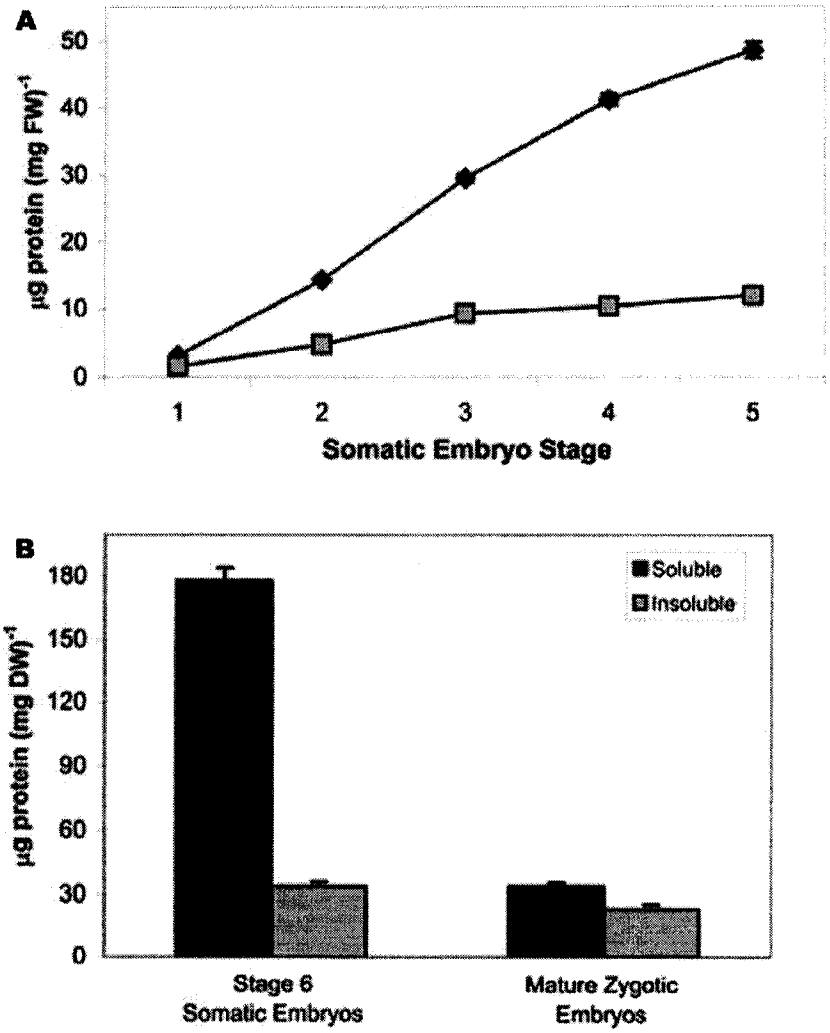


Figure 2.5: Proteins in loblolly pine somatic embryos. Adapted from Brownfield et al 2007. Error bars represent the standard error of at least three independent measurements. A. Buffer soluble (black) and insoluble (grey) protein accumulation during somatic embryo maturation. B. Protein content of mature, desiccated somatic embryos and mature zygotic embryos on a dry weight basis.

increase in total protein throughout development (Figure 2.5). We refer here to two separate pools of proteins: buffer-soluble and buffer insoluble as per the nomenclature of Groome et al (1991) and Stone and Gifford (1997). We noted variation in the timing and level of accumulation of both soluble and insoluble proteins during development (Figure 2.5A). The insoluble protein accumulated to its final levels relatively early (Figure 2.5A), yet individual polypeptides appeared to continue to accumulate beyond stage four (Figure 2.7). In contrast, soluble protein accumulated throughout development and reached the maximum level at the end of development (Figure 2.5A). When the levels of mature, desiccated somatic embryo proteins were compared to those of zygotic embryos, we observed that the somatic embryos contained higher amounts of both soluble and insoluble proteins (Figure 2.5B). The somatic embryos contained 1.5-fold more insoluble protein and five fold more soluble protein. A shift in the ratios of the proteins produced was also observed. In zygotic embryos there were similar amounts of insoluble and soluble proteins present with a soluble to insoluble ratio of roughly three to two. Mature somatic embryos contain over five fold the amount of soluble protein compared to insoluble protein. Thus, the somatic embryos were not only producing more protein overall, but also the protein production was biased more heavily towards soluble proteins.

Increases in soluble protein in the embryos could be due to the accumulation of specific storage proteins or may indicate a general increase in metabolic proteins (enzymes) accompanying active cellular metabolism. To identify differences in storage protein accumulation between zygotic and somatic embryos, we generated electrophoretic profiles for both insoluble and soluble protein extracts in the presence

and absence of β-mercaptoethanol as a reducing agent. The major soluble proteins in mature somatic embryos appear similar to those present in zygotic embryos under reducing and non reducing conditions. Soluble proteins (Figure 2.6) approximately 40-60kD in size under reducing conditions appear to be more abundant in somatic embryos. Conversely, reduced soluble proteins of 14 and 14.5kD in size were not observed in the somatic embryo extracts. Thus, while some differences between the somatic and the zygotic embryos are apparent the soluble electrophoretic protein profiles are generally similar. Therefore we do not believe that the overall increase in soluble protein observed in the somatic embryos is due to an unusual pattern of storage protein accumulation nor to the overproduction of an individual polypeptide or class of proteins.

The profile of phosphate buffer insoluble protein extracted from loblolly somatic embryos also resembles that of zygotic embryos (Figure 2.7). A 47kD insoluble protein is present in both somatic and zygotic embryos and is observed very early in somatic embryo development. This protein persists upon treatment with β–ME and has been previously described in loblolly pine (Groome et al. 1991) and interior spruce (Newton et al. 1992, Flinn et al. 1993). Given this protein's size, structure and solubility it is similar to proteins that have been described as a 7S globulin seed storage proteins (Newton et al 1992, Flinn et al. 1993).

The presence of a 60kD insoluble protein that disappears when treated with β-ME and the concurrent appearance of a 37.5kD and a 22.5kD under reducing conditions is also consistent with previous reports of loblolly zygotic embryos (Groome et al. 1991) and with *Pinus strobus* somatic embryos (Klimaszewska et al.

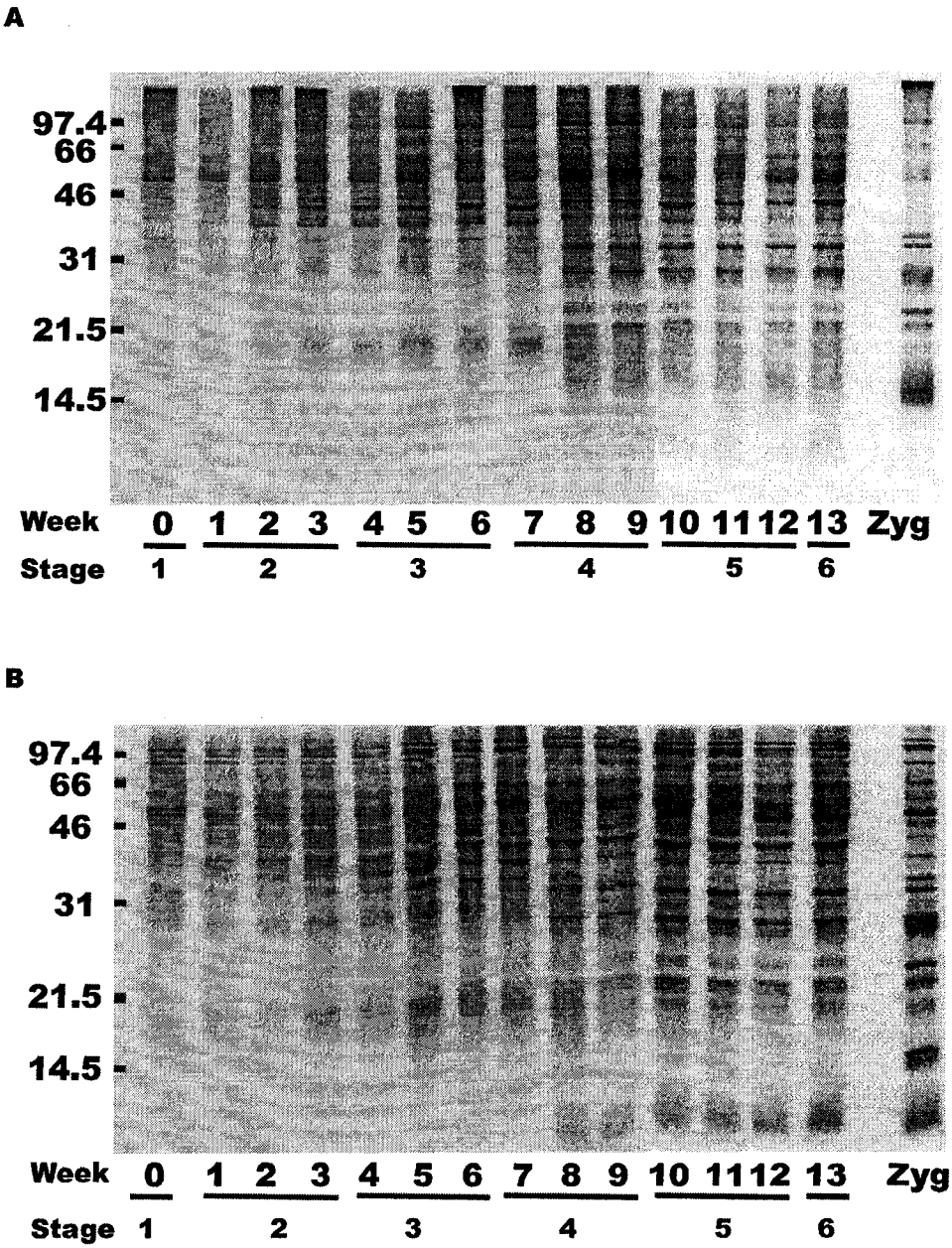


Figure 2.6: Soluble protein profile over somatic embryo maturation time course. Protein was isolated each week followed by SDS-PAGE gels analysis. Protein masses in kD are indicated to the left. Embryo stages as indicated in Table 1. Zyg=mature zygotic embryos. A. SDS-PAGE of protein run under non-reducing conditions. B. SDS-PAGE of protein run under reducing conditions.

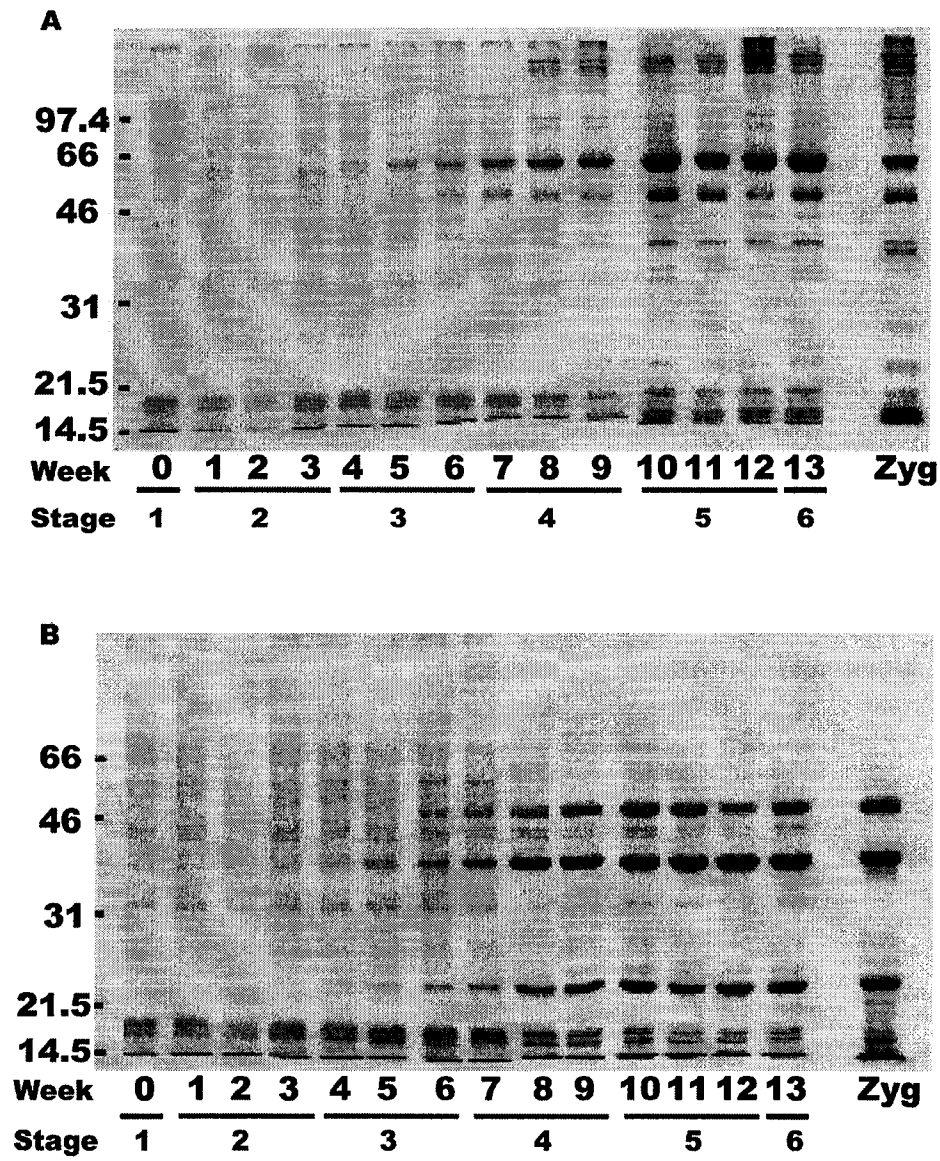


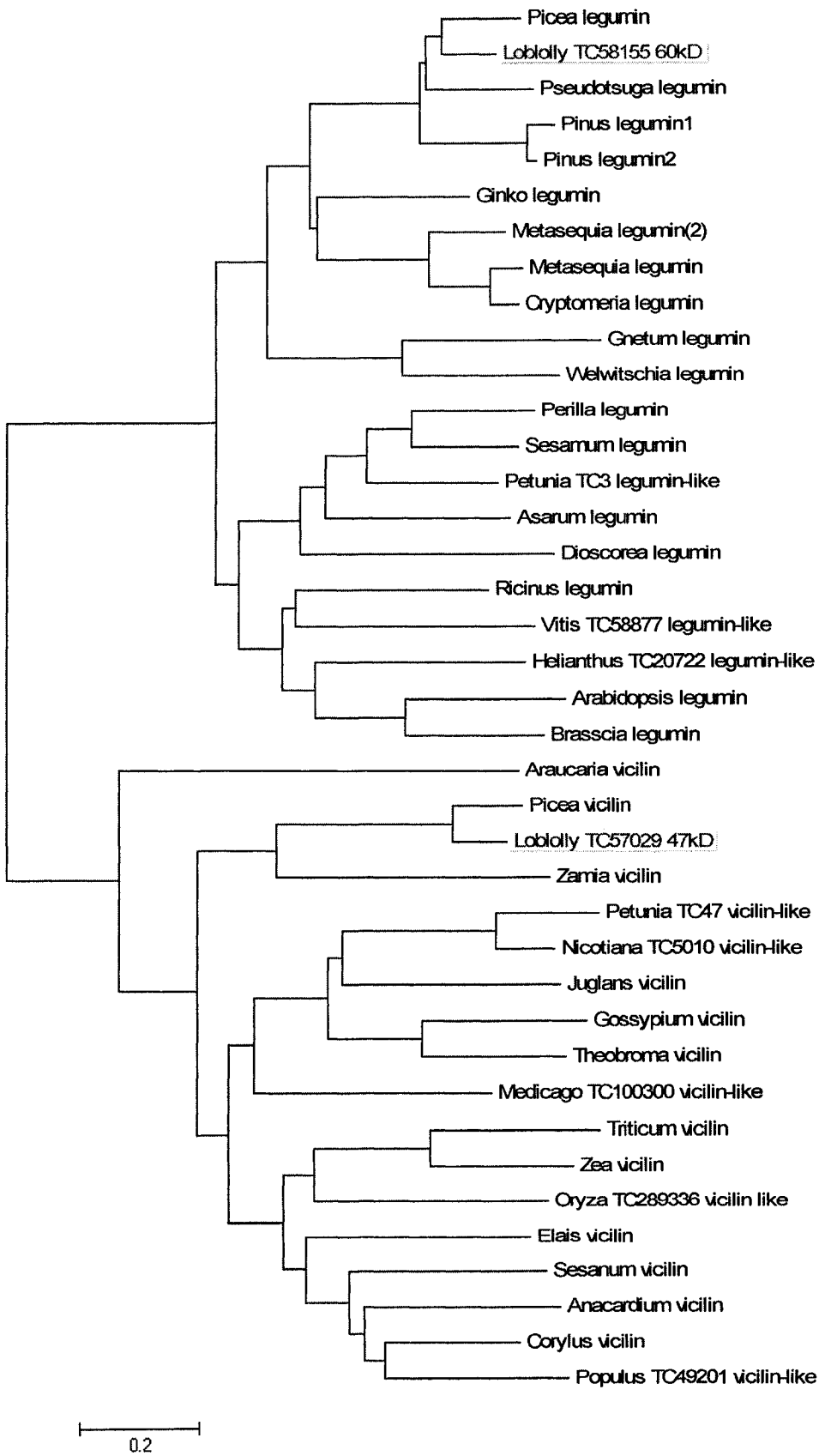
Figure 2.7: Insoluble protein profile over somatic embryo maturation time course. Protein was isolated each week followed by SDS-PAGE analysis. Protein masses in kD are indicated to the left. Embryo stages as indicated in Table 1. Zyg=mature zygotic embryos. A. SDS-PAGE of protein run under non-reducing conditions. B. SDS-PAGE of protein run under reducing conditions.

2004). This 60kD insoluble protein is present in somatic embryos still in liquid culture (Stage 1) increases after early maturation and continues to accumulate until the embryos reach maturation. Under reducing conditions the 37.5kD polypeptide can be visualized during the cell culture while the 22.5kD band is not detected until Stage 2. Although we can identify some differences in protein profiles of somatic and zygotic embryos, they are quite similar overall.

In an attempt to better characterize these seed storage proteins in loblolly pine we completed a phylogenetic analysis of the globulin seed storage proteins (Figure 2.8). Amino acid sequencing and mass spectrometry of the 60kD insoluble protein in *Pinus strobus* showed 100% homology to an 11S pine-globulin-like seed storage protein (Klimaszewska et al. 2004). We then utilized this *Pinus strobus* amino acid sequence to identify a consensus EST (expressed sequence tag) from loblolly pine (TC58155) using the Gene Index (www.compbio.dfci.harvard.edu/tgi/). A loblolly pine consensus EST sequence (TC57029) was also identified for the 47kD insoluble protein using a previously described cDNA in interior spruce (Newton et al. 1992). The open reading frame of each EST was predicted and the loblolly pine amino acid sequences were compared to the globulins of other species.

Similar to what was described by Shewry et al. (1995) and observed by Shutov et al. (2003) the globulin seed storage proteins can be divided into two distinct groups: the 7S vicilin seed storage proteins and the 11S legumin storage proteins. The 47kD loblolly pine protein (TC57029) grouped with the vicilin proteins and is similar in structure to other vicilins. The reported lack of disulphide bridges in the 7S vicilin proteins (Shewry et al. 1995) is consistent with our observations of this protein in

Figure 2.8 Phylogenetic analysis of the globulin seed storage proteins. Amino acid sequences were aligned using Clustal W (Higgins et al. 1994) and tree was constructed using Mega 4 (Tamura et al. 2007). The loblolly pine sequences are the predicted acid sequences of consensus ESTs of the 60kD protein (TC58155) and 47kD protein (TC57029) isolated from the pine Gene Index (www. compbio.dfci.harvard.edu/tgi/). Other predicted amino acid sequences from the Gene Index EST collections are also present and labeled as the genus followed by the tentative consensus identification number used in the Gene Index. Legumin GenBank accession numbers are as follows: Dioscorea (CAA64763), Perilla (AAF19607), Sesamum (AAD42944), Asarum (CAA64761), Picea (CAA44874), Pseudotsuga (AAA68981), Metasequia (CAA64791), Metasequia 2 (X95545), Cryptomeria (CAA64790), Ginko (CAA90641), Gnetum (CAA90642), Welwitschia (CAA90643), Ricinus AAF73008, Arabidopsis (AAL91248), Brassica (BAC80213), Pinus 1 (Q41018), Pinus 2 (Q41017). Vicilin GenBank accession numbers are as follows: Triticum (AAA34269), Zea (CAA41809), Elais (AAK28402), Theobroma (S22477), Gossypium (AAA33069), Juglans (AAF18269), Anacardium (AAM73729), Corylus (AAL86739), Sesanum (AAK15089), Picea (CAA44873), Zamia (CAA90652), Araucaria (AAM81249).



loblolly pine and the size is consistent with globulins annotated as vicilins from interior spruce (Newton et al. 1992, Flinn et al. 1993). Similarly the 60kD loblolly pine protein (TC58155) grouped with legumin proteins and is similar in structure to other 11S globulin seed storage proteins. The legumins are characteristically composed of two subunits of approximately 40kD and 20kD in size linked by a disulphide bridge (Shewry et al. 1995). This is similar to the 60kD protein of loblolly pine described here (and previously describe by Groome et al. 1991) which has subunits of 37.5kD and 22.5kD linked by a disulphide bridge. In addition, an antibody raised against a similar protein in white spruce cross reacted with legumins from several angiosperm species (Misra and Green 1994); this further supports the conclusion that the 60kD protein that dissociates upon treatment with β–ME is an 11S globulin type protein.

The albumins are another separate class of seed storage proteins. In addition to being found in many dicot seeds (Shewry et al. 1995) albumins have also been identified in conifers (Chattai and Misra 1998). The conifer albumins are similar in solubility to the loblolly pine globulins; they are insoluble in 50mM sodium phosphate buffer. The albumins are relatively small in size, generally containing two subunits linked together by a disulphide bridge (Shewry et al. 1995). The smaller albumin subunit is approximately 3-5 kD in size and the larger subunit ranges in size from 8-12 kD (Krebbers et al. 1988). The insoluble protein profile (Figure 2.7) of both zygotic and somatic embryos does not indicate the presence of this class of seed storage protein; however, this may be due to insufficient resolution on the gels. A search of the loblolly pine EST library (www.compbio.dfci.harvard.edu/tgi/) using a

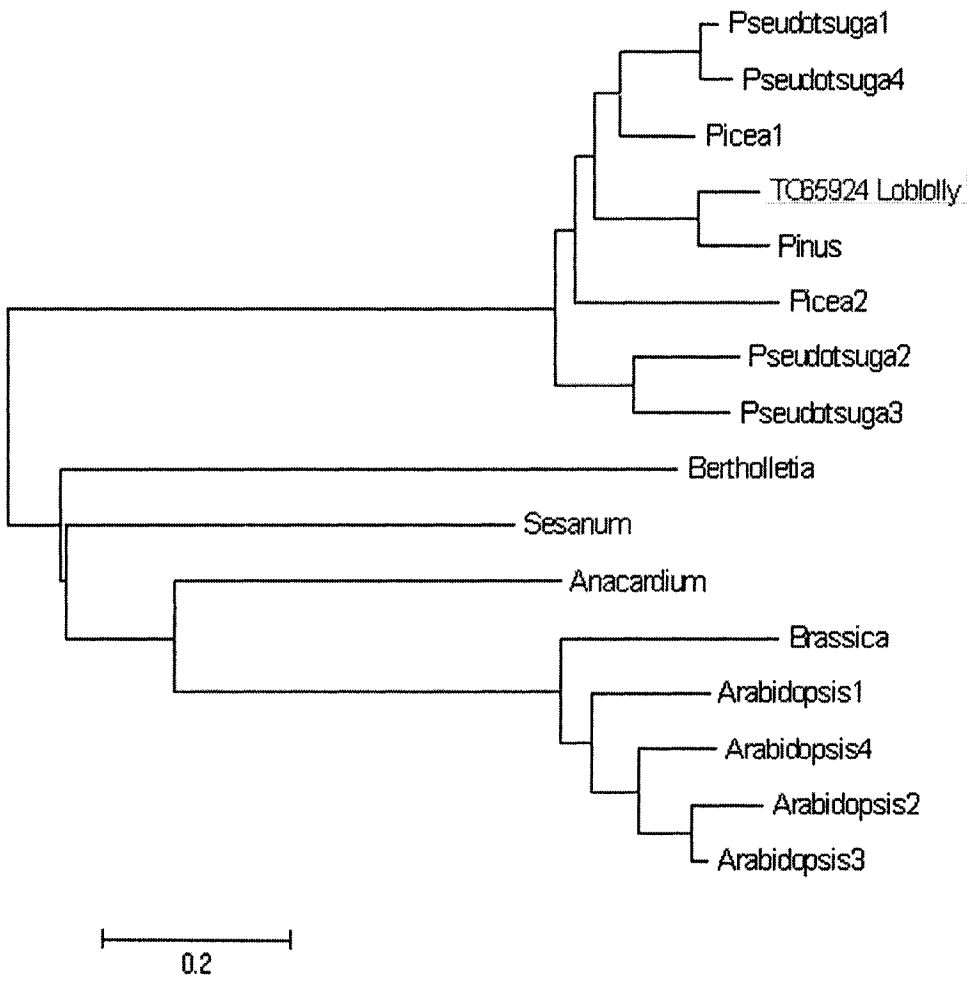


Figure 2.9 Phylogenetic analysis of the 2S albumin seed storage proteins. Amino acid sequences were aligned using Clustal W (Higgins et al. 1994) and tree was constructed using Mega 4 (Tamura et al. 2007). The loblolly pine sequence is the predicted acid sequence of a consensus EST (TC 65924) isolated from the Gene Index (www. compbio.dfci.harvard.edu/tgi/). GenBank accession numbers are as follows Pseudotsuga 1 (AF029970), Pseudotsuga 2 (AF029971), Pseudotsuga 3 (AF029972), Pseudotsuga 4 (AF029973), Picea 1 (X63193), Picea 2 (L47745), Pinus (X62433), Arabidopsis 1 (CAA80870), Arabidopsis 2 (CAA80871), Arabidopsis 3 (CAA80868), Arabidopsis 4 (CAA80869), Brassica (CAA35580), Bertholletia (CAA38362), Anacardium (AAL91665), Sesanum (AAK15088)

Douglas fir 2S albumin cDNA sequence (Chattai and Misra 1998) revealed the presence of albumin ESTs in loblolly pine. A phylogenetic analysis of the predicted amino acid sequence (generated from the consensus EST, TC65924) of the loblolly pine albumin indicated that it was similar to other conifer albumin sequences (Figure 2.9). The phylogenetic analysis also showed a separation of the conifer albumins from the angiosperm albumins; this is consistent with previous reports of a low level of sequence similarity between conifer and angiosperm albumins (Chattai and Misra 1998).

In our study we have attempted to classify the major seed storage proteins observed in the loblolly pine somatic embryos and observed similarities between the electrophoretic protein profiles of somatic and zygotic embryos. However, we also observed a bias towards soluble protein production in somatic embryos, which may have a detrimental effect on their viability. In the megagametophyte, 80% of total protein is buffer-insoluble (King and Gifford 1997). Thus, insoluble proteins comprise the majority of the storage protein reserves for a developing seedling. Since a somatic embryo does not have these reserves present, and the reserves it does possess are qualitatively different, it may be difficult for somatic embryos to survive and produce viable seedlings. In addition, there are less TAGs available to the somatic embryo. TAGs make up 59% of the total storage reserves in the mature loblolly pine seed (Stone and Gifford 1999) yet the somatic embryos contain less TAGs but more protein than their zygotic counterparts. Combined, these data show the storage reserves present in these somatic embryos are vastly different to those that are available to the zygotic embryo.

Attempts have recently been made to make somatic embryos (or their culture environment) more similar to zygotic embryos. Pullman and Buchanan (2003) analyzed megagametophyte and zygotic embryo tissue during development to determine the elemental composition. This provided information for possible ways to make culture conditions more similar to the environment of a developing zygotic embryo. Similarly, Pullman et al. (2003b) analyzed somatic and zygotic embryos to determine differences in metal composition. Subsequently, they modified the media to produce somatic embryos that had a gene expression pattern more similar to zygotic embryos. By first identifying how current somatic embryos differ from their zygotic counterparts we may then be able to find conditions which produce somatic embryo lines that have a higher proportion of viable embryos and exhibit more vigorous growth.

2.4 References

Attree SM, Pomeroy MK, Fowke LC. 1992. Manipulation of conditions for the culture of somatic embryos of white spruce for improved triacylglycerol biosynthesis and desiccation tolerance. Planta 187:395-404.

Attree SM, Pomeroy MK, Fowke LC. 1995. Development of white spruce (*Picea glauca* (Moench) Voss) somatic embryos during culture with abscisic-acid and osmoticum, and their tolerance to drying and frozen storage. Journal of Experimental Botany 46: 433-439.

Beardmore T and Charest PJ. 1995. Black spruce somatic embryo germination and desiccation tolerance. I. Effect of abscisic acid, cold, and heat treatments on the germinability of mature black spruce somatic embryos. Canadian Journal of Forest Research 25: 1763-1772.

Becwar MR, Nagmani R, Wann SR. 1990. Initiation of embryogenic cultures and somatic embryo development in loblolly pine (*Pinus taeda*). Canadian Journal of Forest Research 20: 810-817.

Bewley JD and Black M. 1985. Seeds: physiology of development and germination. Plenum Press New York, New York.

Bomal C and Tremblay FM. 1999. Effect of desiccation to low moisture content on germination, synchronization of root emergence, and plantlet regeneration of black spruce somatic embryos. Plant Cell, Tissue and Organ Culture 56: 193-200.

Brownfield DL, Todd CD, Stone SL, Deyholos MK, Gifford DJ. 2007. Patterns of storage protein and triacylglycerol accumulation during loblolly pine embryo maturation. Plant Cell Tissue and Organ Culture 88: 217-223.

Chattai M and Misra S. 1998. Sequence and expression of embryogenesis-specific cDNAs encoding 2S seed storage proteins in *Pseudotsuga menziesii* [Mirb.] Franco. Planta 206: 138-145.

Dunstan DI, Dong JZ, Carrier DJ, Abrams SR. 1998. Events following ABA treatment of spruce somatic embryos. In Vitro Cell and Developmental Biology (plant) 34: 159-168.

Feirer, RP, Conkey JH, and Verhagen SA. 1989. Triglycerides in embryonic conifer calli: a comparison with zygotic embryos. Plant Cell Reports 8: 207-209.

Flinn BS, Roberts DR, Newton CH, Cyr DR, Webster FB, Taylor IEP. 1993. Storage protein gene expression in zygotic and somatic embryos of interior spruce. Physiologia Plantarum 89: 719-730.

Groome MC, Axler SR, Gifford DJ. 1991. Hydrolysis of lipid and protein reserves in loblolly pine seeds in relation of protein electrophoretic patterns following imbibition. Physiologia Plantarum 83: 99-106.

Gupta PK and Durzan DJ 1987. Biotechnology of somatic polyembryogenesis and plantlet regeneration in loblolly pine. Bio/technology 5: 147-151.

Harms WR, Whitesell CD, DeBell DS. 2000. Growth and development of loblolly pine in a spacing trial planted in Hawaii. Forest Ecology and Management 126: 13-24.

Higgins D, Thompson J, Gibson T. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22: 4673-4680.

Janick J, Velho CC, Whipkey A. 1991. Developmental changes in seeds of loblolly pine. Journal of American Society of Horticultural Science 116: 297-301.

Jeleska S. 1994. Recent developments in somatic embryogenesis. Acta Pharm 44: 397-406.

King JE and Gifford DJ. 1997. Amino acid utilization in seeds of loblolly pine during germination and early seedling growth. Plant Physiology 113:1125-1135.

Klimaszewska K, Morency F, Jones-Overon C, Cooke J. 2004. Accumulation pattern of seed storage proteins in zygotic embryos of Pinus strobes and in somatic embryos from different maturation treatments. Physiologia Plantarum 121: 682-690.

Krebbers E, Herdies L, De Clercq A, Seurinck J, Leemans F, Van Damme J, Segura M, Gheysen G, Van Montagu M, Vandekerckhove J. 1988. Determination of the process sites of an *Arabidopsis* 2S albumin and characterization of the complete gene family. Plant Physiology 87: 859-866.

Mhaske VB, Chengalrayan K, Hazra S. 1998. Influence of osmotica and abscisic acid on triglyceride accumulation in peanut somatic embryos. Plant cell reports 17:742-746.

Misra S and Green M. 1994. Legumin-like storage polypeptides of conifer seeds and their antigenic cross-reactivity with 11S globulins from angiosperms. Journal of Experimental Botany 45: 269-274.

Newton CH, Flinn BS, Sutton BCS. 1992. Vicilin-like seed storage proteins in the gymnosperm interior spruce (*Picea glauca/engelmanii*). Plant Molecular Biology 20: 315-322.

Pullman GS and Buchanan M. 2003. Loblolly pine (*Pinus taeda* L.): stage-specific elemental analyses of zygotic embryo and female gametophytic tissue. Plant Science 164: 943-953.

Pullman GS, Johnson S, Peter G, Cairney J, Xu, N. 2003. Improving loblolly pine somatic embryo maturation: comparison of somatic and zygotic embryo morphology, germination, and gene expression. Plant Cell Reports 21:747-758.

Pullman GS, Montello P, Cairney J, Xu N, Feng X. 2003b. Loblolly pine (*Pinus taeda* L.) somatic embryogenesis: maturation improvements by metal analysis of zygotic and somatic embryos. Plant Science 164: 955-969.

Schultz RP. 1999. Loblolly-the pine for the twenty-first century. New Forests 17: 71-88.

Shewry PR, Napier JA, Tatham AS. 1995. Seed storage proteins: structures and biosynthesis. The Plant Cell 7: 945-956.

Shutov AD, Bäulmlein H, Blattner FR, Müntz K. 2003. Storage and mobilization as antagonistic functional constraints on seed storage globulin evolution. Journal of Experimental Botany 54:1645-1654.

Stasolla C, Kong L, Yeung EC, Thorpe TA. 2002. Maturation of somatic embryos in conifers: Morphogenesis, physiology, biochemistry, and molecular biology. In Vitro Cellular and Developmental Biology 38: 93-105.

Stasolla C and Yeung EC. 2003. Recent advances in conifer somatic embryogenesis: improving somatic embryo quality. Plant Cell, Tissue and Organ Culture 74: 15-35.

Stone SL and Gifford DJ. 1997. Structural and biochemical changes in loblolly pine (*Pinus taeda* L.) seeds during germination and early seedling growth. I. Storage protein reserves. International Journal of Plant Science 158: 727-737.

Stone SL and Gifford DJ. 1999. Structural and biochemical changes in loblolly pine (*Pinus taeda* L.) seeds during germination and early seedling growth. II. Storage triacylglycerols and carbohydrates. International Journal of Plant Science 160: 663-671.

Taber RP, Zhang C, Hu WS. 1998. Kinetics of Douglas-fir (*Pseudotsuga menziesii*) somatic embryo development. Canadian Journal of Botany 76:863-871.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24: 1596-1599.

Xu N and Bewley JD. 1991. Sensitivity to abscisic acid and osmoticum changes during embryogenesis of alfalfa (*Medicago sativa*). Journal of Experimental Botany 42: 821-826.

Xu N and Bewley JD. 1994. Desiccation and the switch from seed development to germination. Alfalfa embryos can synthesize storage proteins after germination if maturation drying is prevented. Seed Science Research 4: 247-255.

Xu N and Bewley JD. 1995. Temporal and nutritional factors modulate responses to abscisic acid and osmoticum in their regulation of storage protein synthesis in developing seeds of alfalfa (*Medicago sativa L.*) Journal of Experimental Botany 46: 675-686.

Chapter 3: Arginase in loblolly pine

3.1 Introduction

A mature pine seed is composed of a diploid embryo surrounded by a maternally derived haploid megagametophyte (Raven et al. 1999). The embryo is situated in the corrosion cavity of the megagametophyte, which is formed during seed development. To improve the number of successfully germinating embryos, loblolly pine seeds are frequently stratified (Carpita et al. 1983). Stratification is the constant exposure of seeds to moisture in a low, non-freezing temperature (2-4°C) for a given period of time (Raven et al. 1999); after the stratification period the seeds are transferred to conditions suitable for germination. Germination occurs over approximately 4 days, and the completion of germination is marked by the emergence of the radicle through the micropyle and the split in the seed coat (Tillman-Sutela and Kauppi 2000). After germination is complete, the seedling moves into early seedling growth, which lasts until the seedling is fully photosynthetic and autotrophic. Until the seedling is autotrophic it must rely on its storage reserves to provide all of the nutrients for growth and development.

The storage reserves of loblolly pine seeds are located both in the embryo and in the megagametophyte, with the majority of the reserves found in the megagametophyte. There are 2 main types of storage reserves present in loblolly pine seeds: storage proteins and triacylglycerols (TAGs). TAGs comprise 59% of the total storage reserves present in the mature loblolly pine seeds (Stone and Gifford 1999).

After radicle emergence, TAG breakdown is initiated by lipases which release the components through a three stage hydrolysis reaction (Bewley and Black 1985). A single TAG, once degraded, releases a molecule of glycerol and 3 fatty acid molecules. The glycerol is further broken down to yield either hexose units or pyruvate. The fatty acids released by lipases proceed through β–oxidation, the glyoxalate cycle, and gluconeogenesis to produce sucrose as an end product. The sucrose is then exported to the developing seedling.

The second major type of storage reserve is storage protein. The majority of the seed storage proteins are contained in the megagametophyte, where approximately 1.9mg out of an average of 2.1mg total protein per seed is located (Stone and Gifford 1997). The storage proteins are contained in protein vacuoles which decrease in number as the proteins within them are hydrolyzed after the completion of germination (Stone and Gifford 1997). Proteinases catalyze the hydrolysis of the storage proteins to reduce them to their constituent amino acids. The amino acids released may be re-used for protein synthesis or may be deaminated to provide the seedling with carbon skeletons (Bewley and Black 1985).

The amino acids released from storage protein hydrolysis are not present in equal amounts. In mature seeds, Arg is the most prevalent amino acid in the buffer-insoluble proteins of the megagametophyte; it accounts for 23% of all amino acids (King and Gifford 1997). The next most abundant amino acids are Glu and Gln together at 21% (these amino acids could not be differentiated from each other during analysis). Assuming that the combined glutamate and glutamine portion was entirely glutamine (which contains an additional nitrogen atom per molecule), approximately

20% of the total nitrogen found in the buffer insoluble megagametophyte storage proteins is present as Glu and Gln (King and Gifford 1997). Similarly, four nitrogen atoms are present in each Arg molecule. Thus, as a conservative estimate, Arg constitutes 45% of the total nitrogen present in the buffer insoluble proteins of the megagametophyte in mature seeds (King and Gifford 1997).

The prevalence of Arg is maintained as the storage proteins are degraded. Arginine is the third most abundant free amino acid in the megagametophytes of mature desiccated seeds, and increases in proportion during early seedling development. By $12DAI_{30}$ (12 days after imbibiton at 30°C, 8 days after radicle emergence) it is the most abundant free amino acid in the megagametophyte, and represents almost half of the stored nitrogen in the insoluble protein of this tissue. Taken together, this shows that Arg most likely plays a significant role in providing the seedling with nitrogen (King and Gifford 1997).

For Arg to be used as a nitrogen source for the developing seedling it first must be catabolized. This is completed by the enzyme arginase. Arginase hydrolyzes Arg into ornithine and urea. The urea is further degraded by urease, releasing ammonia and carbon dioxide. Avila et al. (2001) have proposed that the ammonia is then incorporated into Gln by GS1a, which is one of two isoforms of glutamine synthetase 1 (GS1) identified in pine. Glu is frequently used as a nitrogen donor in transamination reactions; thus, the N released from the storage proteins is made available to other processes and molecules. GS1a RNA has been localized to the cotyledons and upper hypocotyls in Scots pine (Avila et al. 2001), the same location where arginase has been shown to be active in loblolly pine (King and Gifford 1997).

As the seedling moved closer to autotrophy an increase in the asparagine (Asn) content of the seedling was observed (King and Gifford 1997); Avila et al. (2001) believed this increase to be due to GS1b. GS1b uses the ammonia from storage proteins and in conjunction with asparagine synthetase (AS) to produce Asn. Asn, a common nitrogen transport molecule in plants, then supplies other areas of the developing seedling with nitrogen.

GS1b RNA has been localized to the vascular cells (Avila et al. 2001) as has AS (Nakano et al. 2000). The co-localization of AS and GS1b in the vascular cells supports the role of GS1b and AS in the transport of nitrogen in the seedling. In addition, even though Arg is the most prevalent free amino acid in 12 DAI_{30} megagametophytes, Asn is over 5-fold more abundant than arginine in the seedling at the same stage (King and Gifford 1997). This further supports the presence of an Arg to Asn conversion system in the seedling late in early seedling growth.

Arginase is mainly localized to the cotyledons of the developing seedling. Although arginase RNA and protein are also present in the megagametophyte and the root pole, they are much less abundant (Todd et al 2001, Todd et al 2001b). Arginase transcript is detectable at low levels in mature and stratified seeds and increases following germination (at $4DAI_{30}$) until it reaches its maximum at $10DAI_{30}$ (Todd et al 2001b). Arginase protein accumulation and specific activity follow similar patterns, both increasing after germination and then reaching maximum levels. Arginase specific activity peaks at $8DAI_{30}$ (King and Gifford 1997) while protein accumulates for another two days before beginning to decrease (Todd et al 2001b). Taken together, these data indicate that arginase is regulated both temporally and spatially.

Arginase activity is also regulated in part by the megagametophyte. Todd et al. (2002) showed that although arginase expression is initiated in the seedling, maintenance or up-regulation requires the presence of the megagametophyte. Arginase activity can also be induced artificially. If the megagametophyte is removed from a seedling growing in culture, the transcript becomes undetectable 24 hours after megagametophyte removal (Todd et al. 2002). If Arg is exogenously applied to the seedling at the time of megagametophyte removal, arginase transcript levels are maintained at a higher level than control seedlings. Similarly, if Arg is applied to a seedling ($6DIC_{30}$) with low or undetectable arginase transcript levels, a significant increase in transcript abundance occurs (Todd et al. 2003). In all cases mentioned above, arginase specific activity is correlated with RNA abundance. However, the protein levels are not altered by the removal of the megagametophyte or the spraying of Arg; this is most likely due to the relative long life of the arginase protein (Todd et al. 2002, 2003)

An cDNA encoding arginase was previously cloned from a loblolly pine expression library (Todd et al. 2001). The 1366bp cDNA contains four methionine (Met) codons in the first 40bp of the cDNA sequence, all of which are in frame with the first Met. Of these 4 Met codons, 2 are surrounded by translation initiation consensus sequences of plants (Todd et al. 2001). The presumed stop codon is located 1026bp downstream from the first start codon and concludes the 341 amino acid open reading frame. There are also 2 potential polyadenylation signals 193bp and 226bp downstream of the stop codon. Southern analysis indicated that arginase was likely a single copy gene (Todd et al. 2001b)

In the current study, our goal was to isolate the genomic region upstream of the arginase gene in order to understand how arginase was regulated in pine. Currently, no reports of conifer arginase promoters exist. We wanted to identify the minimal promoter as well as cis-regulatory regions. In this study we report the first step in this goal, namely the isolation of DNA sequence from upstream of pine arginase.

Arginase is an enzyme important in N metabolism in plants, and understanding it will allow us to understand how N utilization is regulated in loblolly pine during germination and early seedling growth. N is an important nutrient and is supplied to the seedling by the storage reserves present both in the embryo and the megagametophyte. Conifer somatic embryos lack a megagametophyte and as a result are provided with significantly fewer nutrients than their zygotic counterparts. This lack of storage reserves has been suggested to be the cause of the poor conversion ratios observed for the somatic embryos of some conifer species (Attree et al 1992, Mhaske et al 1998), which includes loblolly pine. If we can understand how N is utilized in the seedling and how this is regulated perhaps we can also use this information to improve the process of somatic embryogenesis.

3.2. Materials and methods

3.2.1 Pine genomic DNA extraction

Loblolly pine seed was provided by Westvaco (Summerville, S.C.) The seed was collected from an open pollinated stand of clonal trees (clone 11-9). Seed was surface sterilized and imbibed in sterile MilliQ water. Embryos were isolated and stored at -80°C until use.

Embryos (10g) were homogenized in liquid nitrogen with sand in a mortar and pestle transferred the powder to plastic beaker with 100mL of grinding buffer (0.3M sucrose, 50mM Tris-Cl pH 8.8, 5mM $MgCl_2$, 5% PVP-40, 2% DIECA added just before use) and solution was stirred until thawed. The suspension was then filtered through 3 layers of Miracloth into four 50mL centrifuge tubes and spun at 3500 rpm for 10 minutes at 4°C in a Beckman JA-20 rotor. After centrifugation, the supernatant was removed and the pellet was resuspended in 4mL of lysis buffer (50 mM Tris-Cl pH 8.0, 20 mM EDTA, 1% N-laurylsarcosine). CsCl was added to lysate to a final concentration of 0.95g/mL and solution was stirred on a hot plate until dissolved. The solution was spun at 15,000 rpm for 20 minutes at 4°C after which the protein and starch were removed with a cotton swab. The supernatant from all four tubes was pooled and ethidium bromide was added to a concentration of 300μg/mL. The solution was transferred into 5mL ultracentrifuge tubes (Corning) and centrifuged at 53,000 rpm at 25°C for 15-20 hours in an ultracentrifuge with a Beckman VTi 65.2 rotor.

Genomic DNA was collected with a 15 gauge needle by puncturing the side of the ultracentrifuge tube. The solution collected was placed in new ultracentrifuge tubes, topped up with fresh CsCl solution (0.95g/mL), and centrifuged at 50,000 rpm for 22 hours at 25°C. The bands containing genomic DNA were again isolated with a 15 gauge needle and extracted several times with 1X TNE (10mM Tris-Cl pH 7.5, 0.1mM EDTA, 10mM NaCl) saturated with butanol. The purified aqueous solution was then dialyzed using dialysis tubing (molecular weight cutoff 6,000-8,000kDa, Corning) overnight against 1X TNE with 1 buffer change.

3.2.2 Genome walking

Genome walking was performed using the Universal Genome Walker kit (Clontech) with some modifications. We digested 2.5μg of genomic DNA with one of DraI, EcoRV, StuI, PvuII, or ScaI. Blunt ended linkers were ligated onto purified digested genomic DNA. Nested PCR was performed using primers anchored in the linker sequence (AP1 and AP2), gene specific primers (first walk:D9 Rev and D15 second walk: D15 and D27; Table 3.1) and Taq. The cycling conditions of the reaction were as follows: 7 cycles of 94°C for 25 seconds, 72°C for 3 minutes, 32 cycles of 94°C for 25 seconds, 67°C for 3 minutes, and a final extension of 67°C for 7 minutes. The products of the first reaction were analyzed by electrophoresis on an agarose gel, diluted 1:50, and used as template for the second reaction. The products were analyzed on an agarose gel and cloned into pCR4TOPO vector for sequencing.

The sequenced inserts were sub-cloned into pCAMBIA 1391 which contains a GUS reporter gene for plant expression.

3.2.3 DNA sequencing

DNA sequencing was performed using the DYEnamic ET terminator cycle (Amersham Pharmacia Biotech) sequencing premix. The reaction consisted of 2-4μL of pre-mix with 5pmol of primer and 100ng of template. The cycling conditions consisted of 30 cycles: 95°C for 30 seconds, 50°C for 15 seconds, 60°C for 60 seconds, and held at 4°C until use. The DNA was subsequently precipitated according to the manufacturer's recommendations and analyzed on an ABI Prism 373 sequencer (Perkin Elmer Biosystems) at the Department of Biological Sciences Molecular Biology Services Unit.

3.2.4 5' RACE

We extracted total RNA using a modified RNeasy protocol (Qiagen). We homogenized 120mg of $10DAI_{30}$ shoot poles under liquid N_2 and 1.2mL of lysis buffer (4 M guanidine isothiocyanate, 0.2M sodium acetate pH5.0, 25mM EDTA, 2.5% PVP-40, 1% β-mercaptoethanol) was added along with 60μL of 20% sarkosyl. The sample was incubated at 70°C for 10 minutes and the resulting lysate was centrifuged through a QIAshredder column for 2 minutes at 18,000g. We added one half volume of absolute ethanol and the mixture was applied to an RNeasy column.

The sample was centrifuged at 8000g for 40 seconds, washed once with 700μL buffer RW1, and twice with 500μL buffer RPE. RNA was eluted into RNase-free water by centrifugation at 8000g for 2 minutes.

5' RACE was completed using the GeneRacer kit (Invitrogen). RNA was treated with calf intestinal phosphatase at 50°C, phenol purified, and ethanol precipitated. The dephosphorylated RNA was then treated with tobacco acid pyrophosphatase at 37°C for 1 hour to remove the 5' cap from the mRNA. The RNA was again phenol purified and precipitated. The GeneRacer oligo was ligated to the 5'end of the mRNA using T4 RNA ligase and incubating for 1 hour at 37°C; after which the RNA was once again ethanol precipitated. The RNA was then reverse transcribed using AMV-RT at 42°C for 50 minutes, 70°C for 15 minutes, and on ice for 2 minutes. The remaining RNA was then degraded by the addition of RNase H and a 20 minute incubation at 37°C. The 5' end of the arginase cDNA was then amplified using nested PCR. The first reaction used the GeneRacer 5' primer and a gene specific primer (D2, Table 3.1). The cycling conditions were as follows: 94°C for 2 minutes, 5 cycles of 94°C for 30 seconds, 72°C for 2.5 minutes, 5 cycles of 94°C for 30 seconds, 70°C for 30 seconds, 72°C for 2 minutes, 20 cycles of 94°C for 30 seconds, 65°C for 30 seconds, 72°C for 2 minutes, and a final extension of 72°C for 10 minutes. Nested PCR was completed using the GeneRacer 5' Nested Primer as well as a second gene specific primer (D15, Table 3.1). The cycling conditions used were as follows: 94°C for 2 minutes, 20 cycles of 94°C for 30 seconds, 65°C for 30 seconds, 72°C for 10 minutes, followed by a final extension of 72°C for 10 minutes.

The products were analyzed on a 1% agarose gel, and cloned into the pCR4TOPO vector for sequencing.

3.2.5 Protoplast preparation

Protoplasts were prepared using the method of Gómez-Maldonado et al (2001). Loblolly pine seedling cotyledons (4-8 DAI_{30}) were isolated and then excised with a single edged razor blade; the prepared tissue (1g) was then incubated overnight in the dark at 25°C in 10mL of an enzyme mixture (0.4% (w/v) cellulose, 0.4% (w/v) macerase, 0.4% sucrose, 0.44% (w/v) K3 medium). The mixture was filtered through a 75μm sieve and the cells isolated by gravity sedimentation. The cells were then washed 3 times with 0.4M sucrose 80mM KCl solution and resuspended to a final concentration of $1X10^5$cells/mL. Protoplast viability was determined by staining cells with PBS containing fluorescein diacetate ($1.2X10^{-7}$ M) (Rotman and Papermaster 1966).

We mixed 1mL of protoplast suspension with 50μg of plasmid DNA and 10μg of sheared salmon sperm DNA. The mixture was placed in a 0.2cm electrode (Bio-Rad) and electroporated at 800V/cm using an electroporator (ElectroCell Manipulator 630, BTX). The protoplasts were then allowed to recover for 10 mins at room temp and diluted with a volume of incubation solution (0.5M mannitol, 88mM sucrose, 0.5mg/L carbenicillin) and incubated in the dark at 24°C for 36 hours.

3.2.6 Bombardment of seedlings

This work was completed at the University of Laval in collaboration with Dr. Armand Séguin of the Canadian Forest Service (Sainte-Foy, Quebec). Seeds were surface sterilized and then the seed coat and outer integuments removed. The embryo (enclosed in the megagametophyte) was placed on culture medium (100mg/mL KNO_3, 745mg/mL KCl, 170mg/mL monobasic KH_2PO_4, 370mg/mL $MgSO_4$-$7H_2O$, 440mg/mL $CaCl_2$-$2H_2O$, 16.9mg/mL $MnSO_4$-H2O, 8.6mg/mL $ZnSO_4$-$7H_2O$, 0.025mg/mL $CuSO_4$-$5H_2O$, 0.083mg/mL KI, 6.2mg/mL H_3BO_3, 0.25mg/mL Na_2MoO_4-$2H_2O$, 27.84mg/mL $FeSO_4$-$7H_2O$, 37.24mg/mL Na_2EDTA, 0.1mg/ml myo-inositol, 0.8mg/mL NH_4NO_3, 30mg/mL sucrose, 5mg/mL activated carbon pH 5.8) for 6-8 days at 30° under continuous light.

We washed 30mg of gold particles (0.6 or 1μm) in 1mL of absolute EtOH for one hour (vortexing twice) and sonicated for 15 minutes. The particles were then rinsed twice with distilled water and resuspended in a final volume of 500μL and stored at room temperature. The particles were again sonicated for 15 minutes before the addition of 5μg of plasmid DNA (1μg/mL in water). $CaCl_2$ (50μL of 2.5M) and spermine (20μL of 0.1M) were added and the mixture was vortexed. The tube was then centrifuged and the supernatant removed before an EtOH wash and the particles resuspended in 50μL of absolute EtOH.

We pipetted 5μL of the DNA mixture on to the centre of the Kapton disc and allowed to stand until the EtOH evaporated. The tissue (surface sterilized seedlings grown in culture (100mg/mL KNO_3, 745mg/mL KCl, 170mg/mL monobasic

KH_2PO_4, 370mg/mL $MgSO_4$-$7H_2O$, 440mg/mL $CaCl_2$-$2H_2O$, 16.9mg/mL $MnSO_4$-H2O, 8.6mg/mL $ZnSO_4$-$7H_2O$, 0.025mg/mL $CuSO_4$-$5H_2O$, 0.083mg/mL KI, 6.2mg/mL H_3BO_3, 0.25mg/mL Na_2MoO_4-$2H_2O$, 27.84mg/mL $FeSO_4$-$7H_2O$, 37.24mg/mL Na_2EDTA, 0.1mg/ml myo-inositol, 0.8mg/mL NH_4NO_3, 30mg/mL sucrose, 5mg/mL activated carbon pH 5.8) for 6-8 days at 30°C under continuous light) was then bombarded under vacuum (26.5 hg) with the helium pressure set to 200psi more than the rupture disc. The tissue was then placed on culture medium (above) for 24 hours at 30°C under continuous light and then stained for GUS activity.

For the staining of GUS, seedlings were placed in cold 90% acetone for 10 minutes. The seedlings were washed with 50mM phosphate buffer (Na_2HPO_4-$7H_20$ pH 7.2) before staining solution was added (1mM X-gluc, 50mM sodium phosphate, 0.5mM $K_3Fe(CN)_6$, 0.5mM $K_4Fe(CN)_6$. Seedlings were stained 12-24 hrs at 37°C in the dark. The staining reaction was stopped at the first appearance of blue colour by the addition of 70% EtOH. After 12-72hrs the 70% EtOH was replaced with 95% EtOH and the seedlings were stored at room temperature.

Table 3.1: Sequences of pine PCR primers used

Name:	Primer Sequence (5' to 3')
AP1 (Fwd)	GTAATACGACTCACTATAGGC
AP2 (Fwd)	ACTATAGGGCACGCGTGGT
Gene Racer 5' Primer (Fwd)	CGACTGGAGCACGAGGACACTGA
Gene Racer 5' Nested Primer (Fwd)	GGACACTGACATGGACTGAAGGAGTA
D2 (Rev)	GCGAGGAGGAGCGAATGCAGGGCC
D9 (Rev)	GTCCCTTGTTATGATGATTCCACCCCA
D15 (Rev)	ACGGTTTTGGCCCTTCTCTATCATTTGTG
D27 (Rev)	TGCGGACTCCACTTTACGGCATTGGGTGG

3.3 Results

3.3.1 Strategies to isolate the genomic region upstream of arginase

To isolate the region upstream of the loblolly pine arginase gene (i.e. the presumptive promoter including all cis-acting regulatory elements conferring the normal transcriptional expression pattern), we first attempted to screen a phage genomic library that was constructed from embryos of mature seeds. We screened plaques with the 5' end of a loblolly pine arginase cDNA (Todd et al 2001), and approximately 40 plaques that hybridized to the probe were isolated. However, sub-cloning inserts from phage into *E.coli* was problematic, and only part of the DNA present in each plaque could be recovered from the bacteria. Sequencing of these partial inserts in *E. coli* suggested that there were two different, but related, DNA sequences that hybridized to the arginase cDNA (data not shown). However, because we were unable to clone and sequence the phage inserts that had hybridized the cDNA probe, we turned to other techniques for isolating the target region upstream of the arginase gene.

As a second approach, we conducted a genome walk in the upstream direction starting from the previously cloned arginase cDNA (Todd et al 2001). Pine genomic DNA was digested with several different restriction enzymes and linkers were ligated to the ends of the fragments. Forward primers that anchored to the linkers and reverse primers that were gene specific were used to amplify the region of DNA between the primers, upstream of the arginase gene.

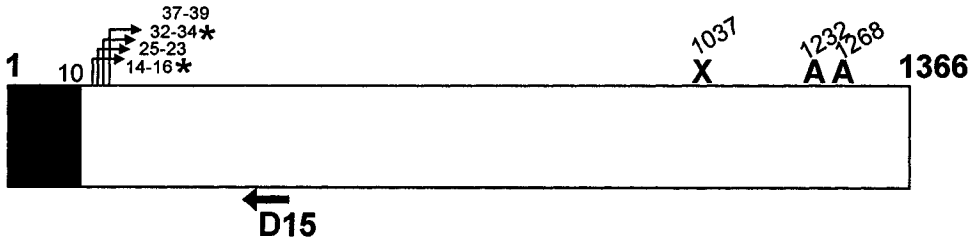


Figure 3.1: The loblolly pine arginase cDNA (Todd et al 2001). The 1366bp cDNA has four possible start codons two of which are surrounded by a translation initiation consensus sequences (*). Assuming the first codon is utilized a 341 amino acid open reading frame is present followed by a stop codon at nucleotide 1037. Two putative polyadenylation sequences (A) are also present at nucleotides 1232 and 1266. The first 10 nucleotides of the cDNA (labeled in black) did not agree with the sequence of the genome walking product. The D15 reverse primer was used to generate the product from the first genome walk as well as RACE.

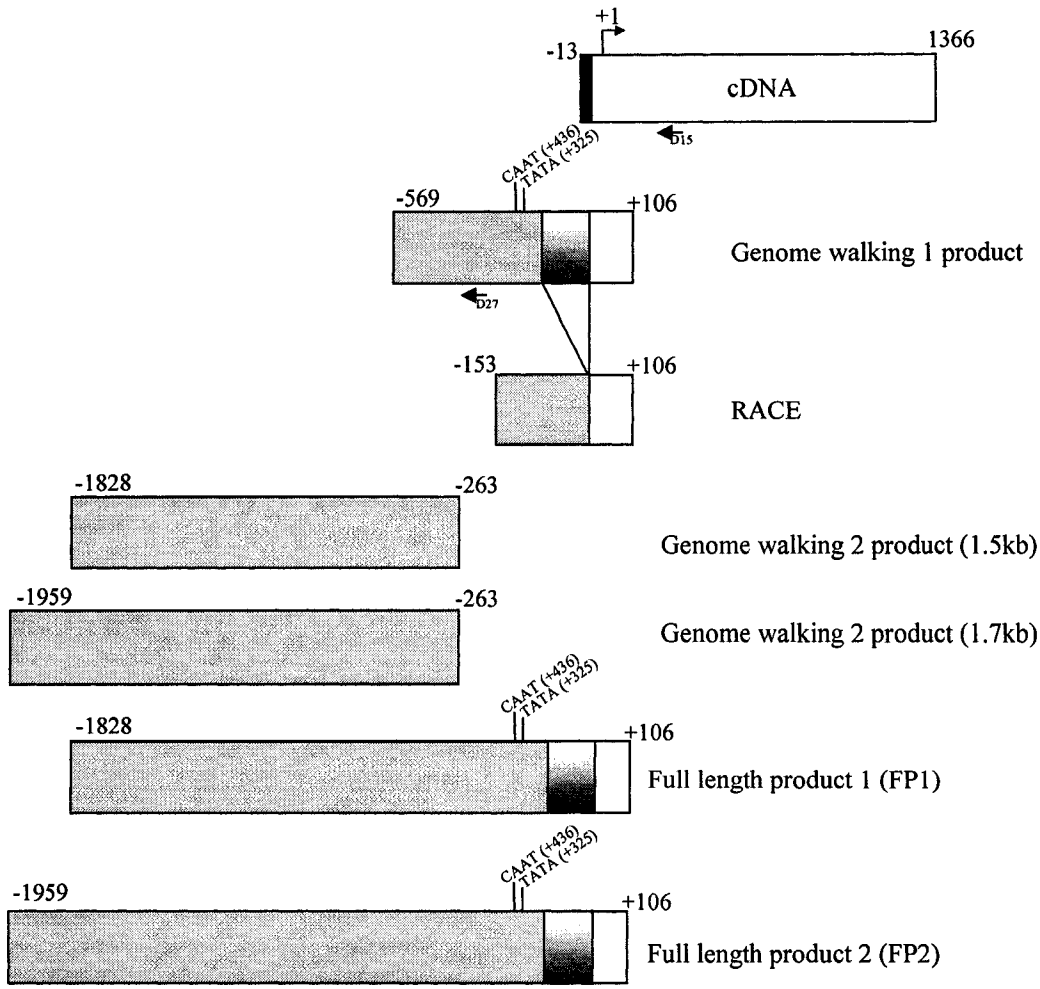


Figure 3.2: Reactions to identify the region upstream of arginase. The previously identified arginase cDNA was used to design reverse primers (D2 and D15) for genome walking. RACE was conducted to confirm the alignment of the genome walking product with the cDNA sequence. Genome walking product 1 was used to design another reverse primer (D27) for a second set of genome walking reactions. Genomic PCR was then used to amplify the full length products (FP1 and FP2) from the genome and these were cloned into pCAMBIA 1391Z. All regions upstream of the start codon are shaded

3.3.2 Genome walking

The arginase cDNA (Figure 3.1) is 1366bp in length and contains four potential start codons in the first 40bp of sequence, two of which are surrounded by translation initiation consensus sequences characteristic of plants (Todd et al 2001). For the purposes of this report, we have assumed that the 5'-most start codon that is also surrounded by a consensus sequence is the translation initiation codon. This shall be referred to as the translation start codon and has been labeled as +1 (Figure 3.1 and 3.2). Beginning from this presumed start codon, a 341 amino acid open reading frame is present, and there are 2 potential polyadenylation signals present as well (Todd et al 2001). The 5' end of this cDNA was used to design the first set of gene specific primers for genome walking.

Two sets of genome walking reactions were completed in order to isolate a sufficient length of sequence for analysis (Figure 3.2). The first set of reactions (D15 reverse primer) produced PCR products of 674bp and 313bp. Sequencing and subsequent alignment of the PCR products revealed that the smaller product was a sub-fragment of the larger product. Both products aligned to the cDNA from the genome walking PCR primer site (located in the cDNA) and continued to align through most of the cDNA as expected. The longest fragment was used for subsequent work (genome walking product 1, Figure 3.2). The alignment also showed a ten bp mismatch in the genomic PCR walking products as compared to the 5' terminus of the cDNA. This apparent mismatch was detected in both genome walking

products, and in repeated sequencing reactions. The cDNA had likewise been sequenced repeatedly during library screening (Todd et al. 2001), making it unlikely that the mismatched region was due to sequencing errors in the cDNA.

cDNAs obtained from libraries are often not accurate in the terminal nucleotides (GeneRacer information, Invitrogen) and therefore we considered it likely that the first 10 nucleotides of the cDNA sequence were inaccurate. Alternatively, it was also possible that the genome walking products were not contiguous with the cDNA or an insertion/intron was present in the genome walking products. In order to determine if the mismatched region was due to a problem in the 5'sequence of the cDNA or if the genome walking products were not contiguous with the cDNA, we conducted 5'RACE to independently determine the sequence of the mRNA, thus eliminating the possibility that the observed differences in genomic and cDNA sequences were due to cloning errors or other artifacts. Primers were created from the cDNA sequence downstream of the mismatched region (D15 reverse primer) and 5' RACE was peformed. The RACE product was approximately 700bp in size and aligned with the sequence of the genome walking PCR, except for a 150bp insertion present in the genomic sequence (Figure 3.2).

To determine whether the 150bp region that we detected in derivatives of genomic DNA but not mRNA was an intron, we compared RT-PCR and genomic PCR products on an agarose gel. We observed that primers in same region amplified a larger fragment from genomic DNA than from cDNA, indicating the presence of an intron in the 5' untranslated region (UTR) of the arginase transcript. The identification of a 53 or 79 nucleotide intron in the 5' UTR of arginase from

Table 3.2: Consensus sequences of the 5' and 3' splice sites in monocots, dicots and the arginase 5' UTR intron. Adapted from Simpson and Filipowicz (1996). Frequency of the presence of each nucleotide at each position is shown with the consensus nucleotide in bold. A. The nucleotide frequency, consensus sequence and arginase sequence for the 5' splice site. B. The nucleotide frequency, consensus sequence and arginase sequence for the 3' splice site.

A

		Exon			Intron					
		-3	-2	-1	1	2	3	4	5	6
Dicots	%G	17	8	**79**	**100**	0	11	3	**50**	10
	%A	**36**	**62**	9	0	0	**69**	**58**	23	22
	%T	15	20	9	0	**99**	15	25	18	**53**
	%C	31	10	2	0	1	4	14	9	16
Monocots	%G	19	7	**78**	**100**	0	19	7	**59**	10
	%A	**40**	**65**	7	0	0	**64**	**44**	17	17
	%T	7	14	7	0	**99**	9	22	13	**53**
	%C	34	14	8	0	1	8	26	10	20
Consensus		A	A	G	G	T	A	A	G	T
Arginase		T	T	G	A	G	G	T	A	T

B

		Intron			Exon		
		-3	-2	-1	1	2	3
Dicots	%G	1	0	**100**	**57**	20	26
	%A	4	**100**	0	21	22	**32**
	%T	32	0	0	11	**44**	28
	%C	**63**	0	0	11	14	14
Monocots	%G	1	0	**100**	**60**	23	23
	%A	3	**100**	0	15	17	20
	%T	15	0	0	11	**40**	**31**
	%C	**80**	0	0	13	20	24
Consensus		C	A	G	G	T	A/T
Arginase		T	A	C	A	G	C

Aspergillus nidulans has also been reported (Borsuk et al 2007). No significant homology was detected upon BLAST alignment of the pine intron sequence with either splice variants of the *A. nidulans* intron

3.3.3 Analysis of the putative intron in the 5' UTR

Sequence analysis indicated that consensus splice sites of higher plants are not present in the arginase intron (Table 3.2). The 5' consensus splice site of AG/GTATGT and the 3' consensus of AG/GT (Simpson and Filipowicz 1996) are not present in the pine intron as seen in Table 3.2. Relative to angiosperms there is less sequence information about gymnosperm introns and intron splice sites, however some information has been reported. Genomic sequence from GS1a has been isolated from Scots pine and the introns detected adhere to the general G/GT 5' consensus and AG/G 3' consensus (Avila et al 2002).

Plant introns generally have a high AT content (Simpson et al 1993). The putative 5' UTR intron we identified has a 53% AT content. While the 53% AT content observed in pine is low, it is within the range of plant AT content for plant introns: 42-89% AT for dicots and 31-80% for monocots (Simpson et al 1993).

3.3.4 Extension of genome walking product isolation

As indicated earlier, the first set of genome walking reactions yielded approximately 600bp of sequence upstream of the start codon of the arginase cDNA.

To extend the length of cloned sequence further upstream, we completed a second set of genome walking reactions (D27 reverse primer). This second set of PCR reactions produced two products of approximately 1.5kb and 1.7 kb in size (Figure 3.2). In order to isolate the full length upstream region (that was identified in 2 separate genome walking reactions) genomic PCR was used to amplify the full length upstream regions. The first full length product (FP1) was 1934bp in length and the second (FP2) was 2065bp in length. Present in both these sequences are putative TATA and CAAT boxes. These are located in a highly homologous region at the 3' end of FP1 and FP2 (Figure 3.3). The TATA box is located approximately 162bp upstream of the intron (318bp upstream of the translation start codon) and the CAAT box is 276bp upstream of the intron (432bp upstream of the translation start codon); TATA boxes are usually located approximately 25-30bp upstream of the transcription initiation site (Joshi 1987).

3.3.5 The repetitive elements of the upstream sequence

Upon sequencing FP1 and FP2 we identified two types of repetitive elements present in both products and denoted them repeat sets one and two (R1, R2, Figure 3.3). FP1 contained two direct tandem repeats (R1A and R1B) that were 147bp in length. FP2 contained a modified version of these elements, with one complete repeat (R1B) and a truncated copy (R1A). Thus, FP1 contained two full copies of repeat one, while FP2 contained only one full length copy but also had the 3' end of repeat one present.

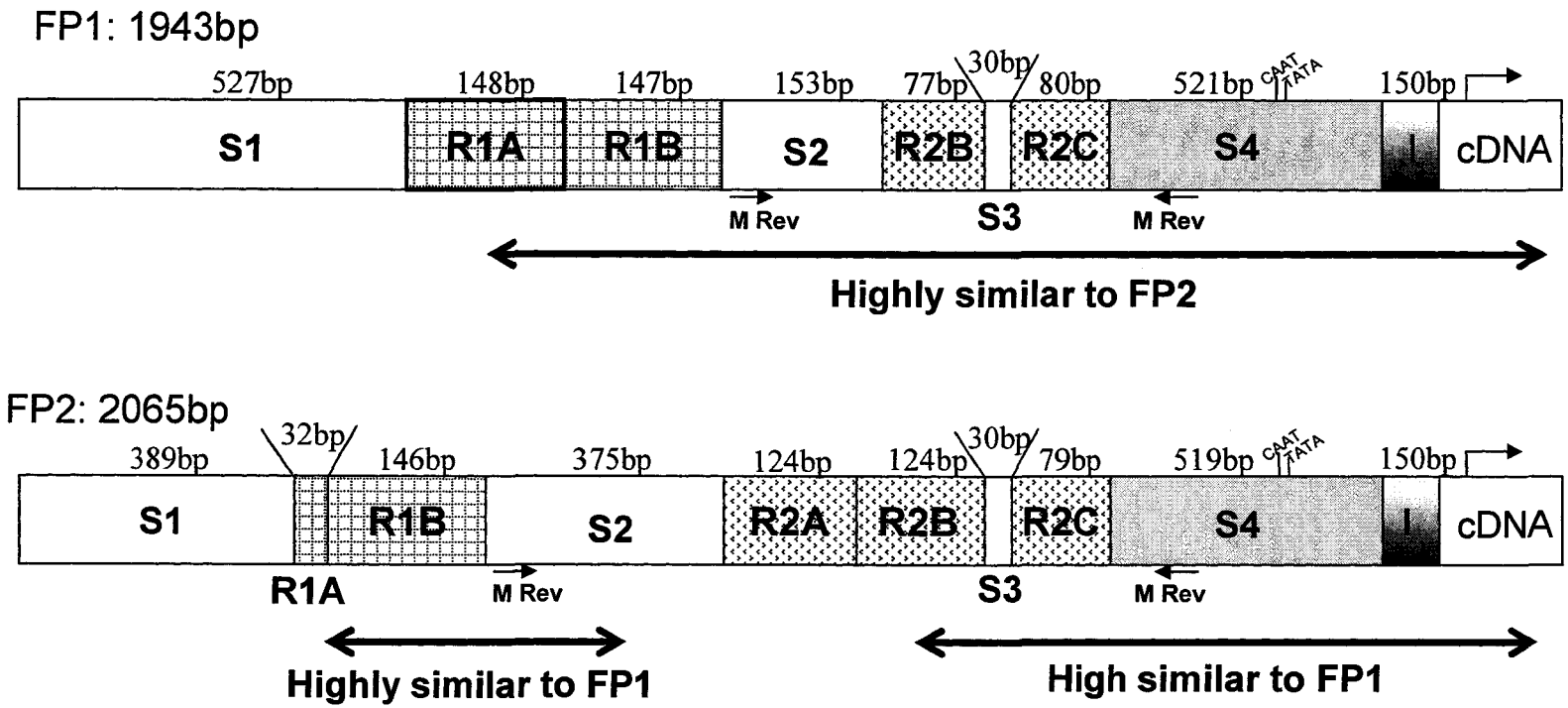


Figure 3.3: Full length genome walking products 1 and 2. S1 is a unique region at the 5' end of each product. Repeat set 1 (R1) is present in 2 copies in bothFP1 and FP2 (R1A and R1B). S2 is a second region that lies between the sets of repeats. Repeat 2 is present as 2 copies (R2B and R2C) in FP1 and 3 copies in FP2 with R2A and R2B expanded to 124bp from the approximately 80bp of the other elements of repeat 2. S3 is a region that spaces R2B and R2C in both products. S4 follows the second set of repeats and is highly similar in both products. The intron (I) is followed by cDNA sequence. Primers M Fwd and M Rev were used to differentiate FP1 and FP2 in PCR.

The second type of repetitive element (R2) was a 77bp direct repeat. FP1 contains two copies of this repeat (R2B and R2C) while FP2 contains three copies (R2A, R2B and R2C). However, R2A and R2B are slightly longer (at 124bp each) than R2C, although all three of these fragments contain the conserved 77bp repeat. BLAST alignment of these 77bp elements in Genbank did not identify significant homology in any other organism. Similarly, repetitive elements were also been identified in the sequence upstream of the cinnamyl alcohol dehydrogenase (CAD) gene of loblolly pine (MacKay 1996).

The majority of the sequence present in FP1 was also present in FP2. Figure 3.3 shows the homologous regions of both FP1 and FP2. The homologous region appears to be split in FP2, resulting in two regions that are both homologous to FP1. The homologous region of FP2 containing R1A and R1B was 89% identical to that found in FP1 (nucleotide identity). Similarly, the region of FP2 containing the 3' end of R2B until the start codon was 97% identical (nucleotide identity). Regions with no significant homology are also present. The terminal 5' regions of the genome walking products were unique sequences with no significant nucleotide homology to each other (Figure 3.3). This unique sequence is 527bp in length in FP1 and 389bp in FP2. Another region with no significant nucleotide homology was also present. This unique region included the extra copy of the second repeated sequence present in FP2 (R2A) and some of its upstream sequence.

3.3.6 Determination of FP1 and 2 as alleles or separate genes

The isolation of two distinct genome walking products adjacent to the arginase gene in the loblolly pine genome led us to question whether arginase was indeed a single copy gene as was predicted by Todd et al (2001b). It was possible that FP1 and FP2 represented two separate genes; alternatively it was also possible that FP1 and FP2 were alleles of a single gene. In order to test this we conducted PCR using DNA isolated from individual megagametophytes. The seed used for the genome walking reactions was from an open-pollinated, clonal population of trees. As the tree population was clonal, all seeds had the same maternal heritage but because the trees were open pollinated they may differ in their paternal lineage. Megagametophytes are haploid seed tissue derived from the maternal tree. Thus, if FP1 and FP2 were alleles of a single gene, both would not be present in a single, haploid megagametophyte. However, if FP1 and FP2 were separate genes, then it would be possible for both to be present in a single megagametophyte. Primers were designed in a polymorphic region to differentiate the products of FP1 and FP2 (M Fwd and M Rev, Figure 3.3), and genomic DNA was isolated from individual megagametophytes. We used DNA from diploid embryos as a positive control (Figure 3.4) to show that it was possible to amplify both products in a single PCR reaction.

In order to test the hypothesis that FP1 and 2 represent separate genes (denoted A and B) we considered several different scenarios, given that the maternal genotype at both hypothetical loci and the linkage relationship of these loci are unknown. Underlying all of these scenarios is the assumption that we could detect only one allele of each gene (A and B), but that the other alleles (denoted a and b)

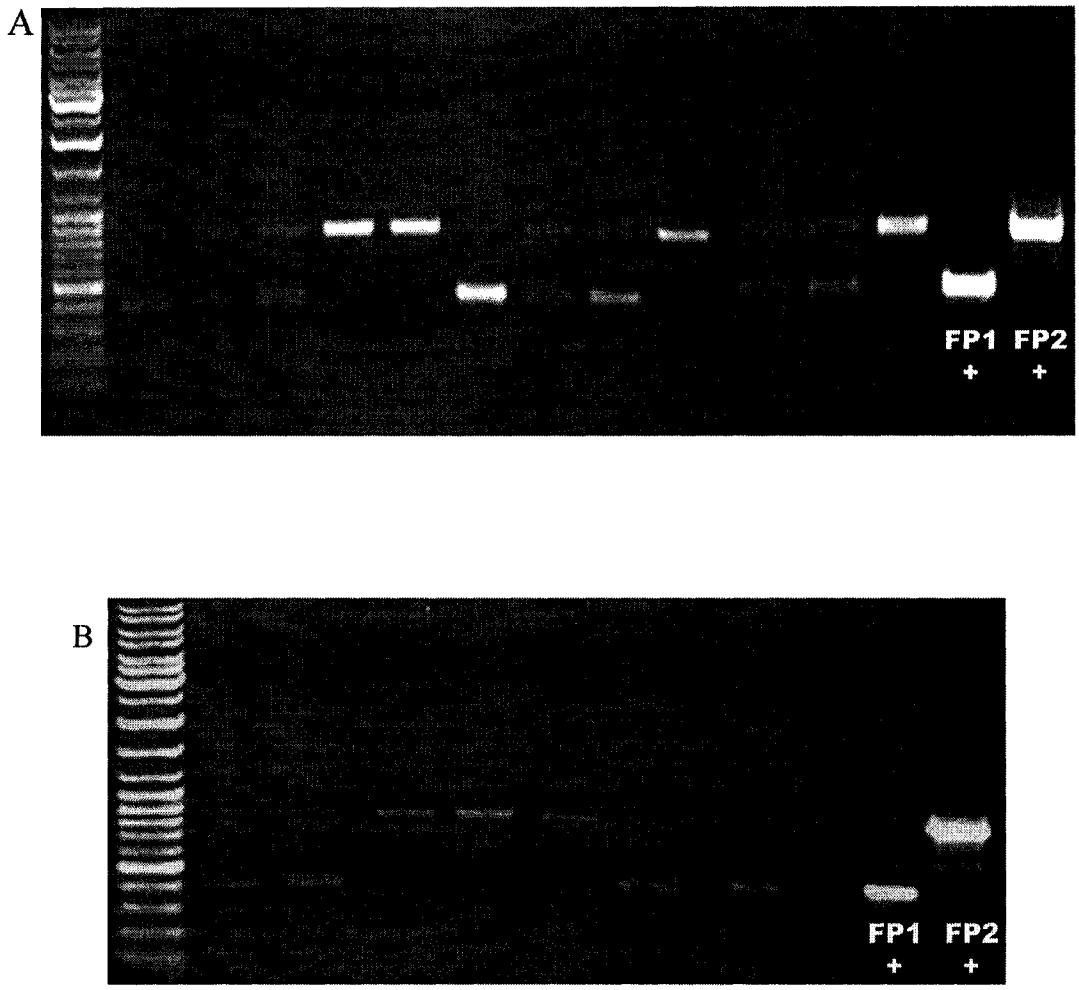


Figure 3.4: PCR of unique regions of FP1 and FP2. A. PCR products from individual embryos. Products FP1 and FP2 are distinguishable as homozygous and heterozygous individuals based on band intensity. B. PCR from individual megagametophytes. Products from FP1 and FP2 did not appear together in any of the 37 megagametophytes successfully amplified.

were not detectable by our PCR assay. Thus, the diploid maternal tree had four possible genotypes: AaBb, AABb, AaBB and AABB; aabb was not a possible maternal genotype as A and B alleles were detected in individual megagametophytes. Furthermore if a given megagametophyte did not yield an amplification product it could have been either because it had the genotype ab, or because of reaction failure for technical reasons independent of genotype.

We conducted PCR genotyping reactions on DNA from 90 individual gametophytes. Of 37 successful amplifications, 17 megagametophytes amplified the FP1 product (A) and 20 megagametophytes amplified the FP2 (B) product. Thus, 53 reactions either failed or were megagametophytes with the undetectable ab alleles. Chi square analyses were conducted for different scenarios (Table 3.3) with three degrees of freedom and a critical value of 7.8 (at $\alpha = 0.05$). Most of the chi squares rejected the null hypothesis that FP1 and FP2 are separate genes even when the critical value was corrected for multiple testing using the Bonferroni procedure (increasing the critical value to 16.2). The set of assumptions in which the null hypothesis was not rejected was that the maternal genotype was AaBb, with allele A and allele b on the same chromosome, with complete linkage, and assuming that all blank lanes were from failed reactions. Thus, two possibilities still exist to explain FP1 and FP2: first that FP1 and FP2 represent two separate genes and are completely linked and the maternal tree is heterozygous at both loci and second that they represent alleles of a single gene.

Table 3.3: Chi square analysis of megagametophyte PCR. Genomic DNA from individual megagametophytes was isolated and PCR was conducted to amplify a polymorphic region to differentiate the either FP1 or FP2. Chi squares test the null hypothesis that FP1 and 2 represent 2 separate genes. The gene for FP1is represented as A, while its undetectable allele is represented by a. Similarly the gene for FP2 is represented as B, while its undetectable allele is represented as b.

A. Assuming the maternal parent is AABb, there is no linkage, that only A and B are detectable and that all blank lanes are ab

	AB	Ab	aB	ab	
frequency	0.25	0.25	0.25	0.25	
total reactions	90	90	90	90	
expected	**22.5**	**22.5**	**22.5**	**22.5**	
observed	0	17	20	53	
chi-sq value	22.50	1.34	0.28	41.34	**65.46**

B. Assuming the maternal parent is AaBb, there is no linkage, that only A and B are detectable and that all blank lanes are failed reactions

	AB	Ab	Ba	ab	
frequency	0.25	0.25	0.25	0.25	
total reactions	37	37	37	37	
expected	**9.25**	**9.25**	**9.25**	**9.25**	
observed	0	17	20	0	
chi-sq value	9.25	6.49	12.49	9.25	**37.48**

C. Assuming the maternal parent is AaBb, complete linkage, that only A and B are detectable and that all blank lanes are ab

	AB	Ab	Ba	ab	
frequency	0	0.5	0.5	0	
total rxns	90	90	90	90	
expected	**0**	**45**	**45**	**0**	
observed	0	17	20	53	
chi-sq value	0	17.42	13.89	∞	∞

D. Assuming the maternal parent is AaBb, complete linkage, that only A and B are detectable and that all blank lanes are failed reactions

	AB	Ab	Ba	ab	
frequency	0	0.5	0.5	0	
total rxns	37	37	37	37	
expected	**0**	**18.5**	**18.5**	**0**	
observed	0	17	20	0	
chi-sq value	0	0.122	0.122	0	0.24

E. Assuming the maternal parent is AABb, there is no linkage, that only A and B are detectable and that all blank lanes are ab

	AB	Ab	aB	ab	
frequency	0.5	0.5	0	0	
total rxns	90	90	90	90	
expected	**45**	**45**	**0**	**0**	
observed	0	17	20	53	
chi-sq value	45	17.42	∞	∞	∞

F. Assuming the maternal parent is AABb, there is no linkage, that only A and B are detectable and that all blank lanes are failed reactions

	AB	Ab	aB	ab	
frequency	0.5	0.5	0	0	
total rxns	37	37	37	37	
expected	**18.5**	**18.5**	**0**	**0**	
observed	0	17	20	0	
chi-sq value	18.5	0.12	∞	0	∞

G. Assuming the maternal parent is AABb, complete linkage, that only A and B are detectable and that all blank lanes are ab

	AB	Ab	aB	ab	
frequency	0.5	0.5	0	0	
total rxns	90	90	90	90	
expected	**45**	**45**	**0**	**0**	
observed	0	17	20	53	
chi-sq value	45	17.42	∞	∞	∞

H. Assuming the maternal parent is AABb, complete linkage, that only A and B are detectable and that all blank lanes are failed reactions

	AB	Ab	aB	ab	
frequency	0.5	0.5	0	0	
total rxns	37	37	37	37	
expected	**18.5**	**18.5**	**0**	**0**	
observed	0	17	20	0	
chi-sq value	18.5	0.12	∞	0	∞

I. Assuming the maternal parent is AaBB, there is no linkage, that only A and B are detectable and that all blank lanes are ab

	AB	Ab	aB	ab	
frequency	0.5	0	0.5	0	
total rxns	90	90	90	90	
expected	**45**	**0**	**45**	**0**	
observed	0	17	20	53	
chi-sq value	45	∞	13.89	∞	∞

J. Assuming the maternal parent is AaBB, there is no linkage, that only A and B are detectable and that all blank lanes are failed reactions

	AB	Ab	aB	ab	
frequency	0.5	0	0.5	0	
total rxns	37	37	37	37	
expected	**18.5**	**0**	**18.5**	**0**	
observed	0	17	20	0	
chi-sq value	18.5	∞	0.12	0	∞

K. Assuming the maternal parent is AaBB, complete linkage, that only A and B are detectable and that all blank lanes are ab

	AB	Ab	aB	ab	
frequency	0.5	0	0.5	0	
total rxns	90	90	90	90	
expected	**45**	**0**	**45**	**0**	
observed	0	17	20	57	
chi-sq value	45	∞	13.89	∞	∞

L. Assuming the maternal parent is AaBB, complete linkage, that only A and B are detectable and that all blank lanes are failed reactions

	AB	Ab	aB	ab	
frequency	0.5	0	0.5	0	
total rxns	37	37	37	37	
expected	**18.5**	**0**	**18.5**	**0**	
observed	0	17	20	0	
chi-sq value	18.5	∞	0.12	0	∞

M. Assuming the maternal parent is AABB, there is no linkage, that only A and B are detectable and that all blank lanes are ab

	AB	Ab	aB	ab	
frequency	1	0	0	0	
total rxns	90	90	90	90	
expected	**90**	**0**	**0**	**0**	
observed	0	17	20	53	
chi-sq value	90	∞	∞	∞	∞

N. Assuming the maternal parent is AABB, there is no linkage, that only A and B are detectable and that all blank lanes are failed reactions

	AB	Ab	aB	ab	
frequency	1	0	0	0	
total rxns	37	37	37	37	
expected	**37**	**0**	**0**	**0**	
observed	0	17	20	0	
chi-sq value	37	∞	∞	**0**	∞

O. Assuming the maternal parent is AABB, complete linkage, that only A and B are detectable and that all blank lanes are ab

	AB	Ab	aB	ab	
frequency	1	0	0	0	
total rxns	90	90	90	90	
expected	**90**	**0**	**0**	**0**	
observed	0	17	20	53	
chi-sq value	90	∞	∞	∞	∞

P. Assuming the maternal parent is AABB, complete linkage, that only A and B are detectable and that all blank lanes are failed reactions

	AB	Ab	aB	ab	
frequency	1	0	0	0	
total rxns	37	37	37	37	
expected	**37**	**0**	**0**	**0**	
observed	0	17	20	0	
chi-sq value	37	∞	∞	0	∞

3.3.7 Transient expression of promoter deletions

To determine if the minimal promoter was present in both FP1 and 2, we fused a deletion series for each of FP1 and 2 to a GUS reporter gene (Figure 3.5). Products were amplified from genomic DNA using reverse primers anchored in unique sequences of FP1 and FP2 and a forward primer anchored in the cDNA. Since we did not know which of the possible start codons is used for translation initiation, and all of the potential start codons are in the same reading frame, both of the codons surrounded by translation initiation codons were included in the promoter-reporter fusions. A total of three products unique to FP1 measuring 1829bp, 1297bp, and 965bp in length were generated. For FP2, we generated 4 unique products: 1938bp, 1756bp 1351bp and 985bp in length. Due to the highly repetitive nature of the sequences it was not possible to produce identically sized PCR products from each of FP1 and FP2. In addition to these unique products, two additional deletions from the region common to both FP1 and FP2 (97% nucleotide identity) were created; the resulting products were 691bp and 279bp in length. All 9 PCR products were individually cloned into the pCAMBIA 1391 vector in a GUS reporter casette.

The pCAMBIA 1391 vector contains a promoter-less GUS reporter gene adjacent to a multiple cloning site making it ideal for promoter analysis. It also contains a hygromycin resistance gene for selection *in planta* that is controlled by a CaMV 35S constitutive promoter. The 35S promoter contains strong enhancer domains which have been reported to produce GUS activity when only a minimal

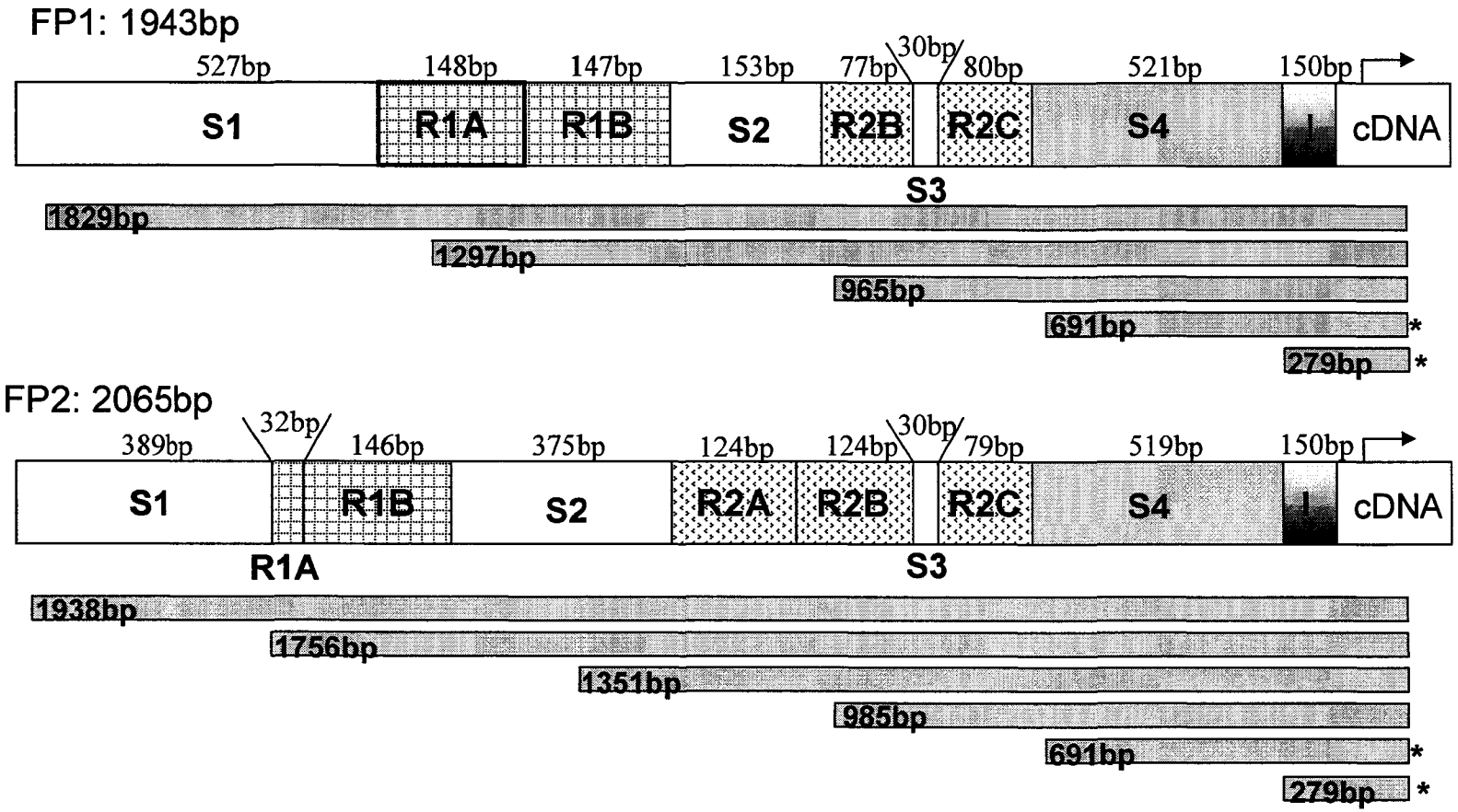


Figure 3.5: Deletions products of FP1 and FP2. Three products unique to FP1were produced: 1829bp, 1297bp and 956bp. Four products unique to FP2 were produced 1938bp, 1756bp, 986bp. Two products (*) of 691bp and 279bp were also produced, these are common to both FP1 and FP2 as they are in the region that is 97% identical.

promoter is present adjacent to the reporter (www.cambia.org). In order to avoid this possibility in our constructs we sub-cloned the region containing our insert, the GUS coding sequence and the NOS terminator into pBluscript. This additional cloning step eliminated the problematic CaMV 35S promoter issue but allowed us to utilize the reporter fusions in certain types of transformation experiments.

In order to quantify expression levels from the reporter fusions, we attempted transient expression using pine protoplasts. Using a previously described procedure (Gómez-Maldonado et al 2001) protoplasts were prepared from loblolly pine cotyledons and stained with fluorescein diacetate to determine their viability. Protoplasts were then electroporated for transient expression analysis. Subsequent staining of the protoplasts indicated that no (or very few) viable cells existed after electroporation. Thus, we attempted to utilize other methods for transient expression. This work was completed by other individuals in collaboration with Dr. Armand Séguin (Canadian Forest Service, Sainte-Foye, Quebec). Transient expression via the bombardment of pine seedlings had only limited success (Figure 3.6). While GUS expression was detectable for each of the full length constructs it was at very low levels, which was unexpected. Arginase is known to be expressed at high levels in the developmental stage of the tissues bombarded (Todd et al. 2001), but the levels of GUS staining were not consistent with high arginase expression.

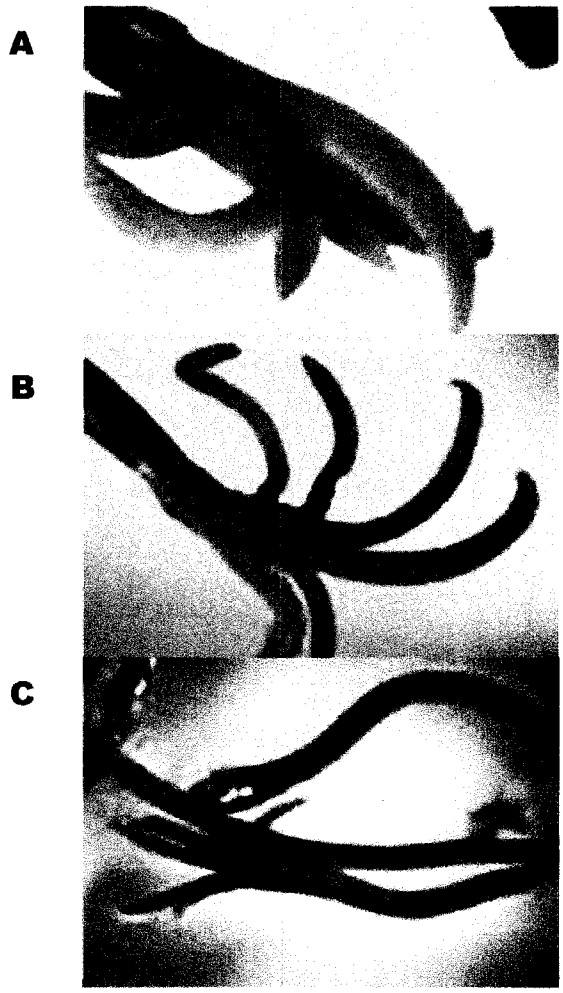


Figure 3.6 : GUS staining of seedlings bombarded with arginase constructs containing the largest promoter fragment. A. Full length FP1-GUS was bombarded and showed low levels of GUS activity. B. Full length FP2-GUS was bombarded and showed low levels of GUS activity. C. Ubiquitin-GUS fragment was bombarded and showed high levels of GUS activity.

3.4 Discussion

3.4.1 Analysis of regions upstream of arginase: FP1 and FP2 5' UTR

To further understand the regulation of arginase in pine, we used genome walking to isolate a genomic DNA fragment immediately upstream of the coding region of this gene. Genome walking has been used successfully to isolate promoters from several species of *Pinus*, including loblolly pine (Ciavatta et al. 2002), radiata pine (Höfig et al. 2003) and Caribbean pine (Mathieu et al. 2003). Using this technique, we isolated 2 different fragments, FP1 and FP2, which both appeared to be adjacent to an arginase coding region in the loblolly pine genome. FP1 and FP2 were 1943bp and 2065bp in length, respectively. Within each of these fragments, we detected a putative intron, a TATA box, a CAAT box, and repetitive sequences. Each of these structural features will be discussed in the following paragraphs.

3.4.2 The 5' UTR intron

Introns in the 5' UTR are often found within many plant genes including the polyubiquitin and sucrose synthase genes (Norris et al. 1993, Fu and Park 1995). The location of the 5' UTR introns relative to the start codon seems to vary along with the size of the intron. The putative 5'UTR intron we identified is located only 4bp upstream of the first possible translation initiation codon. In Arabidopsis, 5' UTR introns are located immediately before the start codon of the *UBQ3*, *UBQ10* and

UBQ11 genes (Norris et al. 1993). Similarly, three rice polyubiquitin promoters have been identified and the introns present are immediately upstream of the start codon. Located further upstream, introns in the 5' UTRs of the potato the *SUS3* and *SUS4* are located 31 and 38bp upstream of the start codon (Fu and Park 1995). The majority of introns in higher plants are less than 150bp in length with very few being longer than 2-3kb (Simpson and Filipowicz 1996). The pine 5' UTR intron is within the range of most introns: 150bp long; however, introns identified in the 5' UTRs of other genes appear to be longer than that of arginase. For example, in Arabidopsis the 5' UTR introns of the *UBQ3*, *UBQ10* and *UBQ11* genes are 375bp, 304bp and 334bp in length (Norris et al. 1993). Similarly, the three rice polyubiquitin introns are 782bp, 962bp, and 1140bp in length and thus are longer than both the pine and previously mentioned Arabidopsis introns (Wang et al. 2000, Sivamani and Qu 2006). Still larger introns have been identified in the 5' UTR of two sugarcane polyubiquitin genes where they are 1360bp and 1374bp in length (Wei et al. 2003) and in maize where the *Ubi1* gene has a 1010bp 5'UTR intron (Christensen et al. 1992). Finally, the potato the *SUS3* and *SUS4* 5' introns are approximately 1.5kb in length (Fu and Park 1995).

Introns present in 5' UTRs have been reported to increase the expression of reporter genes in transient expression. The *UBQ10* promoter had three fold lower reporter gene activity when the intron was not present or was replaced in the equivalent constructs (Norris et al. 1993). Similarly, in rice inclusion of the 5' UTR intron of a polyubiquitin gene in the construct increased transient GUS expression by 20-fold (Sivamani et al. 2006). Removing the 5' UTR intron almost abolished

expression of *SUS3* in transgenic potato (Fu et al. 1995a) and when the leader intron of the potato *SUS4* gene was removed, there was not only a decrease in reporter expression but also the spatial localization of expression in some tissues was altered (Fu et al. 1995b). Thus, some 5'UTR introns appear to have important functions in regulating gene expression. The presence of an intron in the 5' UTR of the arginase gene was demonstrated by genomic PCR (and not by RT-PCR) when primers from the same region were used (Figure 3.2). The constructs designed to test the upstream regions of pine arginase all included the 5' UTR intron. Perhaps 5' UTR introns have both positive and negative effects on gene expression, in order to test this, a second set of constructs that lack the arginase 5'UTR intron should be produced and tested in order to determine if this is true.

The 5' UTR intron may also have a role in the longevity of the arginase mRNA. Borsuk et al. (2007) demonstrated that in *A. nidulans*, the binding of L-arginine to the 5' UTR induces structural changes in the 5' UTR, which alters the splicing of the intron in the UTR to remove either 53 or 72 nucleotides. They also demonstrated the mRNA is degraded more slowly in the presence of L-arginine. Thus, the intron and its alternative splicing affect the stability of the arginase mRNA.

Plant introns generally have a high AT content (Simpson et al. 1993). The putative 5' UTR intron we identified has a 53% AT content. This appears low when compared to those present in the polyubiquitin genes of rice (61%), and Arabidopsis (70% *UBQ3*, 68% *UBQ 10*, 66% *UBQ11*) (Sivamani and Qu 2006; Norris et al. 1993). While the 53% AT content observed in pine is low, it is within the range of plant AT content for plant introns: 42-89% AT for dicots and 31-80% for monocots

(Simpson et al. 1993). Relatively less is known about gymnosperm introns; however, some information is present. Genomic sequence from *GS1a* has been isolated, and the introns that are present show an average of 64% AT in Scots pine (Avila et al. 2002). This, while higher than the AT content of the loblolly pine intron, it is still in the lower range observed for angiosperm introns.

Consensus sequences for intron splicing, which are generally defined in all eukaryotes as G/GT at the 5' end and AG/G at the 3' end, have been identified in several conifer genes (Chiron et al. 200, Avila et al. 2002), as well as in the arginases of human, mouse and rat (Takiguchi et al. 1988, Iyer et al. 2002, Ohtake et al. 1988) and in 5' UTR intron of *A. nidulans* arginase (Borsuk et al. 2007). In contrast, the putative 5' UTR intron we identified appears to lack 5' and 3' consensus splice sites in the expected locations (Table 3.2). However, a 5' consensus splice site was detected in the putative 5' UTR, although its location is not perfectly consistent with the intron/exon boundary predicted by comparison of spliced and unspliced arginase sequences. If the deduced intron/exon boundary of the pine 5' consensus sequence of TTG/AGGTATGA were moved 2 nucleotides to **AG/GTATG**A it would be similar to the **AG/GTAAG**T consensus sequence present in most organisms (the similar nucleotides are highlighted in bold).

3.4.3 TATA and CAAT boxes

We found a TATA box consensus sequence 162bp upstream of the start of the intron (318bp upstream of the translation start codon) and a CAAT box 276bp

upstream of the intron (432bp upstream of the translation start codon) (Figure 4.3). The locations of most TATA boxes and CAAT boxes are reported relative to the transcription start site; however, neither the transcription start site nor the translation start sites have yet been experimentally defined for arginase in loblolly pine. Angiosperm TATA boxes are usually located approximately 25-30bp upstream of the transcription initiation site (Joshi 1987), although the relevance of this estimate to conifers is not certain. Given that there are 162bp between the intron and the putative TATA box it is likely that the transcription initiation site is located in this region. The *GS1a* and *GS1b* TATA boxes from Scots pine (*Pinus sylvestris* L.) are located 35bp and 38bp upstream of the transcription initiation sites; this corresponds to the TATA boxes being 215bp and 106bp upstream of the translation initiation site (Avila et al. 2002, Gómez-Maldonado et al. 2004). A TATA box from a germin-like protein (GLP) gene isolated from Caribbean pine is 78bp upstream of the translation start codon (Mathieu et al. 2003) while the TATA box for the *PR10* promoter from western white pine is 29bp upstream of the start codon (Liu et al. 2005). Thus, there appears to be a significant range for the location of the TATA box relative to the translation start site. The putative TATA boxes in arginase are approximately 318bp upstream of the first ATG which is slightly further upstream than the other TATA boxes from conifers. However, this larger distance may be attributed in part to the presence of a 150bp intron between the TATA box and the translation start site.

In all eukaryotes, the CAAT box is normally located upstream of both the transcription initiation site and the TATA box (although not all promoters contain these consensus sequences). In the genomic sequence upstream of arginase (located in

a region that is conserved between FP1 and FP2) a putative CAAT box is present a further 111bp upstream of the putative TATA box, which corresponds to 432bp upstream of the start codon. In Scots pine the *GS1a* and *GS1b* CAAT box is present 355bp and 230bp upstream of the translation start site (Gómez-Maldonado et al. 2004). Thus, the putative pine arginase CAAT box is slightly further upstream than the CAAT box of *GS1a* relative to the translation start site, but is still located in a position consistent with its expected function in transcriptional regulation.

3.4.4 Repetitive sequences

Sequencing of the genome walking products indicated the presence of two types of repetitive elements present in each of the full length products, FP1 and FP2 (Figure 3.3). BLAST analysis of each repeat showed no homology of these repetitive elements to any sequence in public databases (e value greater than 10^{-5}). Repetitive elements have also been identified in the region upstream of the *CAD* gene in loblolly pine. *CAD* catalyzes the final step in the production of monolignols the monomers of polymerized lignin (MacKay et al. 1995). The *CAD* gene is present as a single copy in loblolly pine, but several alleles of this gene encode enzymes with different mobilities when run on SDS-PAGE gels (MacKay et al. 1995). The upstream regions of four *CAD* alleles were sequenced and in the regions upstream of the *CAD* alleles, five unrelated sets of tandem sequence repeats were identified (MacKay 1996). The individual repeats were generally linked together with no more than 1bp separating each repeat and the repeats ranged in size and copy number; from 2bp to

approximately 66bp and from two to six copies. Three of the sets of repeats were greater than 20bp in length, each of these were present in two copies. All repeats were present in the same copy number and were approximately the same size in the 4 alleles tested. The repeats found in FP1 and FP2 were separated by larger regions of DNA (153 and 375bp) (Figure 3.3). In contrast to the CAD alleles, all repeats of a set were not joined directly together in the upstream region of the arginase gene; in the arginase upstream regions each of the repeats in repeat set 1 were joined together but repeats 2B and 2C were separated by 30bp in both FP1 and FP2. The repeats found in FP1 and FP2 were also not present in the same copy number; there were three copies of repeat two in FP2 but only two in FP1, however, repeat one was present in two copies in both FP1 and FP2. The repeats of FP1 and FP2 were also not all approximately the same size, which was a feature of the CAD allele repeats. Repeat 1A in FP2 was 32bp while repeat 1A in FP1 was 148bp in length. Repeat two also seemed to vary in size; it was approximately 80bp in FP1 and in one of the three copies of repeat two in FP2 but was 124bp in R2A and R2B in FP2 (Figure 3.3). In general, the repeats found in the upstream region of arginase were larger than those of in the *CAD* upstream regions. Thus, sequence repeats also exist in the upstream regions of other genes in loblolly pine; however, the repeats of arginase were not identical in structure nor arrangement to the repeats found in the CAD alleles.

3.4.5 Determination of FP1 and FP2 as alleles of a single arginase or as two genes

It was possible that the distinct FP1 and FP2 fragments we identified represented either two alleles of a single gene or two separate arginase genes. In loblolly pine, high-stringency Southern blotting of genomic DNA with the arginase cDNA suggested that arginase was present as single copy gene; however, because low stringency hybridization conditions were not tested, it is possible that more than one arginase gene is present in loblolly pine (Todd et al. 2001b). Multiple arginase genes are present in several plant species including two in tomato (Chen et al. 2004), two in Arabidopsis (Krumpelman et al. 1995, Todd et al. unpublished data), and a small gene family in soybean (Goldraij and Polacco 1999). Thus, based on a comparison to other species, it is possible that more than one arginase gene is present in loblolly pine. If multiple arginase genes were present in loblolly pine and both encoded proteins, the proteins would be considered isoenzymes. However, it is also possible that the two genome walking products we identified represented alleles of a single arginase gene. If this were the case, the protein products would be considered alloenzymes. Multiple alleles of the *CAD* gene have been identified in loblolly pine (MacKay et al. 1995). Each allele of the *CAD* gene was found to encode an alloenzyme that could be differentiated on an SDS-PAGE gel. When the arginase enzyme was purified from loblolly pine a single protein band was present (Todd et al. 2001); thus, no alloenzymes or isoenzymes appeared to exist. However, if only small differences that did not alter the mobility of the protein existed, neither alloenzymes nor isoenzymes might be distinguished. The seed used to complete the arginase protein purification and the experiments here were from an open pollinated stand of clonal trees. In order to attempt to determine whether alloenzymes or isoenzymes of

loblolly pine arginase exist, individuals from less closely related populations should be compared.

We successfully completed PCR genotyping of 37 individual megagametophytes (Figure 3.4). Because the genome of the diploid maternal tree was unknown, we tested the hypothesis that FP1 and FP2 represented two different arginase genes (i.e. isozymes). We completed chi-square tests for all combinations of the possible assumptions, which included the genotype of the maternal tree at both putative loci, the linkage status of the genes and the reason for lack of a PCR product in some reactions. In total, 16 chi-square tests were completed and the only scenario in which we did not reject the null hypothesis (i.e. that there were two genes) was that in which the maternal tree was heterozygous at two closely linked loci, and that all the blank lanes were failed reactions.

It is possible that if two arginases were present in the pine genome they would be tightly linked. *GS1a* and *GS1b* in maritime pine (*Pinus pinaster*) were found to be closely linked as no recombination was found when 68 F2 individuals were analyzed on 2D gels (Avila et al. 2000). *KN1* and *KN2,* which are sub-groups of class I knox transcription factor genes, are also located close together on the same linkage group in two *Picea* species (Guillet-Claude et al. 2004). However other sub-groups *KN3* and *KN4* are mapped separately both to each other and to the *KN1* and *KN2* groups. Thus, not all genes from a presumed duplication event are closely linked.

At present we are not able to determine conclusively whether FP1 and FP2 represent alleles of single pine arginase gene or if multiple arginases are present.

However, if FP1 and FP2 are associated with two distinct arginase genes, these genes must be linked. This could be determined by screening a second genomic library using a vector from which the inserts are easily cloned. The first genomic library was successfully screened with the 5' end of the cDNA and appeared to yield results consistent with the genome walking reactions. The cDNA could again be used to screen this second library; this would hopefully yield not only the sequences upstream of arginase, but full length genomic sequence(s) of arginase. However, as it is possible that there are two arginase genes closely linked in the genome and it would be beneficial to isolate a single clone containing both genes together if they are linked, a library capable of containing larger inserts should be considered.

3.4.6 Transient expression of FP1 and FP2

The transient expression of FP1 and FP2 was relatively unsuccessful (Figure 3.6). Todd et al. (2001b) have shown that the arginase transcript accumulates in the cotyledons of loblolly pine seedlings up to $6DIC_{30}$ old. However, when tissue from 4 DIC_{30} or $8DIC_{30}$ was bombarded with reporter fusions of FP1 or FP2, only very low levels of GUS activity were detected, which is not consistent with expected maximal transcript accumulation. There are several possible explanations for the low level of GUS expression we observed. As mentioned earlier, perhaps the intron contains negative regulatory elements. Introns in 5' UTRs have been shown to have positive influences on gene expression (Norris et al. 1993, Sivamani et al. 2006), but perhaps the opposite is also true. Alternatively, the cloning of the putative promoters may

have altered their ability to express the reporter gene. The secondary structure of the arginase 5' UTR is believed to be involved in regulating arginase protein levels in *A. nidulans* (Borsuk et al. 2007). It is possible that the presence of the vector sequence does not allow for proper formation of the secondary structure required for normal levels of arginase expression. Similarly, mutations incurred during cloning could affect the formation of the secondary structure or the induction by the metabolite. The *A. nidulans* arginase 5' UTR is bound directly by arginine which causes its secondary structure to alter (Borsuk et al. 2007). If arginine were not capable of binding, induction of arginase would not occur. Examples of a metabolite altering gene expression when bound also exist in plants. The end product of the thiamine biosynthesis pathway in Arabidopsis, thiamine pyrophosphate (TPP), binds and activates an enzyme active early in thiamine biosynthesis pathway (Sudarsan et al. 2003) thus activating its own production. The putative TPP binding domain is located in the 3' UTR. Thus, examples of metabolite regulation using the UTRs exist in plants. Perhaps in the transient expression arginine could not be bound or the required secondary structure to be produced and no expression of the reporter gene occurred. A preliminary experiment to determine if indeed the pine arginase 5' UTR intron was affected by the presence of arginine would be to complete RT-PCR using primers that span the intron under inducing and non-inducing conditions. Finally, in *A. nidulans,* alternative splicing of the intron in the 5' UTR of arginase affected the stability of the transcript (Borsuk et al. 2007). In the pine transient expression assays all constructs tested contained the intron, perhaps the splicing of the transcript was incorrect and this did not allow expression of the reporter gene.

3.5 References

Attree SM, Pomeroy MK, Fowke LC. 1992. Manipulation of conditions for the culture of somatic embryos of white spruce for improved triacylglycerol biosynthesis and dessication tolerance. Planta 187: 395-404.

Ávila C, Muñoz-Chapuli R, Plomion C, Frigerio JM, Cánovas FM. 2000: Two genes encoding distinct cytosolic glutamine synthetases are closely linked in the pine genome. FEBS Letters 477: 237-243.

Avila C, Suarez MF, Gomez-Maldonado J, Canovas FM. 2001. Spatial and temporal expression of two cytosolic glutamine synthetase genes in Scots pine: functional implications on nitrogen development during early stages of conifer development. The Plant Journal 25: 93-102.

Avila C, Cantón FR, Barnestein P, Suárez MF, Marraccini P, Rey M, Humara J, Ordás R, Cánovas FM. 2002. Structural and functional characterization of the 5' upstream region of a gluatamine synthetase gene from Scots pine. Annals of Forest Science 59: 669-673.

Bewley JD and Black M. 1985. Seeds: physiology of development and germination. Plenum Press New York, New York.

Borsuk P, Przykorska A, Blachino K, Koper M, Pawlowicz JM, Pekala M, Weglenski P. 2007. L-Arginine influences the structure and function of arginase mRNA in *Aspergillus nidulans*. Biological Chemistry 388: 135-144.

Carpita NC, Skaria A, Barnett JP, Dunlap JR. 1983. Cold stratification and growth of radicles of loblolly pine (*Pinus taeda*) embryos. Physiologia Plantarum 59: 601-606.

Chiron H, Drouet A, Lieutier F, Payer HD, Ernst D Sandermann H. 2000. Gene induction of stilbene biosynthesis in Scots pine in response to ozone treatment, wounding, and fungal infection. Plant Physiology 124: 865-872.

Chopra S, Del-favero J, Dolferus R, Jacobs M. 1992. Sucrose synthase of Arabidopsis: genomic cloning and sequence characterization. Plant Molecular Biology 18: 131-134.

Christensen AH, Sharrock RA, Quail PH. 1992. Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. Plant Molecular Biology 18: 675-689.

Ciavatta VT, Ergertsdotter U, Clapham D, von Arnold S, Cairney J. 2002. A promoter from the loblolly pine *PtNIP1;1* gene directs expression in an early-embyrogenesis and suspensor-specific fashion. Planta 215: 694-698.

Fowells HA. Silvics of forest trees of the United States. United States forest service, Division of Timber Management Research, U.S. Department of Agriculture, Washington, D.C., U.S.D.A. In: Agricultural Handbook No.271 pg 360-372.
Fu H, Kim SY, Park WD. 1995a. A potato Sus3 sucrose synthase gene contains a context-dependent 3' element and a leader intron with both positive and negative tissue-specific effects. The Plant Cell 7: 1395-1403.

Fu H, Kim SY, Park WD. 1995b. High-level tuber expression and sucrose inducibility of a potato Sus4 sucrose synthase gene require 5' and 3' flanking sequences and the leader intron. The Plant Cell 7: 1387-1394.

Gómez-Maldonado J, Crespillo R, Ávila C, Cánovas FC. 2001. Efficient preparation of maritime pine (*Pinus pinaster*) protoplasts suitable for transgene expression analysis. Plant Molecular Biology Reporter 19: 361-366.

Gómez-Maldonado J, Cánovas FC, Avila C. 2004. Molecular analysis of the 5' upstream region of a gibberelin-inducible cytosolic glutamine synthetase gene (GS1b) expressed in pine vascular tissue. Planta 218: 1036-1045.

Guillet-Claude C, Isabel N, Pelgas B, Bousquet J. 2004. The evolutionary implications o knox-I gene duplications in conifers: correlated evidence from phylogeny, gene mapping, and analysis of functional divergence. Molecular Biology and Evolution 21: 2232-2245.

Harms WR, Whitesell CD, DeBell DS. 2000. Growth and development of loblolly pine in a spacing trial planted in Hawaii. Forest Ecology and Management 126: 13-24.

Höffig KP, Moyle RL, Putterill J, Walter C. 2003. Expression analysis of four *Pinus radiata* male cone promoters in the heterologous host Arabidopsis. Planta 217: 858-867.

Iyer RK, Yoo PK, Kerm RM, Rozengut N, Tsoa R, O' Brien WE, Yu H, Grody WW, Cederbaum SD. 2002. Mouse model for human arginase deficiency. Molecular and Cellular Biology 22: 4491-4498.

Joshi CP. 1987. An inspection of the domain between putative TATA box and translation start site in 79 plant genes. Nucleic Acids Research 15: 6643-6653.

King JE and Gifford DJ. 1997. Amino acid utilization in seeds of loblolly pine during germination and early seedling growth. Plant Physiology 113:1125-1135.

Knight RD and Landweber LF. 1998 Rhyme or reason: RNA-arginine interactions and the genetic code. Chemistry and Biology 5: R215-R220.

Liu JJ, Ekramoddoullah AKM, Piggot N, Zamani A. 2005. Molecular cloning of a pathogen/wound-inducible *PR10* promoter from *Pinus monticola* and characterization in transgenic *Arabidopsis* plants. Planta 221: 159-169.

MacKay JJ, Liu W, Whetten R, Sederoff RR, O'Malley DM. 1995. Genetic analysis of cinnamyl alcohol dehydrogenase in loblolly pine: single gene inheritance, molecular characterization and evolution. Molecular and General Genetics 247: 537-545.

MacKay J. 1996. PhD thesis. A mutation in lignin biosynthesis in loblolly pine: genetic molecular and biochemical analyses. North Carolina State University, Raleigh North Carolina.

Mathieu M, Nuetelings G, Hawkins S, Grenier E, David H. 2003. Cloning of a pine germin-like protein (GLP) gene promoter and analysis of its activity in transgenic tobacco bright yellow 2 cells. Physiologia Plantarum 117: 425-434.

Mhaske VB, Chengalrayan K, Farr AL, Randall RJ. 1998. Influence of osmotica and abscisic acid on trigyceride accumulation in peanut somatic embryos. Plant Cell Reports 17: 742-746.

Nakano K, Suzuki T, Hayakawa T, Yamaya T. 2000. Organ and cellular localization of asparagine synthetase in rice plants. Plant Cell Physiology 41:874-880.

Norris SR, Meyer SE, Callis J. 1993. The intron of *Arabidopsis thaliana* polyubiquitin genes is conserved in location and is a quantitative determinant of chimeric gene expression. Plant Molecular Biology 21: 895-906.

Ohtake A, Takiguchi M, Shigeto Y, Amaya Y, Kawamoto S, Mori M. 1988. Structural organization of the gene for rat liver-type arginase. The Journal of Biological Chemistry 263: 2245-2249.

Pullman GS, Johnson S, Peter G, Cairney J, Xu, N. 2003. Improving loblolly pine somatic embryo maturation: comparison of somatic and zygotic embryo morphology, germination, and gene expression. Plant cell reports 21:747-758.

Raven PH, Evert RF, Eichorn SE. 1999. Biology of Plants 6th Edition. W.H. Freeman and Company/Worth Publishers. New York, New York, USA

Rotman BR and Papermaster RD. 1966. Membrane properties of living cells as studied by enzymatic hydrolysis of fluorogenic esters. Proceedings of the National Academy of Sciences 55: 134-141.

Samson ML. 2000. *Drosophilia* arginase is produced from a third nonvital gene that contains the *elav* locus within its third intron. The Journal of Biological Chemistry 275(6): 31107-31114.

Schultz RP. 1999. Loblolly-the pine for the twenty-first century. New Forests 17: 71-88.

Shaw JR, Ferl RJ, Baier J, St. Clair D, Carson C, McCarty DR, Hannah LC. 1994. Structural features of the maize *sus*1 gene and protein. Plant Physiology 106: 1659-1665.

Simpson GG and Filipowicz W. 1996. Splicing of precursors to mRNA in higher plants: mechanism, regulation and sub-nuclear organization of the spliceosomal machinery. Plant Molecular Biology 32: 1-41.

Simpson CG, Leader DJ, Brown JWS. 1993. Characteristics of plant pre-mRNA introns. In Plant Molecular Biology Labfax ed. RRD Croy 6: 183-251.

Sivamani E and Qu R. 2006. Expression enhancement of a rice polyubiquitin gene promoter. Plant Molecular Biology 60: 225-239.

Stone SL and Gifford DJ. 1997. Structural and biochemical changes in loblolly pine (*Pinus taeda L.*) seeds during germination and early seedling growth. I. Storage protein reserves. International Journal of Plant Science 158:727-737.

Stone SL and Gifford DJ. 1999. Structural and biochemical changes in loblolly pine (*Pinus taeda L.*) seeds during germination and early seedling growth. II. Storage triacylglycerols and carbohydrates. International Journal of Plant Science 160: 663-671.

Takiguchi M, Haraguchi Y, Mori M. Human liver-type arginase gene: structure of the gene and analysis of the promoter region. Nucleic Acids Research 16: 8789-8802.

Tillman-Sutela E and Kauppi A. 2000. Structures contributing to the completion of conifer seed germination. Trees 14: 191-197.

Todd CD, Cooke JEK, Gifford DJ. 2001a. Purification and properties of *Pinus taeda* arginase from germinated seedlings. Plant Physiology and Biochemistry 39: 1037-1045.

Todd CD, Cooke JEK, Mullen RT, Gifford DJ. 2001b. Regulation of loblolly pine (*Pinus taeda L.*) arginase in developing seedling tissue during germination and post-germinative growth. Plant Molecular Biology 45: 555- 565.

Todd CD, Gifford DJ. 2002. The role of the megagametophyte in maintaining loblolly pine (*Pinus taeda* L.) seedling arginase gene expression *in vitro*. Planta 215: 110-118.

Todd CD, Gifford DJ. 2003. Loblolly pine arginase responds to arginine in vitro. Planta 217: 610-615.

Wang J, Jiang J, Oard JH. 2000. Structure, expression and promoter activity of two polyubiqutin genes from rice (*Oryza sativa* L.). Plant Science 156: 201-211.

Wei H, Wang ML, Moore PH, Albert HH. 2003. Comparative analysis of two sugarcane polyubiquitin promoters and flanking sequncences in transgenic plants. Journal of Plant Physiology 160: 1241-1251.

Yu WP, Wang AY, Juang RH, Sung HY, Su JC. 1992. Isolation and sequences of rice sucrose synthase cDNA and genomic DNA isolation. Plant Molecular Biology 18: 139-142.

Chapter 4: Analysis of arginase gene expression in *Arabidopsis thaliana*

4.1 Introduction

Arginine (Arg) is an important amino acid in plants. It is a component of proteins, a medium for the transport and storage of nitrogen (N) (Micallef and Shelp 1989) and precursor for the synthesis of nitric oxide (NO), polyamines and proline (Pro) (Jenkison et al. 1996). Arg is hydrolyzed by the enzyme arginase (L-Arg ureahydrolase, or amidinohydrolyase, EC 3.5.3.1), which produces ornithine and urea. In plants, urea is degraded by urease (Zonia et al. 1995) while ornithine can be utilized in the urea cycle, or as a precursor for the synthesis of both polyamines and proline.

Arginase cDNAs have been cloned from higher plant species, including loblolly pine (*Pinus taeda* L.), soybean (*Glycine max*), tomato (*Lycopersicon esculentum*), and Arabidopsis (*Arabidopsis thaliana*). In soybean, a small family of arginase genes was detected (Goldraij and Polacco 1999) while both tomato and Arabidopsis appear to have two arginase genes. These genes encode functional arginases in both tomato and Arabidopsis (Chen et al. 2004, Krumpelman et al. 1995, Todd et al. unpublished data). In Arabidopsis the first arginase cDNA was identified when it complemented a yeast (*Saccharomyces cerevisiae*) arginase deficient (*car1*) mutant (Krumpelman et al. 1995). The yeast mutant (*car1*, *dur1*, *ura3*-52) was unable to grow on either Arg or urea as a sole N source, but growth on plates with Arg as the sole N source was observed with complementation of an *Arabidopsis* arginase cDNA.

Sequence homology to other arginases as well as the ability of the complemented mutant to accumulate urea in the presence of Arg confirmed its identification as an arginase. This was named *ARGAH1* (At4g08900) in order to differentiate it from *ALTERED RESPONSE TO GRAVITY* (*ARG1*) (C. Todd unpublished data). After the Arabidopsis genome was fully sequenced a second putative arginase (At4g08870) was identified by sequence homology and is referred to as *ARGAH2*. Like *ARGAH1*, the open reading frame (ORF) of *ARGAH2* was able to complement the *car1* mutation in yeast, allowing growth on Arg as the sole N source (C. Todd unpublished data). Todd et al. (unpublished data) have also shown spatial differences in expression of *ARGAH*1 and *ARGAH2* using semi-quantitative RT-PCR (Figure 1). *ARGAH*1 transcripts are expressed in seedlings, as well as in leaves, roots, stems, flowers and siliques whereas *ARGAH2* transcripts are highly abundant in seedlings, leaves, flowers and siliques, but its transcripts are not detected in either the roots or stem. Thus, while *ARGAH2* transcripts are produced in greater abundance in some tissues they are not expressed in all tissues tested like *ARGAH1*.

Physiological roles of arginase have been described in several species including in conifers where arginase is commonly associated with mobilization of seed storage reserves after the completion of germination during early seedling growth. Similarly in soybean, the majority of the storage reserves are located in the cotyledons, where Arg is the predominant form of N stored in amino acids (Micallef and Shelp 1989b), which are degraded during seedling development. Arginase activity has been described during early seedling growth in loblolly pine (King and Gifford (1997) and pea (De Ruiter and Kollöffel 1983). Arginase activity has also

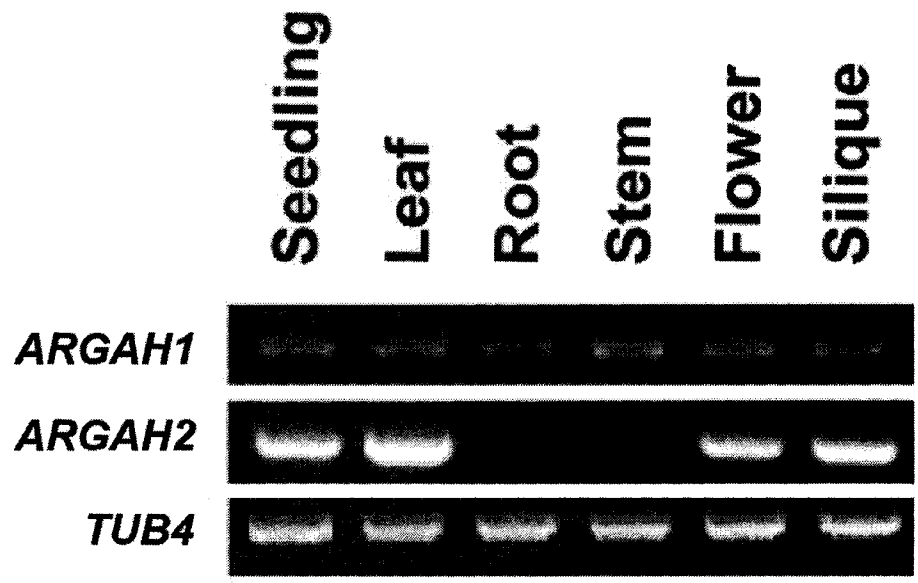


Figure 4.1: Patterns of ARGAH1 and ARGAH2 trancript accumulation in several wild-type Arabidopsis tissues. Adapted from Todd et al. (unpublished data). Seedlings were grown on water and then eiher harvested or transferred to soil until maturity where other tissues were harvested. Tubulin (TUB4) was used as a control.

been detected during early seedling growth in Arabidopsis. Using homozygous T-DNA insertion lines, Todd et al. (unpublished data) have shown that *ARGAH1* enzyme activity is present at a low and constant level throughout germination and for several days of early seedling growth. Seedlings with a functional *ARGAH2* have arginase activity that increases significantly after $4DIC_{22}$ (4 days in culture at 22°C) and reaches its maximum after $8DIC_{22}$. This, in combination with the 2 different patterns of spatial expression suggests that the arginase gene products may have different roles in Arabidopsis. Arginase activity has also been described in unpollinated ovaries and developing fruit (Alabadí et al. 1996) and, similar to Arabidopsis, the two arginases (*LeARG1* and *LeARG2*) present in tomato are differentially expressed in several tissues (Chen et al. 2004). Chen et al. (2004) also demonstrated that in tomato one arginase is induced in response to wounding, jasmonic acid, and the pathogen *Pseudomonas syringae*. This work collectively shows that arginase is involved in several distinct physiological processes in plants and that the two arginases present in some species may have distinct roles.

Ornithine, one of the products of the arginase reaction, is a precursor for the synthesis of both polyamines and Pro in plants. Pro can be synthesized from either glutamate or ornithine, each using a separate pathway (Buchanan et al. 2000). Polyamines are also synthesized via two pathways in plants: the first directly from arginine via arginine decarboxylase and the second from ornithine via ornithine decarboxylase (Bais and Ravishankar 2002). Thus, in some circumstances, arginase can provide ornithine to the plant for the production of proline and polyamines during its lifecycle.

Arginase is also induced in response to wounding. Chen et al. (2004) demonstrated that *LeARG2* but not *LeARG1* was induced in leaves in response to wounding. *LeARG2* in tomato is also induced by the pathogen *P. syringae*, which causes bacterial speck disease (Chen et al. 2004). In some preliminary work Todd et al. (unpublished observations) observed that arginase transcripts also accumulate in response to wounding in Arabidopsis.

In addition to serving as a substrate for arginase, Arg can also be a precursor for nitric oxide (NO) via nitric oxide synthase (NOS) in animals and this pathway may be present in plants (Guo et al. 2003); thus arginase and NOS may compete for Arg as a substrate. NO is a signaling molecule which has been implicated in plant growth, fertility, stomatal movements, hormone signaling and defense in plants (Guo et al. 2003). Several lines of evidence show that NO is involved in plant defense signaling. Tobacco plants resistant to the tobacco mosaic virus had increased NOS activity when exposed to the virus but susceptible tobacco plants did not (Durner et al. 1998). Exposure of either tobacco plants or cell suspensions to mammalian NOS triggered the expression of defense related genes and finally inhibitors of NOS in Arabidopsis compromised the hypersensitive response of plants exposed to *P. syringae* (Delledonne et al. 1998). In animals NO has roles in vasodilation, neurotransmission, and immune responses (Wu and Morris 1998). Inhibition of arginase resulted in increased NO production and increased parasite killing when mice macrophages were infected with trypanosomes (Duleu et al. 2004). Duleu et al. (2004) also demonstrated that trypanosomes directly induced both arginases possibly as a strategy to escape the NO response of the immune system. Thus, arginase may

play several roles in defense: competing with NO production and in the production of polyamines.

In plants arginase has been described in a number of species and some its roles including mobilization of storage reserves in angiosperms and gymnosperms have been investigated. In this work we investigated the role of arginase in the model plant Arabidopsis. We produced transgenic plants containing the upstream region of each arginase and GUS staining of these transgenic plants at several different developmental stages was completed. In addition, we investigated the accumulation of each arginase transcript in pollen and in response to wounding.

4.2. Materials and methods

4.2.1 Cloning of promoter:reporter fusions

Genomic DNA was isolated from Arabidopsis using a modified Doyle and Doyle (1987) procedure. Two to six leaves from mature Arabidopsis plants (cultivar Columbia) were ground under liquid N_2 and then incubated with 7mL of extraction buffer (1M Tris-HCl pH 8.0, 0.25M EDTA, 1.5M NaCl, 5% (w/v) CTAB, 0.2% β-Mercaptoethanol with 0.5g PVP-40 added before use) at 65°C for 30-60 minutes. An equal volume of chloroform-isoamyl alcohol (24:1) was added and the mixture was centrifuged at 10,000g for 20 minutes. The aqueous phase was removed and the nucleic acids were precipitated with cold isopropanol (2/3 volume). The solution was centrifuged at 10,000g for 15 minutes and the pellet was washed for 20 minutes with wash buffer (76% ethanol, 10 mM NH_4OAc). The tube was once again centrifuged, and the pellet was air dried and resuspended in 1mL TE buffer pH8.0. RNase A was added to a final concentration of 10μg/mL and the solution was incubated at 37°C for 30 minutes. The DNA was again precipitated, air dried, and resuspended in TE buffer pH 8.0.

PCR was carried out using 100-150ng of Arabidopsis genomic DNA. All *ARAGAH1* products were amplified using AT Reverse 4 primer and AT Fwd 4 (*ARGAH1*-A) or AT Fwd 6 (*ARGAH1*-B) (Table 4.1). *ARGAH1*-A was amplified using the following cycling conditions: 94°C 3 minutes, 45 cycles of: 94°C 1 minutes, 48°C 1 minute, 72°C 1 minute and a final extension of 72°C for 10 minutes.

ARGAH1-B was amplified using the above protocol with the following changes: annealing temperature 50°C and an extension of 4 minutes with a temperature of 72°C. All *ARGAH2* products were amplified using AT8870 Rev and either AT8870 Fwd 1 (*ARGAH2*-A) or AT8870 Fwd 4 (*ARGAH2*-B) (Table 4.1), with the following modifications to the PCR conditions described above. For *ARGAH2*-A and *ARGAH2*-B the annealing temperature was 45°C and the extension was 3 minutes at 72°C. All PCR products were cloned into the pCR4TOPO vector for sequencing and then sub-cloned into pCAMBIA 1303.

4.2.2 Plant transformation

Competent *Agrobacterium tumefaciens* (strain GV3101) were prepared as described by Weigel and Glazebrook (2002). An overnight culture (1mL) was used to inoculate 200mL of LB. This was cultured overnight (at 37°C shaking 200rpm) and the cells were harvested the next morning by centrifugation at 5000rpm for 10 minutes at room temperature. The pellet was washed with sterile TE pH 8.0, resuspended in $1/10^{th}$ the original volume of TE and frozen in liquid N_2 and stored at -80°C in 250µL aliquots. Competent Agrobacterium were thawed on ice, 10µL of plasmid DNA was added and the mixture was incubated on ice for 5 minutes then transferred to liquid N_2 for 5 minutes. The tube was then placed in a 37°C water bath for 5 minutes, 1mL of LB was added and placed on a shaker at room temperature for 2-4 hours. The cells were plated on LB plates containing 50µg/µL of kanamycin and 25-50µg/µL of gentamycin and incubated at 28°C for 2 days. Positive clones were

identified through PCR using primers pCAM 1303 Fwd and pCAM 1303 Rev (Table 4.1). The cycling conditions were 95°C for 5 minutes, followed by 30-45 cycles of: 94°C 30 seconds, 55°C 30 seconds, 72°C 90 seconds-3 minutes, followed by a final extension of 72°C for 10 minutes.

Arabidopsis thaliana was transformed according to the method described by Weigel and Glazebrook (2002). Briefly, an 5mL *Agrobacterium* culture (5mL) was incubated at 28°C for 2 days. The culture was centrifuged at 4000 rpm for 10 minutes at room temperature and resuspended in 1mL of infiltration medium (1/2 MS, 5% sucrose, 0.5μL Silwet L-77 was added before use). The solution was applied by pipet to unopened flowers every 4 days for 16 days. To screen for transformed seedlings, the seeds were surface sterilized and incubated at 4°C for 2 days and then screened on ½ MS media with 25μg/mL of hygromycin for approximately 21 days (room temperature, constant light). Seedlings that exhibited resistance to hygromycin were transferred to soil and grown (at 22°C, 16 hours light) to seed.

To confirm the presence of a transgene within putative transformants, we surface sterilized seeds, incubated them at 4°C for 2 days and then germinated them (22°C, 16 hours light) on ½ MS with 25μg/mL of hygromycin for 14 days. Individual seedlings were harvested and frozen at -80°C until use. Genomic DNA was isolated using a modified protocol from Dellaporta et al. (1983). Seedlings were ground in 125μL of extraction buffer (100mM Tris-Cl pH 8.0, 50mM EDTA, 500mM NaCl with 10mM β–Mercaptoethanol added immediately before use), 10μl of RNase A (10mg/mL) and 10μL of 20% SDS were added. The mixture was incubated at 65°C for 10mins. KOAc was added (50μL of 5M) and the tube was then incubated on wet

ice for 20 minutes and centrifuged for 15 minutes at 8000rpm at 4°C. The supernatant was added to 150μL of isopropanol in a fresh tube and incubated at -20°C for an hour. The mixture was centrifuged for 20 minutes at 12,500 rpm and the supernatant removed by aspiration. The pellet was washed with 70% EtOH, air dried, and resuspended in 40μL of TE pH 8.0.

The DNA was quantified using a Nanodrop spectrophotometer (ND 1000) and approximately 250μg of genomic DNA was used as the template for a PCR reaction. Primers 1303 forward and reverse (Table 4.1) were used to amplify the insert for each line of T_2 and T_3 seedlings. Cycling conditions were as follows: 94°C for 1 minute followed by 30 cycles of 94°C for 1 minute, 1 minute at 50°C and 3 minutes at 72°C, a final extension of 72°C for 10 minutes. PCR products were run on a 1% agarose gel and compared to expected fragment sizes.

4.2.3 GUS staining

Seedlings were isolated (and washed in water if grown in soil) and placed in cold 90% acetone for 10 minutes. The seedlings were then washed with 50mM phosphate buffer (Na_2HPO_4-$7H_20$ pH 7.2) before the staining solution was added (1mM X-gluc, 50mM sodium phosphate, 0.5mM $K_3Fe(CN)_6$, 0.5mM $K_4Fe(CN)_6$). Seedlings were stained 12-24 hours at 37°C in the dark. The staining reaction was stopped at the first appearance of blue colour by the addition of 70% EtOH. After 12-72hrs the 70% EtOH was replaced with 95% EtOH and the seedlings were stored at room temperature.

4.2.4 Wounding of Arabidopsis plants

Seeds were surface sterilized, incubated at 4°C for 2 days and then germinated on ½ MS media with 25ug/mL of hygromycin for approximately 21 days (at 22°C, 16 hours light). Seedlings exhibiting resistance to hygromycin were transferred to soil and grown again at 22°C with 16 hours of light. Plants were grown to the rosette stage and were wounded with forceps around the leaf margins, not cutting off the mid-vein. Control seedlings were not wounded. Plants were covered with a clear plastic cover and left for 8 hours at (22°C) in growth chamber. Leaves were then harvested and either GUS stained as above or frozen under liquid nitrogen and stored at -80°C until use.

4.2.5 Quantitative PCR

Pollen was isolated from wild-type Arabidopsis (cultivar Columbia) plants using a modified version of the procedure described by Honys and Twell (2003). Inflorescences from approximately 300 plants (grown at 22°C 16 hours light) were harvested and mixed with 150mL of cold 0.3M mannitol and shaken for 1 minute. The pollen suspension was filtered first through a sterile sieve containing 250μM mesh followed by a sieve containing 75μM mesh. The pollen was concentrated by centrifugation (1600 rpm, 5 minutes, 4°C) using a Sorvall RT 7 centrifuge and RTH-

750 rotor; the pellet was then stored at -80°C. Leaf tissue from the same plants was also harvested and stored at -80°C

RNA was isolated using the Qiagen RNeasy Plant Mini kit. Tissue was ground under liquid N_2 using a mortar and pestle and mixed with 1mL of buffer RLT containing β –mercaptoethanol. The liquid was spun through a QIAshredder spin column and an RNeasy mini column, and then washed with buffer (RW1 and RPE) before air drying. The RNA was then resuspended in 30μL of ddH_2O. The RNA was then DNase treated (DNA-free (Ambion)). Briefly, we added buffer (0.1volume) and rDNase I (1μL, 2 units) to the RNA and incubated at 37°C for 30 minutes. DNase Inactivation Reagent was then added (1volume) and incubated at room temperature for 2 minutes before centrifuging at 10,000 x g and then isolating the supernatant. The RNA was then quantified using a NanoDrop spectrophotometer (ND 1000).

cDNA was synthesized using RevertAid H Minus First Strand cDNA Synthesis Kit (Fermentas). 750ng of total RNA was mixed with 0.5μg of oligo$(dT)_{12-18}$ primer (Invitrogen) . The mixture was then incubated at 70°C for 5 minutes and then chilled on ice before adding 4μL of 5X reaction buffer (250 mM Tris-HCl (pH 8.3 at 25°C), 250 mM KCl, 20 mM $MgCl_2$, 50 mM DTT), 1μL of RNase Out (Invitrogen) and 2μL of 10mM dNTP mix. The mixture was then incubated at 37°C for 5 minutes before the addition of 1μL (200 units) of RevertAid H Minus M-MuLV reverse transcriptase. The mixture was incubated at 42°C for 60 minutes and then stopped by heating at 70°C for 10 minutes and finally cooled on ice.

Quantitative real-time PCR was performed in an Applied Biosystems 7500 Fast Real-Time PCR System. Primers were tested against a range of template

dilutions (1:4 to 1:16384) to ensure qPCR efficiency was not proportional to template quantity. For quantitative PCR reactions, 2.5µl of a 1:4 dilution of the reverse-transcription reaction was used in a total volume of 10µl with 0.4µM of each forward and reverse gene-specific primer (Table 4.1), 0.2mM dNTPs, 0.25X SYBR Green, 1X ROX and 0.075U Platinum Taq (Invitrogen). Reactions were carried out at 60°C. Threshold cycles (C_T) were determined using 7500 Fast Software. C_T values were calculated using Ubiquitin 10 (Ubq 10) as an endogenous control (Ubq set #2, Table 4.1) and $\Delta\Delta C_T$ values were generated using the leaf tissue as a reference. Data was expressed quantity relative to leaf ($2^{-\Delta\Delta C_T}$) (Livak and Schmittgen 2001). The standard deviation was calculated for each sample and 1 standard deviation is shown on the graph.

4.2.6 RNA extraction, cDNA synthesis and RT-PCR

RNA was isolated from 1-2 leaves using the Qiagen RNeasy Plant Mini kit (as described above) and cDNA was synthesized from 250ng of total RNA using the RevertAid H Minus First Strand cDNA Synthesis Kit (Fermentas) according to the manufacturer's instructions (as above).

RT-PCR was performed by adding 2µL of cDNA to 1µL of 10mM dNTPs, with 1µL of each 10µM primer and 5µL of 10X PCR buffer (100mM Tris-HCl, 500mM KCl, 15mM $MgCl_2$). *ARGAH1* was amplified using Arg 1 Fwd and Arg 1 Rev primers; similarly *ARGAH2* was amplified using Arg 2 Fwd and Arg 2 Rev (Table 4.1). The cycling conditions were as follows: 94°C for 1 minute, 28 cycles of

94°C for 1 minute, 62°C for 1 minute, 72°C for 90 seconds, followed by a final extension of 72°C for 10 minutes. Tubulin was used as a positive control for each tissue and treatment; this was amplified using the Tub 5' and Tub 3' primers (Table 4.1). The cycling conditions were as follows: 94°C for 1 minute, 30 cycles of 94°C for 1 minute, 55°C for 1 minute, 72°C for 90 seconds, followed by a final extension of 72°C for 10 minutes. The reactions were then visualized on a 1% TAE agarose gel.

Table 4.1: Sequences of Arabidopsis PCR primers used

Name:	Primer Sequence (5' to 3')
AT Fwd 4	AGGATCCCAATCTAACCAATCATGTCCAAAAGTCGTATC
AT Fwd 6	AGGATCCTGCGGTAGGTGTGATGACTTTAGCTTTAGGT
AT Rev 4	AACCATGGATCTGAAATTGGATCGGAACTTGAAGATCAA
AT8870 Fwd 1	GGATCCACGACTATAAAAATTTTGAATACACAAATACCA
AT 8870 Fwd 4	GGATCCTTTTGAAGAGTGTTTTGTTAGTGGTTGA
AT8870 Rev	CATGGATCTGATTTGTAGTAAGTCAGTGAGGTATAAGTT
pCAM 1303 Fwd	TGTGGAATTGTGAGCGGATA
pCAM 1303 Rev	TGCCCAACCTTTCGGTATAA
Arg 1 qRT Fwd	TGCACGATGAAACATAACAAAAAAC
Arg 1 qRT Rev	ATGGTTGCAGCTAAGCTTGTTAGA
Arg 2 qRT Fwd	AGAGCTAGCCGCAAAAATGTCA
Arg 2 qRT Rev	ACCATGTCTATGAGACCACACTTATTG
Ubquitin10 Fwd 2	GGCCTTGTATAATCCCTGATGAA
Ubquitin 10 Rev 2	AGAAGTTCGACTTGTCATTAGAAAGAAA
Arg 1 Fwd	GAGTGGCCGAAACAGAGATT
Arg 1 Rev	AGCCAGACTTATTGCATTCAAGAAC
Arg 2 Fwd	ATGTGGAAGATTGGGCAG
Arg 2 Rev	CCCAATTGATATTAAGCTTATTCGATACC
Tub 5'	TTGCGTTCTTCGTTTCCCTGG
Tub 3'	GAGGGTGCCATTGACAACATC

4.3 Results

4.3.1 Transgenic Arabidopsis GUS staining

Two arginase genes have been previously identified in the Arabidopsis genome: *ARGAH1* (At4g08900) and *ARGAH2* (At4g08870). To identify regions of DNA that regulate the transcription of these genes, we isolated two genomic fragments for each arginase gene immediately upstream of their respective start codons. The shorter genomic fragments (*ARGAH1*-A, *ARGAH2*-A) were approximately 250bp in length and the longer fragments (*ARGAH1*-B, *ARGAH2*-B) were approximately 1kb in length (Figure 4.2). These fragments were ligated into the pCAMBIA 1303 vector containing a β-glucuronidase (GUS) reporter gene to examine the expression pattern conferred by these four genomic fragments.

Because arginase activity has been reported in $7DIC_{22}$ seedlings (7 days in culture at 22°C; C. Todd, unpublished data), we examined GUS expression at this stage in T_3 generation seedlings containing our reporter constructs. The plants were grown on water, ½ MS or ½ MS + 1% sucrose in order to determine if the availability of nutrients, especially carbon and nitrogen, altered the observed staining pattern. Seedlings grown on water had small cotyledons, a relatively short root, and no visible leaves (Figure 4.3A). Seedlings grown on ½ MS medium were larger in size relative to seedlings grown on water and had small leaves, and roots that were longer than the water grown seedlings. The seedlings grown on ½ MS + sucrose medium had leaves

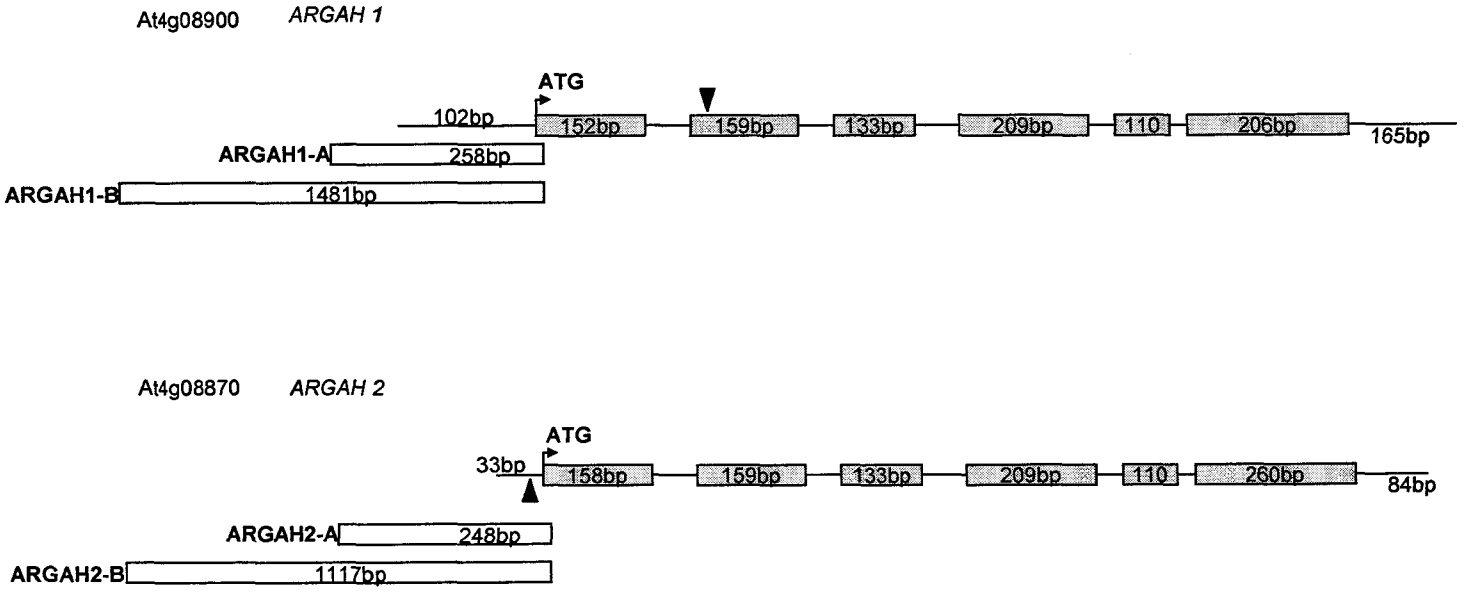


Figure 4.2: The Arabidopsis arginase genes. Predicted exons are indicated by the boxes and the intron sequences by the lines based on sequence information from TAIR. The predicted size of the 5' and 3' UTRs are labeled by their size at their respective ends. The location of the T-DNA insertion for each line (SALK_057987 and SAIL_181C11) is marked by the black triangle. Upstream fragments include the start codon and the length of the upstream region is indicated. Upstream regions were of different sizes due to the inability to design primers that amplified fragments of the same size.

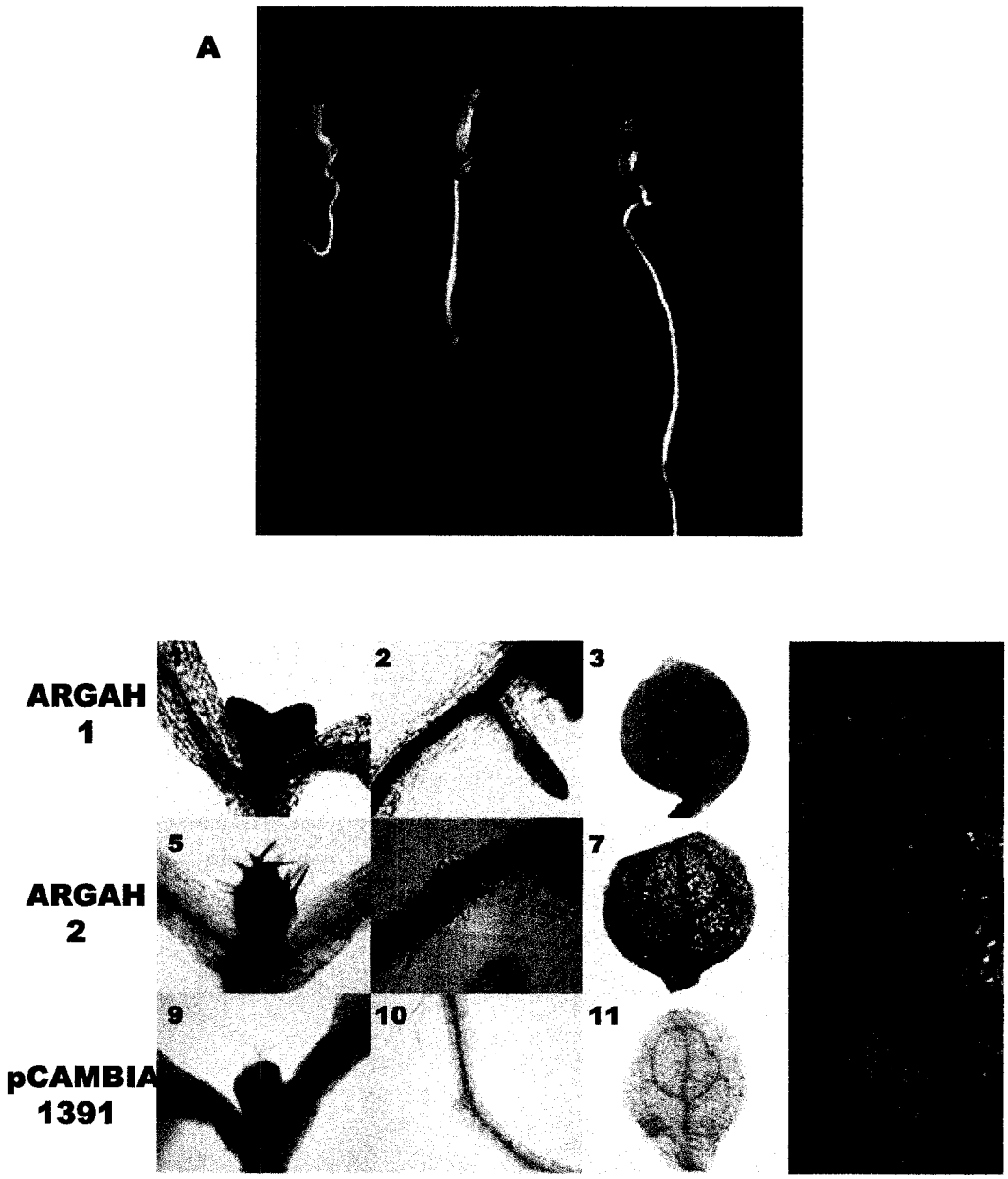


Figure 4.3: GUS staining of 7 day old Arabidopsis seedlings containing either an approximately 1kb upstream fragment of ARGAH1 or ARGAH2, or the pCAMBIA 1391 vector. A. Seedlings grown on water, 1/2 MS or 1/2 MS + 1% sucrose. B. GUS staining of seedlings grown on MS+ sucrose. The staining patterns observed were similiar for the other medias tested and for the smaller promoter fragments. Panels 1,5 and 9 are Arabidopsis seedling leaves (10X magnification). Panels 2,6 and 10 are primary roots (10X magnification). Panels 3,7 and 11 are cotyledons (10X magnification) as are panels 4,8 and 12 (20X magnification).

that were further developed but not fully expanded and longer roots compared to either the water or ½ MS seedlings. GUS expression was observed throughout the young leaves (if present) with the darkest staining observed at the leaf tip. Staining was also observed in the vasculature of the roots and in the vasculature of the cotyledons (Figure 4.3B). No GUS expression was detected in the hypocotyls of any line on any media. The staining pattern was similar for seedlings grown on all three types of media with the staining pattern becoming more intense as the richness of the media increased. GUS staining patterns were similar for all fragments tested (Figure 4.3B), indicating that the smallest promoter fragments (*ARGAH 1*-A and *ARGAH 2*-A) contained sufficient cis-regulatory elements to direct the GUS expression pattern we described above. No GUS staining was present in any of the lines of control seedlings, which were transformed with pCAMBIA 1391 (a promoter-less GUS construct similar to pCAMBIA 1303).

The GUS staining pattern of the four sets of transgenic lines at later developmental stages were examined next. All *ARGAH1* and *ARGAH2* GUS rosette-stage plants (before bolting) had staining in young leaves but not in the older leaves or cotyledons (Figure 4.4 panels A and D). This indicated that *ARGAH1-A* and *ARGAH2-A* were sufficient to confer this pattern of expression. After bolting, young leaves of shoots emerging from the axils of both rosette and cauline leaves also stained consistently in *ARGAH1* and *ARGAH2* lines (Figure 4.4). No staining was observed in comparable tissues or stages of negative control, pCAMBIA 1391 plants.

In flowers, we observed distinct reporter gene expression patterns for several of the transgenic lines (Figure 4.5). Plants bearing the smallest *ARGAH1* fragment

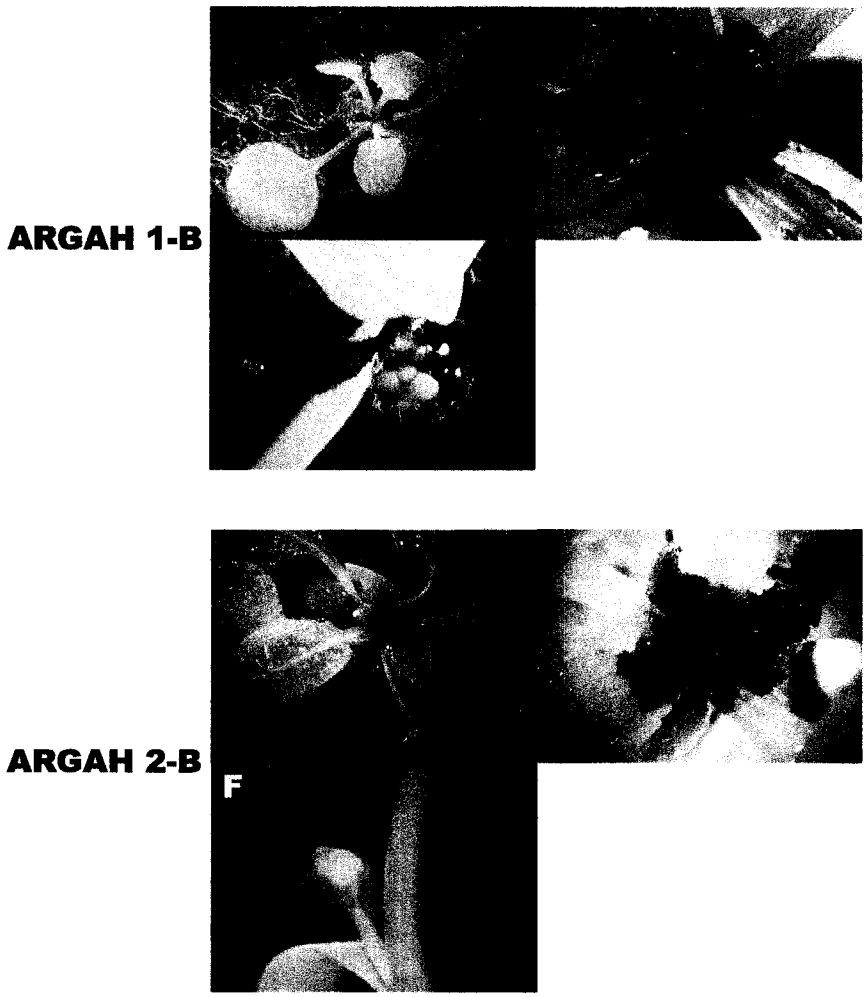


Figure 4.4: GUS staining of rosettes and young leaves containing an approximately 1kb promoter fragment during development. A, D. GUS stained rosette stage plants. B,E. Young leaves emerging from the axils of the rosette exhibit staining in ARGAH1 and ARGAH2 lines. C,F. Young leaves from the axils of cauline leaves exhibit staining in ARGAH1 and ARGAH2 lines.

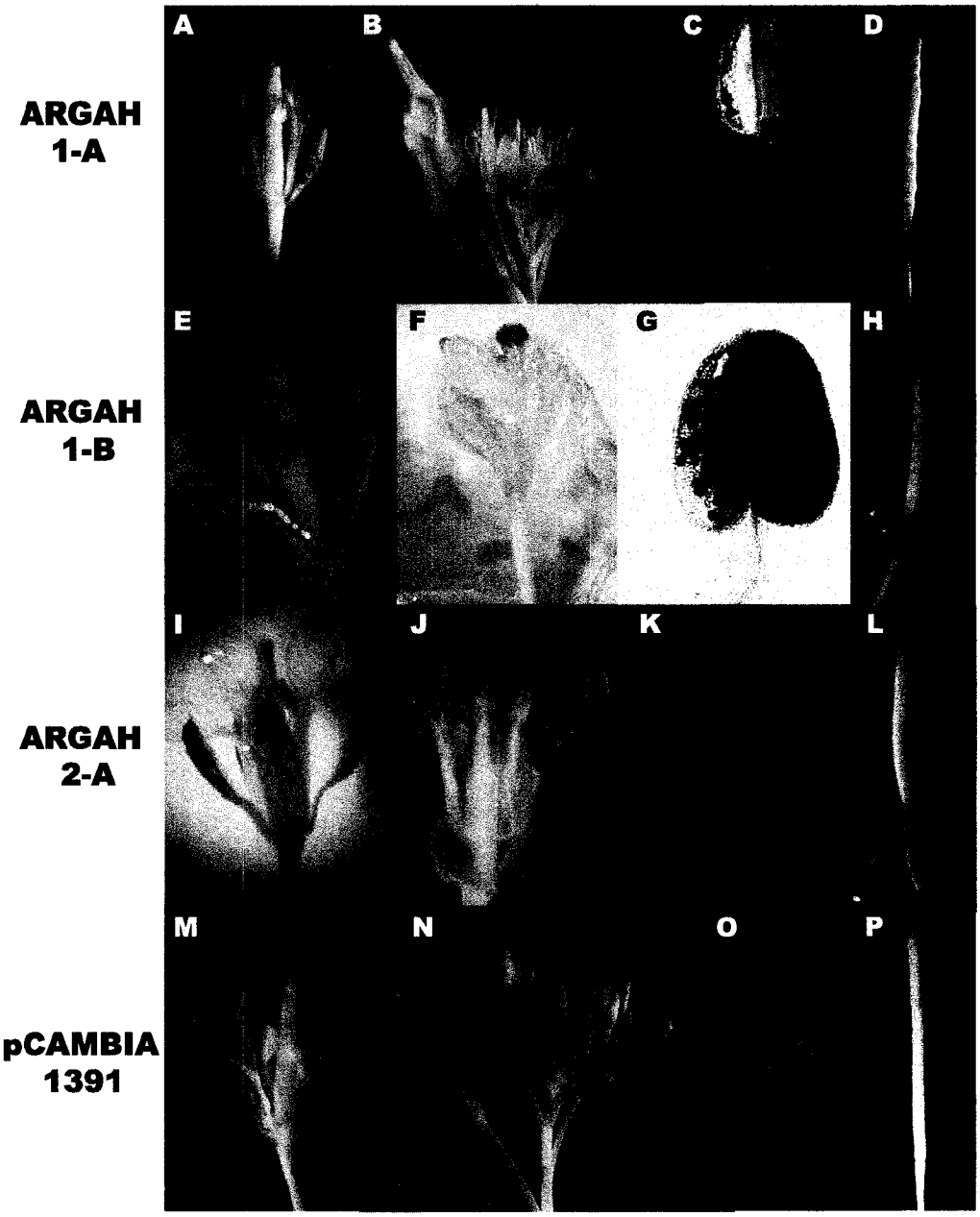


Figure 4.5: GUS staining of flowers containing a 250bp (ARGAH1-A) or approximately 1kb (ARGAH 1-B and ARGAH 2-B) promoter fragment or the pCAMBIA 1391 vector containing a promoter-less GUS gene. ARGAH2-B are not showns as the patterns of staining are similiar to ARGAH2-A.

(*ARGAH1-A*) exhibited faint staining in the pedicel, peduncle and in the abscission zone. *ARGAH1-B* lines stained in the same manner but also appeared to stain within the anthers. Upon closer inspection of the anthers it became apparent that GUS staining was localized to pollen inside the anthers (Figure 4.5); this was not observed in any of the four independent *ARGAH2-B* staining lines nor in any of the six *ARGAH1-A* lines that stained elsewhere in the flowers. GUS staining in pollen was also not observed in any of the *ARGAH2*-GUS or pCAMBIA 1391 lines tested. Thus, it appeared that the larger, *ARGAH1-B* fragment contained cis-elements sufficient to direct expression in pollen.

Both of the *ARGAH2* fragments (i.e. *ARGAH2*-A and *ARGAH2*-B) had similar GUS expression patterns. In these lines, we detected staining in the vasculature of sepals. Within the majority of lines (5/6 *ARGAH2*-A lines) GUS staining was also detected in the style, uniformly distributed beneath the stigma (Figure 4.5, panel I). Of the six *ARGAH2-A* lines one showed a variant on this pattern, with some staining observed in the ovary, just below the style, and in the connective of the anther, as well as the vasculature of the sepals and filament (Figure 4.5, panel J). Staining of the abscission zone of siliques was found in all arginase promoter-GUS constructs tested, however, this staining was not observed in the siliques of any pCAMBIA 1391 plants.

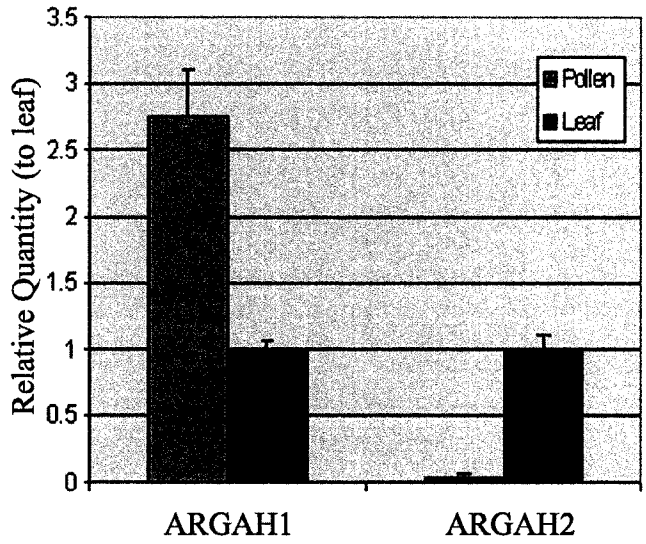


Figure 4.6: Relative quantitative expression of ARGAH1 and ARGAH 2 in pollen. Expression was calculated using Ubquitin 10 as an endogenous control and expressed relative to leaf expression (2 exp (-ddCT)). Error bars represent the standard deviation.

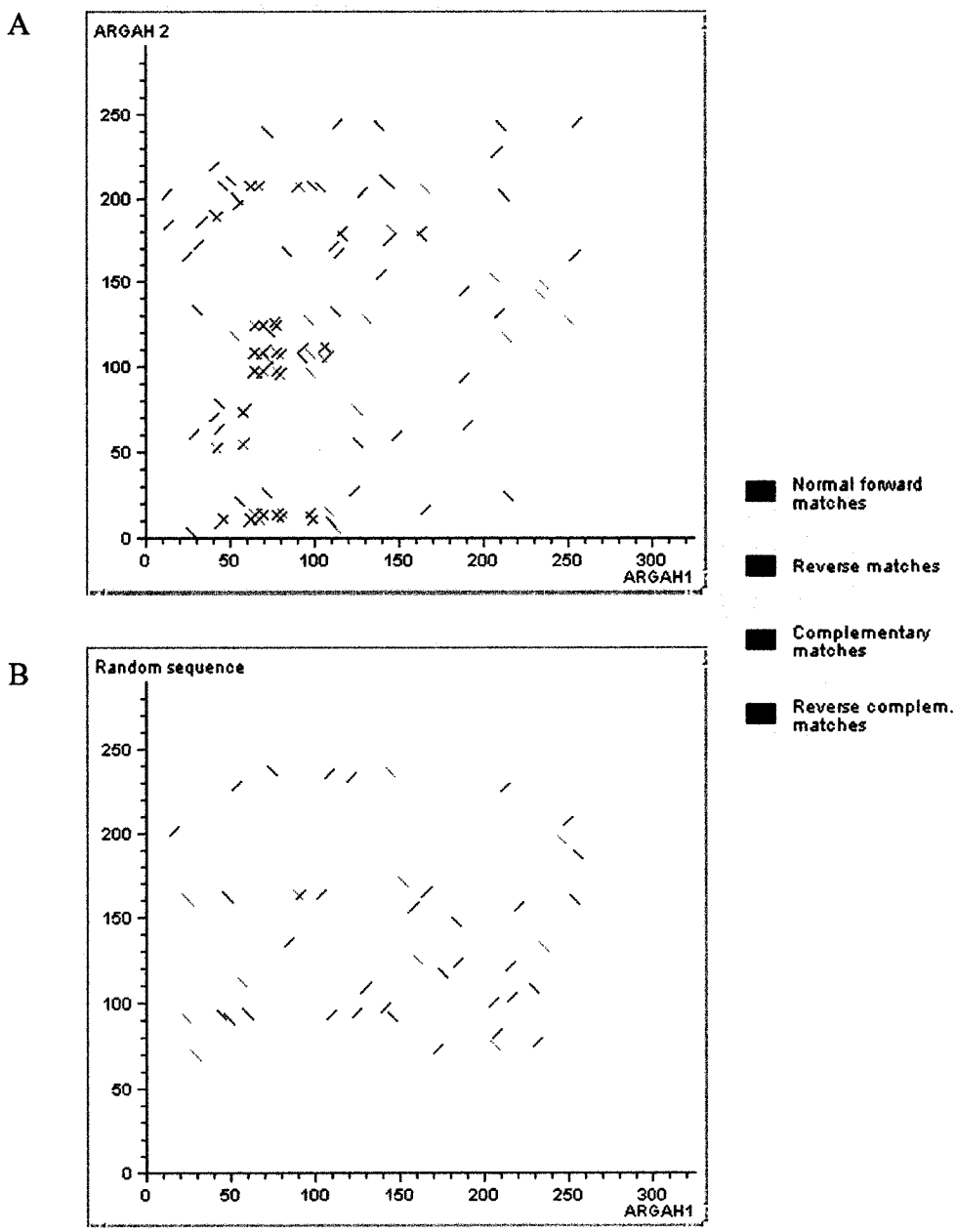


Figure 4.7: Dot plot analysis of the ARGAH1-A and ARGAH2-A sequences. Appoximately 250bp regions were compared and regions of similarity that were 6bp long or greater were marked. A. ARGAH1-A and ARGAH2-A sequence dot plot comparison. B. ARGAH1-1 sequence compared to random DNA sequence (generated by www.bioinformatics.org). Dot plots were made using the Bionautics Molecular Niology web portal (www.bionautics.com).

4.3.2 Arginase transcript abundance in pollen

In order to confirm *ARGAH1* transcript expression in pollen, we conducted quantitative reverse-transcriptase PCR (Figure 4.6). Pollen was isolated from wild-type plants, and primers specific to both *ARGAH1* and *ARGAH2* were designed. *ARGAH1* was expressed approximately three fold higher in pollen than leaf tissue but *ARGAH2* expression was over thirty fold less in pollen than in the leaf. Thus, *ARGAH1* but not *ARGAH2* appeared to be relatively highly expressed in pollen compared to vegetative tissues.

4.3.3 Cis- acting element motif searches of the upstream regions

ARGAH1-A and *ARGAH2-A* fragments conferred similar patterns of GUS expression, including staining young leaves, root and cotyledon vasculature of $7DIC_{22}$ seedlings, young leaves of rosette plants and in the young leaves of shoots emerging from the axils of rosettes and cauline leaves. Therefore, the sequences were compared to identify regions of DNA common to both fragments. BLAST sequence alignment of *ARGAH1-A* and *ARGAH2-A* indicated that there were no significant sequence similarities between the two DNA fragments (data not shown). A dot plot analysis of the 250bp upstream fragments generated by Bionautics Molecular web portal (www.bionautics.com) indicated that there were several small regions of nucleotide identity (Figure 4.7). The frequency of regions that had six bp or more of nucleotide

identity was greater when *ARGAH1*-A and *ARGAH1*-B were compared than when *ARGAH1-A* was compared to random DNA sequence (generated by www.bioinformatics.org). Thus, while there were no large regions of sequence that shared similarity, there were many small regions. These small regions may represent shared cis-acting regulatory motifs.

The presence of previously described cis-element motifs within the sequences of *ARGAH1*-A, *ARGAH1*-B, *ARGAH2*-A, and *ARGAH2*-B was investigated using the PLACE (plant cis-acting regulatory DNA elements) (Prestridge 1991, Higo et al. 1999) and PlantCare (Lescot et al. 2002) databases of cis-acting regulatory elements. Table 4.2 shows the location of some of the elements identified. The PlantCare database found no motifs (other than TATA and CAAT boxes) common to both the *ARGAH1*-A and *ARGAH2*-A sequences. However, the PLACE database found several elements that were present in both *ARGAH1-A* and *ARGAH2-A*, these included two elements previously identified in light-regulated genes, and two elements which were previously identified in seed storage protein genes. The light responsive elements found in both *ARGAH1*-A and *ARGAH2*-A included an I box and an Inr (initiator) element. Inr elements have been found in a photosynthesis light responsive gene without a TATA box in tobacco (Nakamura et al. 2002), while I boxes have been found in the promoters of light responsive genes in Arabidopsis (Donald and Cashmore 1990) and tomato (Rose et al. 1999). A SEF (soybean embryo factor), 1 of the seed storage motifs, was present in *ARGAH1*-A and *ARGAH2*-A was previously identified in a soybean (*Glycine max*) 7S globulin gene (Lessard et al. 1991), and the second element (E box) was identified in the napin gene of *Brassica*

Table 4.2: The locations of several cis-acting elements identified in the regions upstream of *ARGAH1* and *ARGAH2*. N refers to the presence of any nucleotide while R refers to the presence of a purine (A or G) and Y refers to the presence of a pyrimidine (T or C).

A: The locations of the seed storage motifs

Motif	Consensus	*ARGAH1*	*ARGAH2*
E box	CANNTG	-130	-137
			-684
			-719
			-941
Soybean Embryo Factor	RTTTTR	-195	-45
(SEF)		-200	-241
		-285	-615
		-567	-641
		-841	-690
		-966	-922
		-1358	-951

B: The locations of the light responsive motifs

Motif	Consensus	*ARGAH1*	*ARGAH2*
I box	GTAA	-150	-80
		-382	-980
		-409	
		-436	
		-524	
		-935	
		-943	
		-948	
		-1024	
Initatior (Inr)	YTCANTYY	-119	-43
		-303	-367
		-1137	-1075
		-1147	

C: The locations of the pollen related motifs

Motif	Consensus	*ARGAH1*	*ARGAH2*
GTGA	GTGA	-42	-24
		-54	-99
		-370	-652
		-637	-674
		-643	
		-765	
		-1144	
		-1263	
		-1471	
POLLEN1LELAT52	CGTCA	-98	-72
		-145	-445
		-319	-509
		-467	-665
		-574	
		-678	
		-706	
		-1047	
		-1077	
		-1138	
		-1347	
		-1382	
		-1439	

D: The locations of the wounding related motifs

Motif	Consensus	*ARGAH1*	*ARGAH2*
W box	TGACY	-580	-20
		-767	

napus (Stålberg et al. 1996). PLACE analysis of the larger fragments (*ARGAH1*-B and *ARGAH2*-B) identified several additional I boxes, and Inr elements in each upstream region. However, no new light responsive element types were identified in the upstream regions present in the longer constructs. Several additional copies of the SEF motif were present in the both the *ARGAH1*-B and *ARGAH2*-B upstream regions and similarly more copies of the E box motif were present in the *ARGAH2*-B sequence; however, no additional copies were present in the *ARGAH1*-B sequence.

The PLACE database (Prestridge 1991, Higo et al. 1999) also identified two types of pollen-related cis-acting elements in the fragments upstream of the Arabidopsis arginase genes (Table 4.2). A motif (GTGA) that was previously identified in two tomato late anther tomato (LAT) genes (Twell et al. 1991) and a tobacco (*Nicotiana tabacum*) pollen10 gene (Rogers et al. 2001) was found in both *ARGAH1* and *ARGAH2* genes. However, while this motif was found four times in the 1117bp upstream of *ARGAH2* it was present nine times in the 1481bp upstream of *ARGAH1*. In addition, this motif was only identified once more when 350bp of further upstream sequence of *ARGAH2* was searched using PLACE (data not shown). Thus, in comparable sized upstream regions *ARGAH1* contained nine copies of this motif while *ARGAH2* had five copies; this element was present in greater abundance in lines where GUS staining in pollen was observed.

The second type of cis-acting element related to pollen was also identified in the regions upstream of the Arabidopsis arginases (Table1). This POLLEN1LELAT52 element was first identified in the pollen-specific lat52 gene in tomato (Bate and Twell 1998) and then in an endo-β-mannanase gene in tomato

(Filichkin et al. 2004). Similar to the previously mentioned pollen cis-acting element, the *ARGAH 2*-B construct contained four copies of this motif while *ARGAH1*-B contained thirteen copies. A motif search of 350bp further upstream of *ARGAH2* revealed the presence of an additional three copies of this motif. Therefore, in comparable sized upstream regions *ARGAH1* contained thirteen copies of this motif while *ARGAH2* had seven. Also, similar to the other pollen cis-acting element, there was a difference in the number of elements present in the two *ARGAH1* fragments used. There were four copies of the POLLEN1LELAT52 element were present in the *ARGAH1*-A construct and thirteen copies were present in the *ARGAH1*-B construct.

A W box element was identified in the sequence of *ARGAH2*-A, but not in *ARGAH1*-A; this was consistent with wound inducibility of this Arabidopsis arginase gene. However, further upstream (in *ARGAH1*-B), there were two copies of the W box. The W box was the only element present in a group of 26 promoters from genes that were coordinately expressed in response to systemic acquired resistance (Malek et al. 2000); thus, it appears this is an important cis acting element. Gel shift assays have shown that the W box is bound by a family WRKY transcription factors (Rushton et al. 1996, Yamamoto et al. 2004). In addition, these transcription factors increase in response to a several defense related stimuli (Li et al. 2004) including an oligopeptide elicitor from the pathogenic fungus *Phytophthora megasperm* (Rushton et al. 1996) to the fungus *Trichoderma viride* in tobacco (Yamamoto et al. 2004).

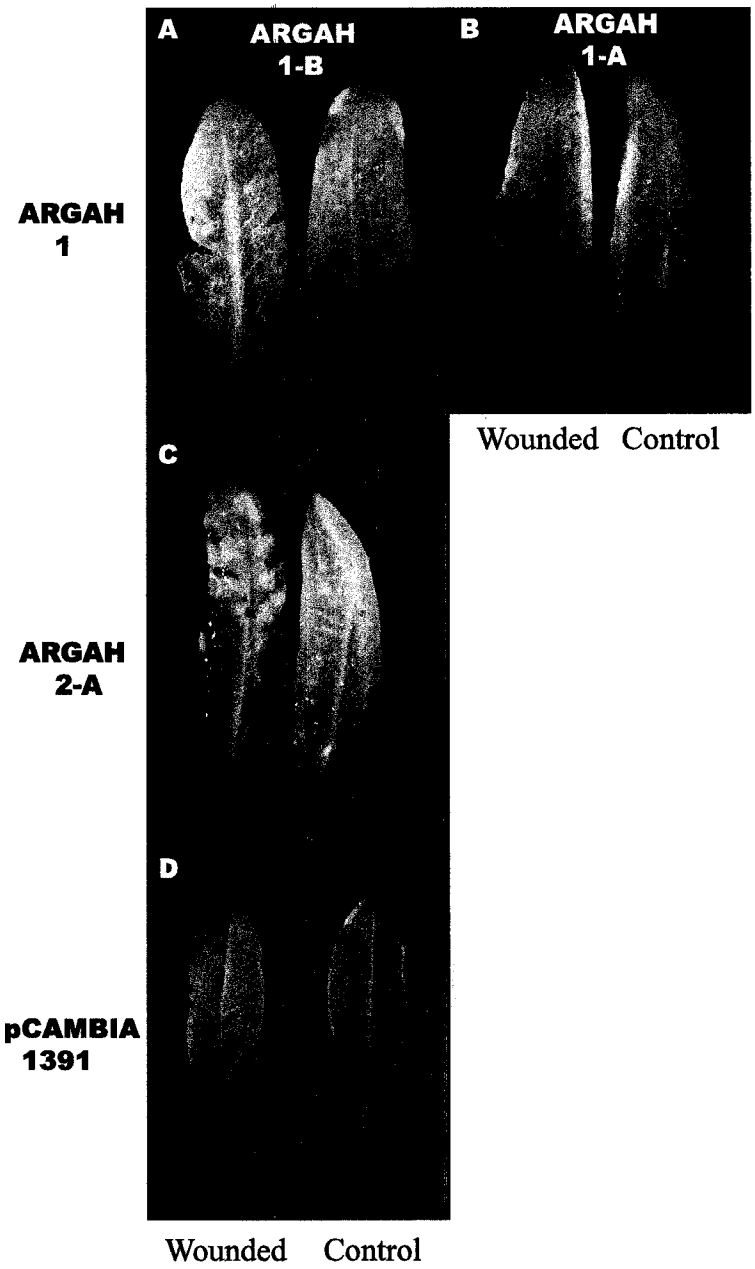


Figure 4.8 : GUS staining of wounded Arabidopsis leaves that contain ARGAH1 and ARGAH2 250bp promoter fragments. A. Wounded and control leaves from plants containing the shortest ARGAH1 fragment generally do not stain in response to wounding. However, when staining is observed (B) it is diffuse and does not directly surround the wound sites. C. ARGAH 2 is wound induced GUS staining is distinct and surroundes the wound sites. D. No GUS staining is visible in wounded pCAMBIA 1391 leaves.

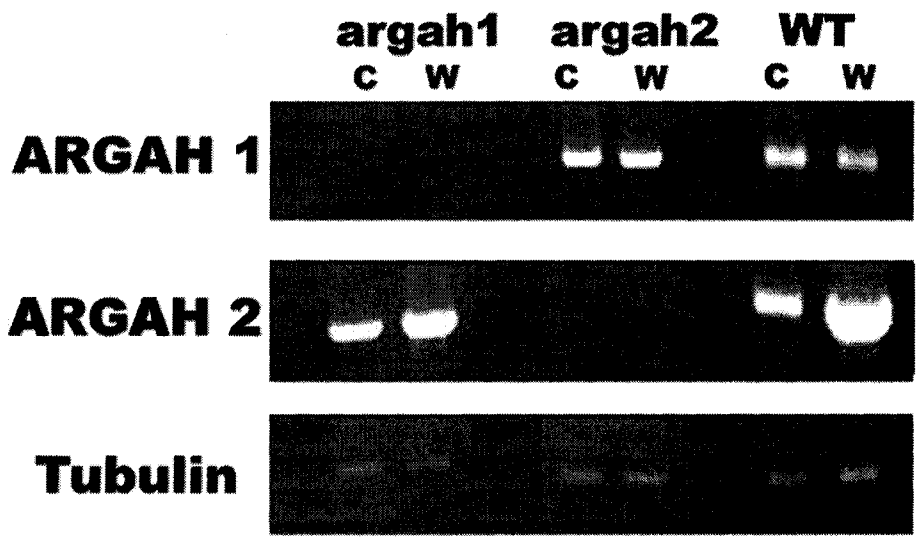


Figure 4.9: RT-PCR of ARGAH1, ARGAH2 and TUBULIN 4 in wounded and control leaves of wild-type and T-DNA insertion line (argah1 and argah2) plants. Plants were grown to the rosette stage and using forceps were wounded along the leaf margins.

4.3.4 Wounding and arginase in Arabidopsis

In addition to the presence of cis-acting motifs previously identified in wound inducible genes, wound induction has been reported for one of the two arginase genes of tomato (Chen et al. 2004). To determine if any of the Arabidopsis reporter gene fusions we produced were responsive to wounding, we conducted GUS staining of wounded rosette leaves our transgenic plants. Figure 4.8 shows that *ARGAH2-A* had intense GUS staining immediately surrounding the wound sites; this was not observed in any *ARGAH1-A*-GUS lines. In some *ARGAH1-A* lines, a faint, diffuse staining was detected, which did not directly surround the wound sites as was observed in *ARGAH2*-A GUS leaves. This pattern of staining was observed in most (5/8) *ARGAH1-A* lines but in none of the five *ARGAH 1*-B lines; thus the presence of the larger promoter fragment appears to abolish the diffuse staining pattern observed in wounded *ARGAH1-A* plants.

In order to confirm that wounding did induce *ARGAH2* (C Todd unpublished data) expression we conducted semi-quantitative RT-PCR using homozygous *ARGAH 1* (SALK 0157987) and *ARGAH2* (SAIL 181_C11) T-DNA insertion lines (Todd et al. unpublished data). The lines had a T-DNA insertion in the 5' UTR of *ARGAH1* and in the second exon of *ARGAH2* (www.arabidopsis.org). Primers unique to *ARGAH1* and *ARGAH2* were designed and RT-PCR was conducted on wounded and control rosette leaves. Figure 4.9 shows that *ARGAH2* but not *ARGAH1* transcript accumulates in response to wounding. This accumulation was observed in both

argah1 and wild-type lines; no transcript was detected in *argah2* control or wounded leaves. *ARGAH1* levels did not appear to be altered in response to wounding; this was observed in both the wild-type and *argah2* plants. As expected, the *ARGAH 1* transcript was not detected in *argah1* leaves. As a positive control we used the Arabidopsis *TUB4* gene; with equal loading of the samples, the transcripts of this gene appeared to remain constant in regardless of the treatment or the genotype.

4.4 Discussion

4.4.1 Cis-acting regulatory elements

The isolation of the upstream regions of *ARGAH1* and *ARGAH2* in Arabidopsis allowed us to look at the pattern of expression that the regions conferred. We observed GUS staining in transgenic lines containing approximately 250bp of genomic sequence upstream of either *ARGAH1* or *ARGAH2*. Thus, it appeared that these fragments contained sufficient regulatory elements to confer staining. The staining patterns observed in $7DIC_{22}$ (days in culture at 22°C) were very similar for the *ARGAH1* and *ARGAH2* lines. However, dot plot analysis (Figure 4.6) of the 250bp upstream regions indicated that while regions of nucleotide similarity did exist, the *ARGAH1* and *ARGAH2* upstream sequences were generally not highly conserved. Therefore, the similar aspects of the expression patterns observed in *ARGAH1* and *ARGAH2* derived constructs could not be explained on the basis of extensive sequence homology.

A search for cis-acting regulatory elements in the PLACE database (Prestridge 1991, Higo et al. 1999) identified several elements common to both *ARGAH1* and *ARGAH2*. PLACE identified two types of elements previously identified in seed storage proteins (Table 4.2). The first was an E box which is characteristically found in angiosperm 2S storage protein genes (Stålberg et al. 1996). The second type of element found in both *ARGAH1*A and *ARGAH2*A was an element identified in a soybean 7S globulin storage protein: a soybean embryo motif (SEF) (Lessard et al.

1991). However, the storage protein genes would be expected to be expressed during seed development when the plant is producing storage reserves. Arginase has been shown to be involved in storage reserve mobilization in both gymnosperm and angiosperm species (King and Gifford 1997, De Ruiter and Kollöffel 1983). Storage reserve deposition and catabolism would likely not occur concurrently in the plant, thus it seems counter-intuitive that these elements would be involved in the expression of arginase during this stage of development. Nevertheless, these conserved cis-elements may have a role in the expression of arginase at other developmental stages or have different functions in different contexts.

Light responsive elements were identified in all the arginase upstream regions (Table 4.2). In Arabidopsis, arginase activity increases in after $2DIC_{22}$. This increase is significantly larger and continues in seedlings that have a functional *ARGAH2* (i.e. wild-type and *argah1* T-DNA lines) (Todd et al. unpublished data). Germination occurred at $2DIC_{22}$; perhaps light perception plays a role in arginase expression. In mammals, circadian rhythms influence arginase activity (Soler et al. 1988, Fuentes et al. 1994) and microarray data have indicated that is also be true in Arabidopsis (Harmer et al. 2000). However, the induction of arginase in response to light still needs to be determined experimentally.

4.4.2 Staining patterns and arginase expression in pollen

Staining of young leaf tissue was observed in several developmental stages (Figure 4.4). Young leaves of $7DIC_{22}$ seedlings stained as did young rosette leaves

and young leaves surrounding the axillary shoots of both cauline and rosette leaves in both *ARGAH1* and *ARGAH2* lines. The transport of Arg into young leaves to supply them with N would explain the presence of arginase in these tissues. Staining was also present in several vascular tissues: in the roots and cotyledons of $7DIC_{22}$ seedlings (Figure 4.3), in the sepal vasculature and in the vascular bundle of the filament (Figure 4.5). The major amino acids involved in N transport are glutamate, glutamine, aspartate and asparagine (Buchanan et al. 2000); Arg transport does occur although to a seemingly lesser extent (Lea and Ireland 1999). However, examples of Arg transport do exist: the majority of the storage proteins are located in the megagametophyte of the loblolly pine seed (Stone and Gifford 1997); however during storage protein mobilization Arg accumulates to a greater extent in the cotyledons of the developing seedling (where arginase is highly active) as compared to the megagametophyte (King and Gifford 1997). Thus, Arg appears to be transported from the megagametophyte to the cotyledons following germination. In addition, Arg accumulation (along with glutamine and asparagine) was observed following stem girdling in poplar (Cooke et al. 2003). Together these examples indicate that Arg is at least occasionally transported and perhaps arginase is active in the vascular tissues. However, the presence of a functional arginase in vascular tissue would need to be determined experimentally.

When the GUS staining patterns were compared with the RT-PCR data generated for *ARGAH1* and *ARGAH2* (Figure 4.1, C. Todd unpublished data), there were some differences. RT-PCR showed that *ARGAH1* is expressed at a low level in all tissues tested (seedlings, leaves, roots, stems, flowers and siliques) (Figure 4.1).

However, staining was not observed in stems or mature leaves and was inconsistent in mature roots (data not shown). *ARGAH2* RT-PCR showed relatively higher expression compared to *ARGAH1*, and transcript was present in seedlings, leaves, flowers and siliques. With the exception of the presence of the transcript in leaves the *ARGAH2* RT-PCR data agrees with the observed staining pattern: *ARGAH2* being expressed in the seedling, flowers, and siliques. Two possible explanations exist for the inconsistencies between the GUS staining and the RT-PCR data. First, the entire promoter may not be contained in the *ARGAH* constructs. Second, there may some post-transcriptional regulation of arginase.

An analysis of the Arabidopsis microarray expression database using GENEVESTIGATOR gene chronologer (Zimmerman et al. 2004) showed that *ARGAH2* transcript abundance was the highest in unopened flowers when compared to other developmental stages (Figure 4.10). GENEVESTIGATOR gene atlas also indicated that in inflorescences *ARGAH2* expression was the highest in the sepals, siliques, and cauline leaves but absent in pollen; this data supports the observed *ARGAH2* GUS staining patterns. Unfortunately, no GENEVESTIGATOR data was available for *ARGAH1*. In addition, we queried the Arabidopsis electronic fluorescent protein (eFP) browser (www.bbc.botany.utoronto.ca/efp) (Schmid et al. 2005); according to eFP*ARGAH1* and *ARGAH2* transcripts were present in virtually identical patterns with the highest levels of transcript detected in the cotyledons of the seedling, all apical tissues of the rosette and in young leaves (Figure 4.11). However, the observed patterns of staining and transcript accumulation from RT-PCR were not identical for *ARGAH1* and *ARGAH2*. Combining the *in silico* analyses, we notice that

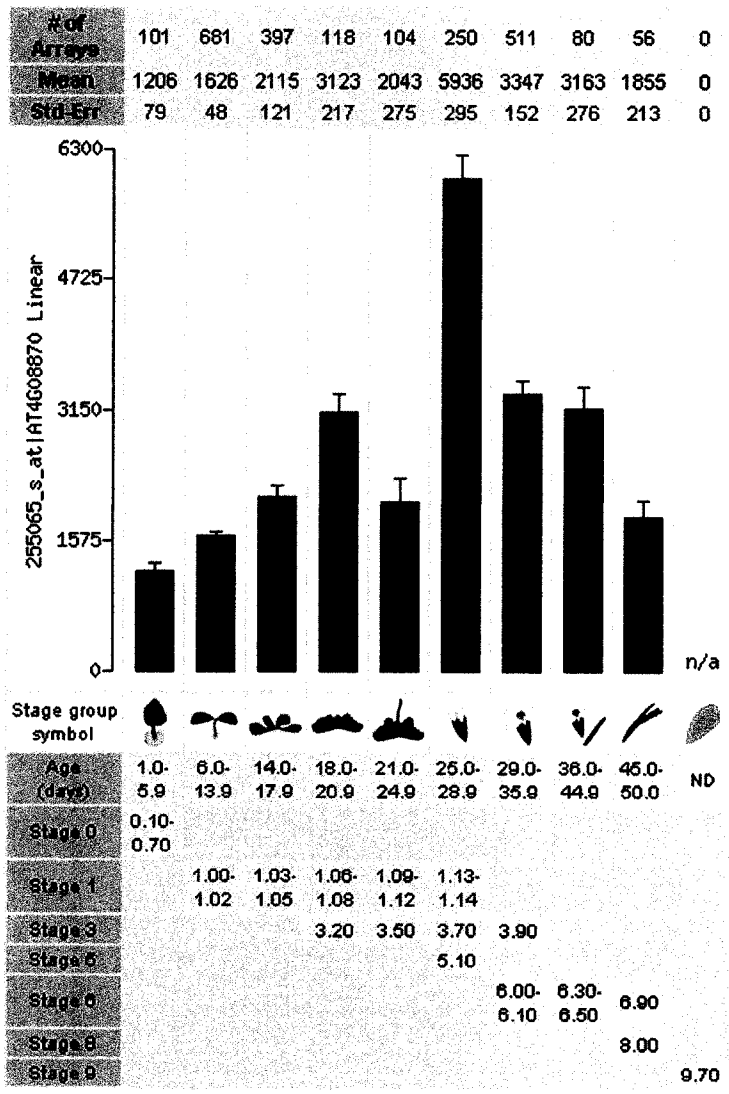


Figure 4.10: ARGAH2 (At4g08870) GENEVESTIGATOR gene chronologer profile (www.arabidopsis.org) (Zimmerman et al 2004). ARGAH2 transcript appears to be most abundant in unopened flowers followed by a lower level in rosettes, opened flowers, and opened flowers with young siliques.

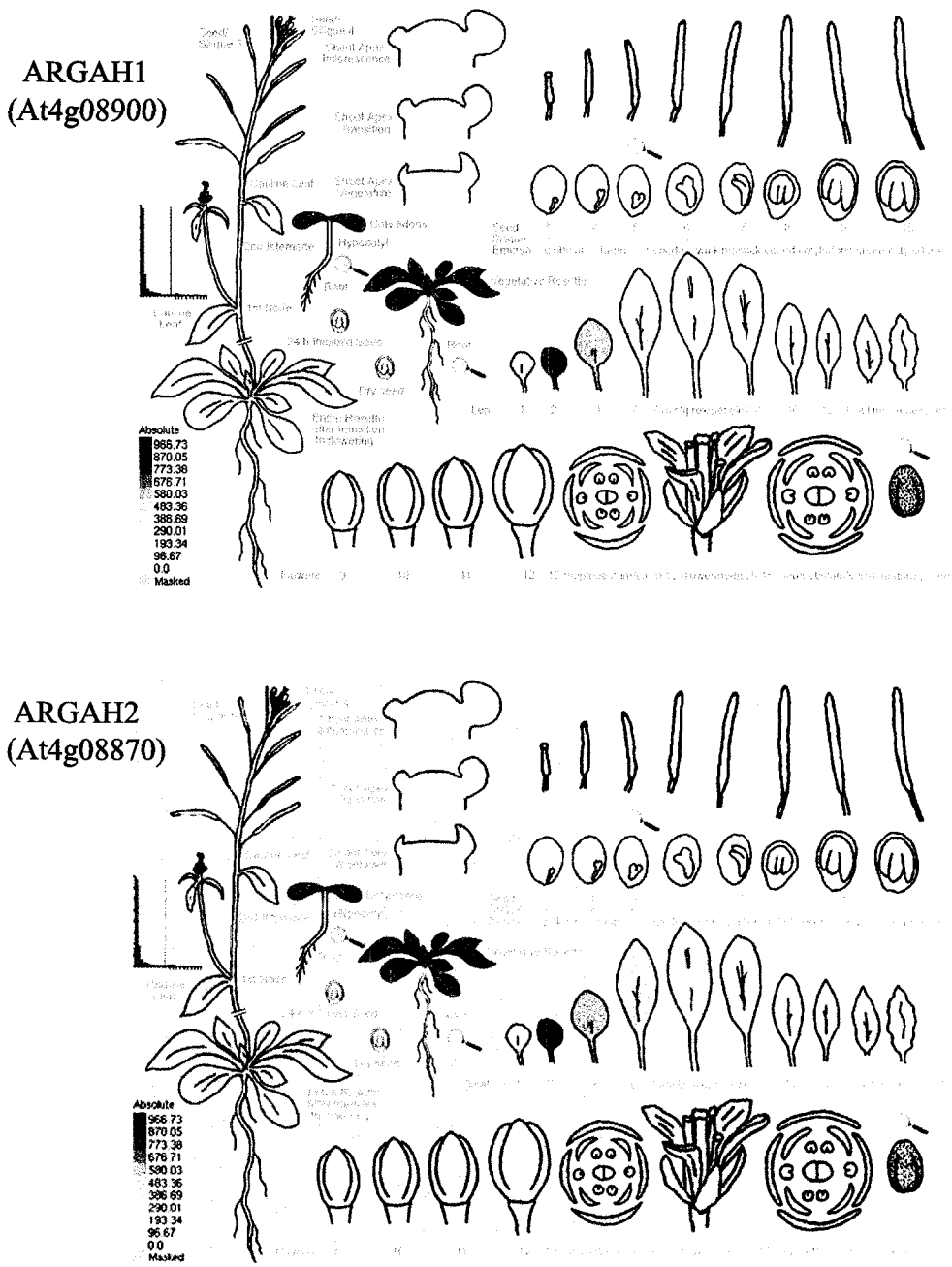


Figure 4.11: ARGAH1 and ARGAH2 patterns of expression from eFP (electronic fluorescent protein) browser (www.bbc.botany.utoronto.ca/efp) (Schmid et al. 2005). ARGAH1 and ARGAH2 appear to have identical transcript accumulation patterns with highest abundance in the cotyledons of seedlings, apical parts of the rosette, and young leaves.

ARGAH1 and *ARGAH2* are expressed in the cotyledons of seedlings and in young leaves, which agrees with our observed staining pattern. However, GUS staining was not observed in rosette leaves of any of the *ARGAH1* or *ARGAH2* lines but both *ARGAH1* and *ARGAH2* transcripts accumulate according to RT-PCR and eFP data (and *ARGAH2* GENEVESTIGATOR). Again, as our transgenic plants did not contain the full upstream region, it is possible that elements were missing from our constructs; alternatively there may some post-transcriptional regulation.

GUS staining was observed in the pollen of *ARGAH1*-B plants but was not observed in any of the *ARGAH2* lines tested. Similarly, quantitative PCR indicated that *ARGAH2* but not *ARGAH1* transcript was present in pollen. Thus, it appears that only one form of arginase in Arabidopsis is actively transcribed in pollen. To our knowledge the role of arginase in pollen has not been previously investigated. Amino acid analysis of various tissues in tomato revealed that the proline (Pro) content in pollen was 60 fold higher than any other organ tested and that proline represented greater than 70% of the total amino acids found in pollen (Schwacke et al. 1999). Similarly, in *Petunia* proline comprises almost 60% of the total amino acid pool in pollen (Hong-qi et al. 1982). It has been suggested that this high level of proline is utilized in the production of cell wall components (hydroxyproline rich glycoproteins) necessary for pollen tube extension (Hong-qi et al. 1982). Thus, with such a demand for Pro in mature and germinating pollen, it would be reasonable to hypothesize that the function of arginase in pollen would be in the production of proline. In Arabidopsis, methyl jasmonate (MeJA) is important in pollen development (Benedetti

et al. 1998). Perhaps one of the roles for MeJA in pollen development is to stimulate arginase which, in turn, increases the proline content.

The PLACE database (Prestridge 1991, Higo et al. 1999) identified two types of pollen related cis-acting elements in both of the Arabidopsis arginase genes. The GTGA motif that was previously identified in tomato (Twell et al. 1991) and tobacco (Rogers et al. 2001) has been shown to be important in pollen specific expression. Transient expression assays of this GTGA motif in tobacco pollen showed that changing the motif from GTGA to TATG resulted in a two fold decrease in reporter gene activity (Rogers et al. 2001). This element was detected four times in the 1117bp upstream of *ARGAH2* (ARGAH2-B) but was present nine times in the 1481bp upstream of *ARGAH1 (ARGAH1-B)*. Thus, given the relative abundance and demonstrated importance in pollen expression other species it is possible that this element is important *ARGAH1* expression in pollen. In our GUS staining we did not observe pollen staining in any of the six staining lines of *ARGAH1*-A. In comparing the location of this cis-acting element in the *ARGAH1* upstream sequence we found that of the nine copies of the GTGA element in the *ARGAH1*-B construct only three are present in *ARGAH1*-A construct. From this data we cannot determine if a low level of GUS (below detection) was present in *ARGAH1*-A pollen; however, it would be interesting to determine.

The POLLEN1LELAT52 element was the second type of pollen related cis-acting element identified in the Arabidopsis arginases. This AGAAA element was first identified in the pollen specific lat52 gene in tomato (Bate and Twell 1998) and then in an endo-mannase gene (*LeMAN*) also in tomato (Filichkin et al. 2004). Similar

to the previous pollen cis-acting element the *ARGAH2*-B construct contained four copies of this motif while *ARGAH1*-B contained 13 copies. Thus, increased abundance of the POLLEN1LELAT52 element appears to coincide with the appearance of GUS staining in the *ARGAH2* lines. Interestingly, the upstream region of the *LeMAN5* gene of tomato also contained an E-box element which is conserved in the promoters of storage-protein gene promoters (Stålberg et al. 1996). Endo-β-mannanases degrade the mannan polymers of cell walls and have been associated with the mobilization of stored mannans in the cell walls of the endosperm (Filichkin et al. 2004). Thus, it is interesting that two genes associated with the mobilization of seed storage reserves after germination appear to be expressed in pollen.

4.4.3 Arginase and wounding

GUS staining and RT-PCR indicated that *ARGAH2* but not *ARGAH1* is induced in response to wounding (Figures 4.7 and 4.8). Of the tomato arginases, *LeARG2* is induced in response to wounding, treatment with methyl jasmonate and to the pathogen *Pseudomonas syringae*. It would be interesting to compare the upstream sequences of the tomato arginases to those of Arabidopsis in order to determine if what if any regions of homology are shared; however these sequences are not available. It is possible that *ARGAH2* and *LeARG2* may share motifs as they both appear to be wound inducible. In tomato, *LeARG2* was also shown to be induced in response to *P. syringae* (Chen et al. 2004). A GENEVESTIGATOR response viewer analysis (Zimmermann et al. 2004) indicated that Arabidopsis *ARGAH2* responds not

only to *P. syringae* but also to the fungus *Botrytis cinerea* and to a lesser degree to the fungus *Erysiphe orontii*. It would be interesting to confirm the induction of arginase in response to these pathogens and then possibly look at arginase expression and activity in some of the defense signaling mutants to attempt to determine how arginase is induced in response to wounding.

In our investigation of arginase in Arabidopsis we have shown that the upstream regions of *ARGAH1* and *ARGAH2* confer staining in 7day old seedlings, as well as in rosettes and flowers. In these upstream regions we have also identified several types of cis-acting elements including those that are light-responsive. We have also shown that arginase (*ARGAH1*) transcript accumulates in pollen and have hypothesized that arginase may play a role in the synthesis of Pro which is highly abundant in pollen (Schwacke et al. 1999). Arginase (*ARGAH2*) transcript also induced in response to wounding. Thus, the different forms of arginase appear to have different roles in the plant but each hydrolyzes Arg and allows the plant to utilize the products: ornithine and urea.

4.5 References

Alabadí, D, Agüero MS, Pérez-Amador MA, Carbonell J. 1996. Arginase, arginine decarboxylase, ornithine decarboxylase, and polyamines in tomato ovaries. Plant Physiology 112: 1237-1244.

Bais HP and Ravishankar GA. 2002. Role of polyamines in the ontogeny of plants and their biotechnological applications. Plant Cell, Tissue and Organ Culture 69: 1-34.

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

Benedetti CE, Costa CL, Turcinelli SR, Arruda P. 1998. Differential expression of a novel gene in response to coronatine, methyl jasmonate, and wounding in the *Coi1* mutant of Arabidopsis. Plant Physiology: 116: 1037-1042.

Buchanan BB, Gruissem W, Jones RS. 2000. Biochemistry & Molecular Biology of Plants. American Society of Plant Biology. Rockville MD.

Chen H, McCaig BC, Melotto M, He SY, Howe GA. 2004. Regulation of plant arginase by wounding, jasmonate, and the phytotoxin coronatine. Journal of Biological Chemistry 279: 45998-46007.

Cooke JEK, Brown KA, Wu R, Davis JM. 2003. Gene expression associated with N-induced shifts in resource allocation in poplar. Plant, Cell and Environment 26: 757-770.

Dellaporta SL, Wood J, Hicks JB. 1983. A plant DNA minipreparation: version II. Plant Molecular Biology Reports 1: 19-21.

Delledonne M, Xia Y, Dixon RA, Lamn C. 1998. Nitric oxide functions as a signal in plant disease resistance. Nature 394: 585-588.

De Ruiter H and Kollöffel C. 1983. Arginine catabolism in the cotyledons of developing and germinating pea seeds. Plant Physiology 73: 525-528.

Donald RGK, Cashmore AR. 1990. Mutation of either G box or I box sequences profoundly affects expression from the Arabidopsis rbcS-1A promoter. EMBO Journal 9: 1717-1726

Doyle JJ and Doyle JL. 1987. A rapid DNA isolation procedure from small quantities of fresh leaf tissues. Phytochemistry Bulletin 19:11-15.

Duleu S, Vincendeau P, Courtois P, Semballa S, Lagroye I, Daulouède S, Boucher JL, Wilson KT, Veyret B, Gobert AP. 2004. Mouse strain susceptibility to trypanosome infection: and arginase-dependent effect. The Journal of Immunology. 172 :6298-6303.
Durner J, Wendehenne D, Klessig DF. 1998. Defense gene induction in tobacco by nitric oxide, cyclic GMP and cyclic ADP-ribose. Proceedings of the National Academy of Sciences 95: 10328-10333.

Filichkin SA, Leonard JM, Monteros A, Liu PP, Nonogaki H. 2004. A novel endo-beta-mannanase gene in tomato *LeMAN5* is associated with anther and pollen development. Plant Physiology 134: 1080-1087

Fuentes JM, Pascual MR, Salido F, Soler G, Madrid JA. 1994. Oscillations in rat liver cytosolic enzyme activities of the urea cycle. Archives of the International Society of Physiology, Biochemisty and Biophysics 102: 237-41.

Gander ES, Holmstroem KO, De Paiva GR, De Castro LAB, Carneiro M, Grossi de Sá MF. 1991. Isolation, characterization and expression of a gene coding for a 2S albumin from *Bertholletia excelsa* (Brazil nut). Plant Molecular Biology 16: 437-448.

Goldraij A, and Pollaco JC. 1999. Arginase is inoperative in developing soybean embryos. Plant Physiology 119: 297-303.

Guo FQ, Okamoto M, Crawford NM. 2003. Identification of a plant nitric oxide synthase gene involved in hormonal signaling. Science 302: 100-103.

Harmer SL, Hogenesch JB, Straume M, Chang HS, Han B, Zhu T, Wang X, Kreps JA, Kay SA. 2000. Orchestrated transcription of key pathways in Arabidopsis by the circadian clock. Science 290: 2110-2113.

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

Hong-qi Z, Croes AF, Linskens HF. 1982. Protein synthesis in germinating pollen of Petunia: role of proline. Planta 154: 199-203.

Honys D and Twell D. 2003. Comaprative analysis of the Arabidopsis pollen transcriptome. Plant Physiology 132: 640-652.

Jenkinson CP, Grody WW, Cederbaum SD. 1996. Comparative Properties of Arginases. Comparative Biochemistry and Physiology 114B: 107-132.

King JE and Gifford DJ. 1997. Amino acid utilization in seeds of loblolly pine during germination and early seedling growth. I. Arginine and arginase activity. Plant Physiology 113: 1125-1135.

Krumpelman PM, Freyermuth SK, Cannon JF, Fink GR, Polacco JC. 1995. Nucleotide Sequence of *Arabidopsis thaliana* Arginase Expressed in Yeast. Plant Physiology 107: 1479-1480.

Lea PJ and Ireland RJ. 1999. Nitrogen Metabolism in Higher Plants. In: Plant Amino Acids, Biochemistry and Biotechnology. arcel Dekker Inc. New York, New York.

Lescot M, Déhais P, Moreau Y, De Moor B, Rouzé P, Rombauts S. 2002. PlantCare: a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Research 30: 325-327.

Lessard PA, Allen RD, Bernier F, Crispino JD, Fujiwara T, Beachy RN. 1991. Multiple nuclear factors interact with upstream sequences of differentially regulated beta-conglycinin genes. Plant Molecular Biology 16: 397-413.

Li J, Brader G, Palva T. 2004. The WRKY70 transcription factor: a node of convergence for jasmonate-mediated and salicylate-mediated signals in plant defense. The Plant Cell 16: 319-331.

Livak KJ and Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C}{}_T$ method. Methods 25: 402-408.

Malek K, Levine A, Euglem t, Morgan A, Schmid J, Lawton KA, Dangl JL, Dietrich RA. 2000. The transcriptome of *Arabidopsis thaliana* during systemic acquired resistance. Nature Genetics 26: 403-410.

Micallef BJ and Shelp BJ. 1989. Arginine metabolism in developing soybean cotyledons. I. Relationship to nitrogen nutrition. Plant Physiology 90: 624-630.

Micallef BJ and Shelp BJ. 1989b. Arginine metabolism in developing soybean cotyledons. III.Utilization. Plant Physiology 91: 170-174.

Nakamura M, Tsunoda T, Obokata J. 2002. Photosynthesis nuclear genes generally lack TATA-boxes: a tobacco photosystem I gene responds to light through an initiator. Plant Journal 29: 1-10.

Perez-Amador MA, Leon J, Green PJ, Carbonell J. 2002. Induction of the arginine decarboxylase *ADC2* gene provides evidence for the involvement of polyamines in the wound response in Arabidopsis. Plant Phys 130: 1454-1463.

Prestridge DS. 1991. SIGNAL SCAN: A computer program that scans DNA sequences for eukaryotic transcriptional elements CABIOS 7: 203-206.

Rogers HJ, Bate N, Combe J, Sullivan J, Sweetman J, Swan C, Lonsdale DM, Twell D. 2001. Functional analysis of cis-regulatory elements within the promoter of the tobacco late pollen gene 10. Plant Molecular Biology 45: 577-585

Rose A, Meier I, Wienand U. 1999. The tomato I-box binding factor *LeMYBI* is a member of a novel RT class of Myb-like proteins. Plant Journal 20: 641-652.

Rushton PJ, Torres JT, Parniske M, Wernert P, Hahlbrock K, Somssich IE. 1996. Interaction of elicitor-induced DNA binding proteins with elicitor response elements in the promoters of parsley PR1 genes. EMBO Journal 15: 5690-5700.

Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Schölkopf B, Weigel D, Lohmann JU. 2005. A gene expression map of *Arabidopsis thaliana* development. Nature Genetics 37: 501-506.

Schwacke R, Grallath S, Breitkreuz KE, Stransky E, Stransky H, Frommer WB, Rentsch D. 1999. *LeProT1*, a transporter for proline, glycine betaine, and γ–amino butyric acid in tomato pollen. The Plant Cell 11: 377-391.

Soler G, Bautista JM, Madrid JA, Salido GM. 1988. Circadian rhythms in enzymatic activity or rat liver arginase and glucose-6-phosphate dehydrogenase. Chronobiologia 15: 205-212.

Stålberg K, Ellerstom M, Ezcurra I, Ablov S, Rask L. 1996. Disruption of an overlapping E-box/ABRE motif abolished high RT transcription of the napA storage-protein promoter in transgenic *Brassica napus* seeds. Planta 199:515-519

Stone SL and Gifford DJ. 1997. Structural and biochemical changes in loblolly pine (*Pinus taeda* L.) seeds during germination and early-seedling growth. I. Storage protein reserves. International Journal of Plant Science 158: 727-737.

Twell D, Yamaguchi J, Wing RA, Ushiba J, McCormick S. 1991. Promoter analysis of genes that are coordinately expressed during pollen development reveals pollen-specific enhancer sequences and shared regulatory elements. Genes & Development 5: 496-507.

Weigel D and Glazebrook J. (2002). Arabidopsis a laboratory manual. Cold Spring Harbour Laboratory Press. Cold Spring Harbour N.Y. USA.

Wu G and Morris SM. 1998. Arginine metabolism: nitric oxide and beyond. Biochemistry Journal 336: 1-17.

Yamamoto S, Nakano T, Suzuki K, Shinshi H. 2004. Elicitor-induced activation of transcription via W box-related cis-acting elements from a basic chitinase gene by WRKY transcription factors in tobacco. Biochimica et Biophysica Acta 1679: 279-287.

Zimmermann P, Hirsch-Hoffmann M, Hennig L, and W Gruissem. 2004. GENEVESTIGATOR: Arabidopsis Microarray Database and Analysis Toolbox Plant Physiology 136: 2621-2632

Zonia LE, Stebbins NM, Polacco JC. 1995. Essential role of urease in germination of nitrogen-limited *Arabidopsis thaliana* seeds. Plant Physiology 107: 1097-1103.

Chapter 5: Arginase and methyl jasmonate

5.1 Introduction

Arginase (L-Arginine ureahydrolase, or amidinohydrolyase, EC 3.5.3.1) hydrolyzes arginine (Arg) to produce urea and the non-protein amino acid ornithine. Arginine is not only a component of proteins but also an important nitrogen (N) storage molecule for many plants (Micallef and Shelp 1989). In addition, Arg and its degradation products are involved in the urea cycle and the production of polyamines and proline.

Arginase activity has been described in many plant species. In tomato (*Lycopersicon esculentum*), arginase activity was detected in unpollinated ovaries (Alabadí et al. 1996) and gene expression has been detected in several other tissues as well (Chen et al. 2004). In tomato, two arginase genes, *LeARG1* and *LeARG2,* have been cloned. Both proteins, when expressed in *E. coli*, produced urea and had the highest specific activity against L-Arg compared to other substrates tested. The two tomato arginases are expressed in flower buds, mature unopened and opened flowers and immature green fruit. *LeARG1* was also expressed in the roots but neither *LeARG1* nor *LeARG2* transcript was detected in stems or leaves of healthy tomato plants (Chen et al. 2004).

In Arabidopsis (*Arabidopsis thaliana*) two arginase genes have also been described. The first arginase, *ARGAH1* (At4g08900), was identified when its cDNA complemented the yeast (*Saccharomyces cerevisiae*) *car1* mutant (Krumpleman et al.

1995) The yeast mutant (*car1*, *dur1*, *ura*3-52) was arginase deficient and unable to grow on either Arg or urea as a sole N source, but was able to grow on plates with Arg as the sole N source with complementation. The ability of the cDNA to complement the mutant and accumulate urea, as well as its sequence homology to other known arginases confirmed its annotation as an arginase. After the Arabidopsis genome was sequenced a second putative arginase, *ARGAH2* (At4g08870), was identified by sequence homology. Todd et al. (unpublished data) have shown that like *ARGAH1*, the open reading frame (ORF) of *ARGAH2* was able to complement the *car1* mutation in yeast, allowing growth on Arg as the sole N source. *ARGAH1* and *ARGAH2* are so named in order to differentiate them from *ALTERED RESPONSE TO GRAVITY* (*ARG*1) gene. *ARGAH1* and *ARGAH2*, like the arginases of tomato, have distinctive spatial expression patterns. Using semi-quantitative RT-PCR Todd et al. (unpublished data) detected *ARGAH*1 transcripts in all tissues tested including seedlings, leaves, roots, stems, flowers and siliques. *ARGAH 2* is highly expressed in seedlings, leaves, flowers and siliques but the transcript was not detected in either the roots or stem. Thus, in both tomato and Arabidopsis the two arginases are differentially expressed with respect to tissue type.

The role of the arginase enzyme in plants has been investigated in several species. In loblolly pine and pea arginase is involved in the mobilization of storage reserves during early seedling growth (King and Gifford 1997, De Ruiter and Kollöffel 1983 and others). Arginase acts to hydrolyze the Arg released when storage proteins are degraded. As Arg is a major component of seed storage proteins and a major storage reserve of N (De Ruiter and Kollöffel 1983), arginase is an important

enzyme in providing the developing seedling with sufficient N until it is photoautotrophic. The tissue specific expression of arginase in Arabidopsis and tomato raises some interesting questions about its function in those tissues. In tomato arginase was also shown to be induced in response to wounding, methyl jasmonate (MeJA) and the pathogen *Pseudomonas syringae* (Chen et al. 2004). Jasmonates have been implicated in defense response signaling by inducing the expression of many defense related genes (Walters et al. 2002). JA induces the expression of proteinase inhibitors involved in protecting the plant from insect damage and genes involved in synthesis of and antimicrobial secondary metabolites including alkaloids, terpenoids and flavonoids (Pauw and Memelink 2005). Pre-treatment of some plant species with jasmonates has been shown to have a protective effect when the plants are later exposed to pathogen attack (Cohen et al. 1993). Thus, it appears that arginase may have a role in pathogen response. In tomato *LeARG2* transcript accumulates in response to MeJA treatment and wounding (Chen et al 2004). We have previously shown (Chapter 4) that *ARGAH2* transcript accumulates in response to wounding and here we show that the situation in *ARGAH2* is also induced in response to MeJA and quantify the increase in expression using quantitative PCR. In addition, we have conducted microarrays using an *ARGAH2* T-DNA insertion line to identify genes that are differentially expressed in response to the loss of *ARGAH2* and MeJA treatment.

5.2 Materials and methods

5.2.1 Methyl jasmonate treatment of Arabidopsis plants

Seeds were surface sterilized, incubated at 4°C for 2 days and then grown on ½ MS for approximately 21 days. Seedlings exhibiting resistance to hygromycin were placed into soil and grown (at 22°C, 16 hours light) to the rosette stage (before bolting). Plants were treated with MeJA using the method of Perez-Amador et al (2002): seedlings were sprayed with 0.01% (v/v) MeJA in 0.1% EtOH (v/v) and 0.01% (v/v) Tween or 0.1% EtOH (v/v) with 0.01% (v/v) Tween for control seedlings. Plants were covered with clear plastic cover and left for 8 hours at (22°C) in growth chamber. Leaves were then harvested and frozen under liquid nitrogen and stored at -80°C until use.

5.2.2 RNA extraction, cDNA synthesis and RT-PCR

RNA was isolated from 1-2 leaves using the Qiagen RNeasy Plant Mini kit as per the manufacturer's instructions. Tissue was ground under liquid N_2 using a mortar and pestle and mixed with 1mL of buffer RLT containing β –mercaptoethanol. The liquid was spun through a QIAshredder spin column and an RNeasy mini column, and then washed with buffer (RW1 and RPE) before air drying. The RNA was then resuspended in 30μL of water. The RNA was then DNase treated using the DNA-free protocol (Ambion). Buffer was added (0.1 volume) to the RNA and 1μL (2 units) of

rDNase I and then incubated at 37°C for 30 minutes. A single volume of DNase Inactivation Reagent was then added and incubated at room temperature for 2 minutes before centrifuging at 10,000 x g and then isolating the supernatant. The RNA was then quantified using a NanoDrop spectrophotometer (ND 1000).

cDNA was synthesized using the RevertAid H Minus First Strand cDNA Synthesis Kit (Fermentas). 250ng of total RNA was mixed with 0.5μg of oligo$(dT)_{12-18}$ primer (Invitrogen) . The mixture was then incubated at 70°C for 5 minutes and then chilled on ice before adding 4μL of 5X reaction buffer, 1μL of RNase Out (Invitrogen) and 2μL of 10mM dNTP mix. The mixture was then incubated at 37°C for 5 minutes before the addition of 1μL (200 units) of RevertAid H Minus M-MuLV reverse transcriptase. The mixture was incubated at 42°C for 60 minutes and the reaction was stopped by heating at 70°C for 10 minutes and finally cooled on ice.

cDNA (2μL) was added to 1μL of 10mM dNTPs, 1μL of each 1μM primer and 5μL of 10X PCR buffer (100mM Tris-HCl, 500mM KCl, 15mM $MgCl_2$). *ARGAH 1* was amplified using Arg 1 Fwd and Arg 1 Rev primers (Table 5.1); similarly *ARGAH 2* was amplified using Arg 2 Fwd and Arg 2 Rev (Table 5.1). The cycling conditions were as follows: 94°C for 1 minute, 28 cycles of 94°C for 1 minute, 62°C for 1 minute, 72°C for 90 seconds, followed by a final extension of 72°C for 10 minutes. Tubulin was used as a positive control for each tissue and treatment; this was amplified using the Tub 5' and Tub 3' primers (Table 5.1). The cycling conditions were as follows: 94°C for 1 minute, 30 cycles of 94°C for 1 minute, 55°C for 1 minute, 72°C for 90 seconds, followed by a final extension of 72°C for 10 minutes. The reactions were then visualized on a 1% agarose gel.

5.2.3 Micorarrays

The following microarray experiment was completed twice (i.e. 2 biological replicates) and the resulting data was pooled. Each biological replicate also included a dye reversal (technical replicate, as described below). Figure 5.1 shows an outline of the experiment performed.

Arabidopsis seedlings (*argah2* and wild-type Columbia) were grown on soil and treated with MJ or a control solution as described above. All apical (above soil) tissues were harvested, and tissues from at least three individual plants were pooled for extraction; tissue was stored at -80°C until use. Total RNA was isolated using the RNeasy Plant Mini protocol (Qiagen, as above) and was used to synthesize cDNA using the reverse transcriptase from the RevertAid H Minus First Strand cDNA Synthesis Kit (Fermentas) and the 3D Array900 microarray labeling kit (Genisphere, Hatfield, PA) .We incubated 2.5μg of total RNA, 1pmole of polyA oligomers (Cy3 or Cy5, Genisphere), 4μL of 5X Reaction buffer (250 mM Tris-HCl (pH 8.3 at 25°C), 250 mM KCl, 20 mM $MgCl_2$, 50 mM DTT, Fermentas) ,10mM dNTP mix (Genisphere), 1μL Superase-In RNase inhibitor (Genisphere) and 200U MMLV reverse transcriptase (Fermentas) at 42°C for two hours and subsequently stopped by the addition of 1μL of 1M NaOH, 100mM EDTA. Labeled cDNA from wild-type and *argah2* plants (either control or MJ treated) were mixed together with 2.5μL 2X SDS-Based Hybridization Buffer (Genisphere) and hybridized to the slides at 55°C

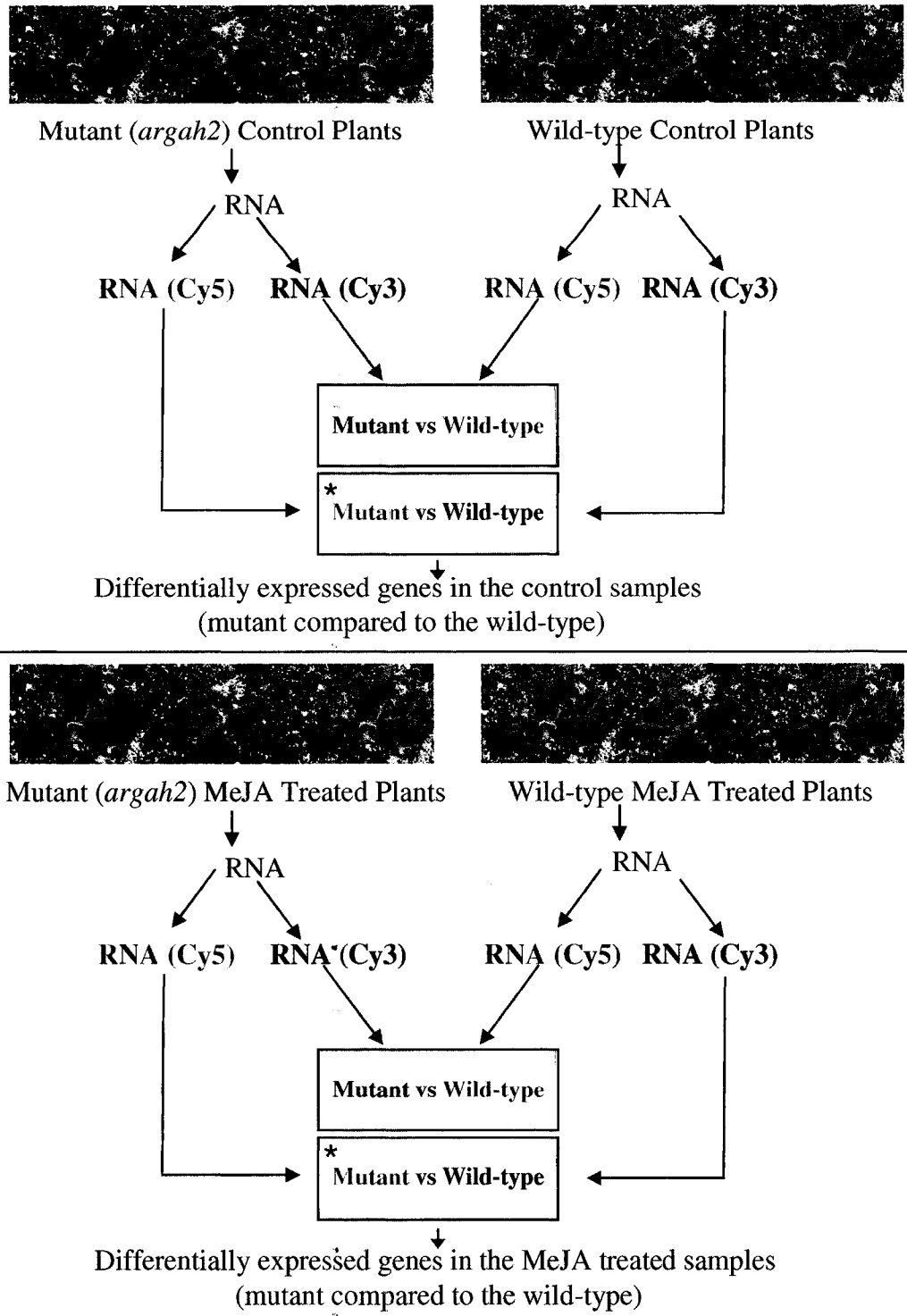


Figure 5.1: An overview of a single biological replicate of the microarray experiment. The above experiment was completed twice (two biological replicates) and each experiment was completed with a dye flip (*) to avoid bias in the microarrays due to differences in labeling efficiencies and/or in signal recording. Each RNA extraction was completed on a pool of at least three individual plants (either control or MeJA treated).

overnight in the dark under a 22 x 60 mm cover slip in a sealed humid 50mL centrifuge tube. Unbound cDNA was then removed by washing twice in pre-warmed (42°C) 2XSSC 0.2% SDS, once in room temperature 2XSSC and once in room temperature 0.2XSSC. Secondary hybridization mixes with 2.5μL of each Alexa Fluor Oyster dye Capture reagent (Genisphere) and 25μL 2X SDS-Based Hybridization Buffer (Genisphere) were prepared and incubated at 80°C for ten minutes. The mixture was added to the slides and the hybridizations were performed at 55°C for 4 hours in the dark under a 22 x 60 mm cover slip in a sealed humid 50mL centrifuge tube and washes repeated as above. To avoid bias in the microarray results as a consequence of dye-related differences in labeling efficiency and/or differences in recording fluorescence signals, dye labeling for each comparison was reversed in two individual hybridizations. Slides were immediately scanned using an ArrayWorx[e] scanner (Applied Precision, Seattle, WA).

Microarrays were analyzed using the TM4 microarray software (Saeed et al. 2003). Median signal intensity was calculated using Spotfinder and the data was normalized with Lowess in MeV. Data was analyzed using SAM to identify differentially expressed genes. A two class unpaired SAM was completed on all data with an estimated false discovery rate of 0 to identify genes that were differentially expressed (*argah2* vs wild-type) in both MeJA and control samples. SAM analyses (1 class with an estimated false discovery rate of 0) were also completed separately on MeJA and control data to identify genes that were differentially expressed (*argah2* vs wild-type) in each treatment (MeJA and control). To identify additional genes for

microarray validation, a t-test in MIDAS was completed using an α value of 0.05 and 3 degrees of freedom on the methyl-jasmonate treated data.

5.2.4 Quantitative PCR

Primers were tested against a range of template dilutions (1:4 to 1:16384) to ensure qPCR efficiency was not proportional to template quantity. Experimental replicates used the RNA samples used for microarrays. RNA was extracted (RNA plant mini kit, Qiagen) and pretreated with DNase I (DNA free, Ambion). Reverse transcription was completed (RevertAid H Minus First Strand cDNA Synthesis Kit, Fermentas) using 2.5μg of total RNA and oligo$(dT)_{12\text{-}18}$ (Invitrogen) as above. Real-time PCR was performed in an Applied Biosystems 7500 Fast Real-Time PCR System. For quantitative PCR reactions, 2.5μl of a 1/25 dilution of the reverse-transcription reaction was used in a total volume of 10μl with 0.4μM of each forward and reverse gene-specific primer (Table 5.1), 0.2 mM dNTPs, 0.25X SYBR Green, 1X ROX and 0.075U Platinum Taq (Invitrogen). Reactions were carried out at 65° C. Threshold cycles (C_T) were determined using 7500 Fast Software. C_T values were calculated using Ubiquitin 10 (Ubq 10) as an endogenous control (Ubq 10 Fwd and Ubq 10 Rev, Table 5.1) and $\Delta\Delta C_T$ values were generated using wild-type samples as a reference (Livak and Schmittgen 2001). Log_2 ratios show relative expression values between SAIL and wild-type samples.

Table 5.1: Sequences of Arabidopsis PCR primers used

Name:	Primer Sequence (5' to 3')
Arg 1 qRT Fwd	TGCACGATGAAACATAACAAAAAAC
Arg 1 qRT Rev	ATGGTTGCAGCTAAGCTTGTTAGA
Arg 2 qRT Fwd	AGAGCTAGCCGCAAAAATGTCA
Arg 2 qRT Rev	ACCATGTCTATGAGACCACACTTATTG
Ubquitin10 Fwd	TACTCTTCACTTGGTCCTGCGTCTTC
Ubquitin10 Rev	GGCCCCAAAACACAAACCACCAA
Arg 1 Fwd	GAGTGGCCGAAACAGAGATT
Arg 1 Rev	AGCCAGACTTATTGCATTCAAGAAC
Arg 2 Fwd	ATGTGGAAGATTGGGCAG
Arg 2 Rev	CCCAATTGATATTAAGCTTATTCGATACC
Tub 5'	TTGCGTTCTTCGTTTCCCTGG
Tub 3'	GAGGGTGCCATTGACAACATC
At1g74710 Fwd	GCGATTAATTGAAGAAAGAGTAACATTTG
At1g74710 Rev	ACAACATTGCTTTCTTATTGTGAGAAC
At1g10640 Fwd	CTCCCATTGATTGTTTGAAGATTG
At1g10640 Rev	TTGCGAAAAATCCGTTAACACA
At1g14880 Fwd	TTCTCTATGTGATCATCATTTGTCTTATGT
At1g14880 Rev	CAAAAACAAACAGAAGATTCAACAACAC
At1g33960 Fwd	TCTTCCACCCTCACACTTTGC
At1g33960 Rev	TCCGCGTTAGAAACACAAAGTC
At1g70840 Fwd	TCTCCAAAGAGATCGATGAACATC
At1g70840 Rev	TGCGTGCTTCATAACATACATATACAC
At1g74710 Fwd	GCGATTAATTGAAGAAAGAGTAACATTTG
At1g74710 Rev	ACAACATTGCTTTCTTATTGTGAGAAC
At2g14610 Fwd	GATCCTCGTGGGAATTATGTGAAC
At2g14610 Rev	TCCTTTATGTACGTGTGTATGCATGA
At2g22780 Fwd	CGTTCCATCGAACCACTTTGT
At2g22780 Rev	TGTTCTCACTTTTATAGCCCAACTACA
At2g29350 Fwd	TGTCCACATGTTCTGTTTCTAAGAGA
At2g29350 Rev	GGTCCATCCGGAGAGTGAACT
At2g31880 Fwd	CAGAGGCCGAACAGTAAAGATGT
At2g31880 Rev	TGCCAAGATTGTGTTCATAGTTATTTC
At2g32680 Fwd	CGCAGAAAGCGTTAGGACTGT
At2g32680 Rev	CGCCTCTGCTGTTTTTCTCAT
At2g33150 Fwd	TTTCACCTCACCTCTCTCTTTTCTG
At2g33150 Rev	TGGTTGTGTGCTTTATTCGAAACT
At2g38770 Fwd	GGGTGATAAGTGAGTTGATGTTTGA
At2g38770 Rev	TTCCTCATGTCATGTATCACTTACCA
At2g42690 Fwd	ATGGAATGGGAAGAAAGGAGAGTT
At2g42690 Rev	AACTCGCATGACTTGTTCACTAATG
At2g47730 Fwd	GCAGTAGTGACCCTCTCTCTTTCTTC
At2g47730 Rev	AAGAGGAGGATGATGATGACATAAATAA
At3g29970 Fwd	CCTTCACTCAACCGTTTCTTCTC
At3g29970 Rev	GAGATACATATAAACGAGGATCTGATAACAA
At5g03090 Fwd	TCTCGACGACTGCAGCAGAT
At5g03090 Rev	AGGCTCATCGGAGGAGAGGTA
At5g03350 Fwd	CATGACATTTGGAGTTGGACCTT
At5g03350 Rev	ATACATTAGAAAAAGCATAACCGAACAG
At5g20230 Fwd	TGCTACTTTTCTGGTCGCTTTTG
At5g20230 Rev	ATACATAACTAGGTTTCGAGTGTGACTCA
At5g49460 Fwd	TGTTGTCATGTATGTTTGTTCACTTCTT
At5g49460 Rev	TGACTCATAGAATCCCAAAAAATGG
At5g64080 Fwd	CCTGAATCTTTTTTCACCAGCTAGA
At5g64080 Rev	CACTCCAAAACATACATAGATTAACCCTAA

5.3 Results

5.3.1 MeJA and arginase transcript accumulation

We conducted semi-quantitative RT-PCR using homozygous *argah1* (SALK 0157987) and *argah2* (SAIL 181_C11) T-DNA insertion lines (Todd et al, unpublished data). The *argah1* line had a T-DNA inserted into the 5' UTR, while in the *argah2* line, the T-DNA was inserted in the second exon. We designed PCR primers that specifically amplified only either *ARGAH1* or *ARGAH2*, and RT-PCR was completed on plants that were treated once with either a MeJA or control solution. Figure 5.2 shows that *ARGAH2*, but not *ARGAH1*, transcript accumulates in response to MeJA. No *ARGAH1* transcript was detected in *argah1* homozygous mutants, and no *ARGAH2* transcript was detected in *argah2* homozygotes. The increased transcript abundance of *ARGAH2* upon treatment with MeJA was observed in both *argah1* and wild-type plants, confirming the specificity of the RT-PCR primers.

In order to quantify the MeJA-induced *ARGAH2* transcript accumulation, we completed quantitative PCR (Figure 5.3). *ARGAH1* expression was lower in the *argah1* mutants compared to the wild-type plants and *ARGAH1* did not increase in any genotype following MeJA treatment. However, *ARGAH2* transcripts were present at greater than 150 fold higher levels in both *argah1* and wild-type MeJA treated plants when compared to wild-type control plants. Thus, *ARGAH2* transcript

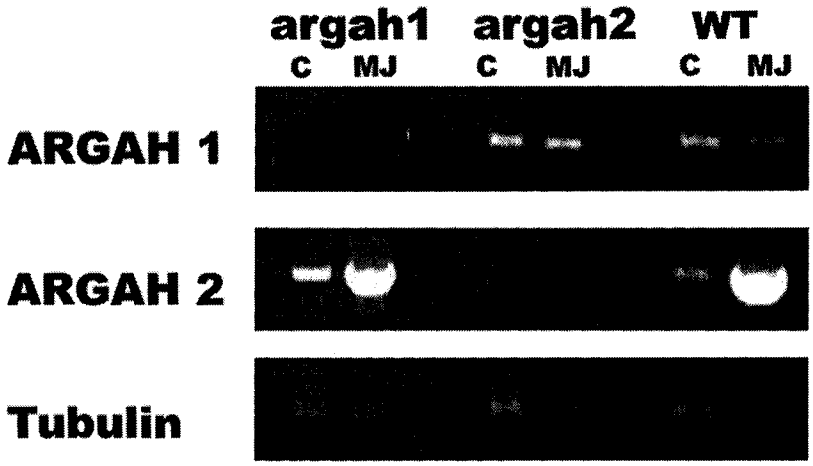


Figure 5.2: RT-PCR of ARGAH1, ARGAH2 and TUBULIN in methyl jasmonate treated and control leaves of wild-type and T-DNA insertion lines (argah1 and argah2) plants. Plants were grown to rosette stage and treated with either methyl jasmonate (MJ) or control (C) solution and covered with a clear lid for 8 hours.

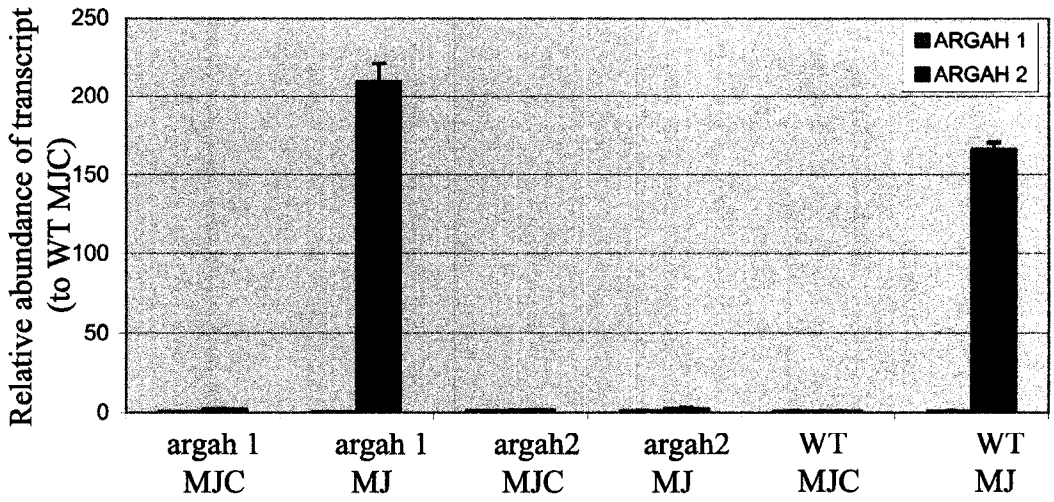


Figure 5.3: Quantitative PCR of ARGAH1 and ARGAH2 in homozygous argah1and argah2 T-DNA lines along with wild-type Arabidopsis in response to methyl jasmonate treatment. Rosette plants were sprayed either with a methyl jasmonate solution (MJ) or control solution (MJC). Transcript is expressed relative to the level of transcript detected in wildtype methyl jasmonate control plants.

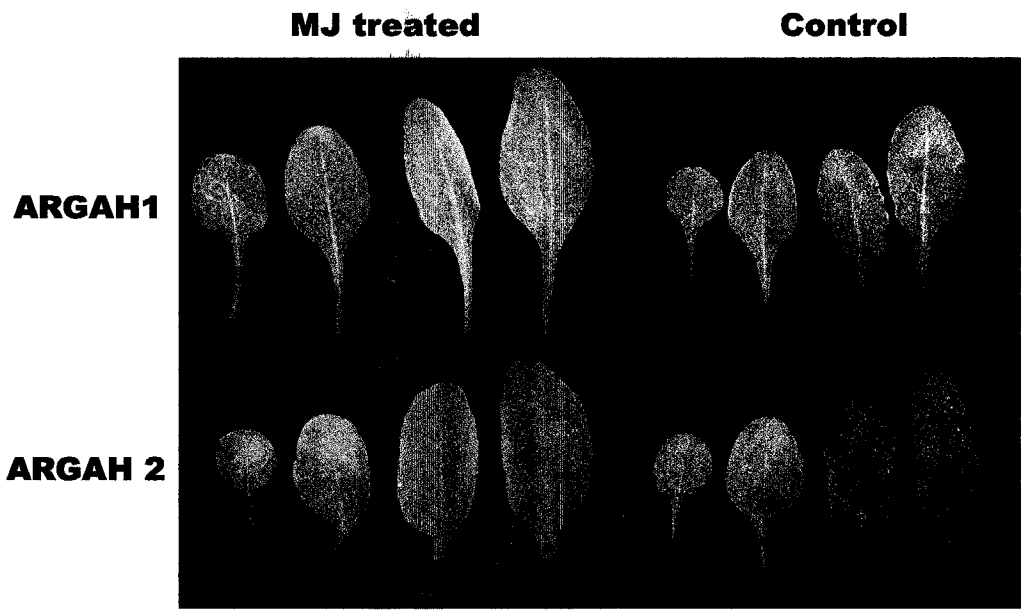


Figure 5.4: GUS staining of methyl jasmonate treated ARGAH1-2 and ARGAH2-2 GUS promoter constructs. Rosette plants, before bolting, were treated 4 times over the course of 8 hours with either MJ or the control solution. Leaves were then harvested and GUS stained. No staining was observed in the leaves of ARGAH 1-2-GUS or ARGAH 2-2-GUS lines.

abundance increased upon treatment with methyl jasmonate while *ARGAH1* transcript abundance did not.

Having demonstrated that *ARGAH2* transcript abundance is increased by MeJA, we wanted to determine whether the putative arginase promoter fragments we previously described contained sufficient cis-elements to confer a MeJA inducible expression pattern. Transgenic plants containing either an approximately 250bp or 1kb upstream fragment from upstream of either the *ARGAH1* or *ARGAH2* start codons fused to the GUS reporter were analyzed. Lines containing the smaller fragments were denoted *ARGAH1-A* or *ARGAH2-A*, while the larger fragments were denoted *ARGAH1-B* and *ARGAH2-B*. T_3 generation plants from each of these constructs were sprayed once with either a MeJA solution or a control solution, however, no MeJA induced GUS staining was detected in any of the plants from 23 independent lines assayed in two replicate experiments. Likewise, we did not observe any staining even when we increased the intensity of the treatment, by treating plants with MeJA 4 times over the course of eight hours (Figure 5.4).

5.3.2 Cis-acting MeJA response elements

To investigate possible explanations for our failure to detect MeJA-inducibility of our reporter gene fusions, we analyzed the sequences of the fragments used in our transgenic constructs, as well as other genomic DNA from further upstream of our fragments. This analysis was conducted using the PlantCare (Lescot et al 2002) and PLACE (Prestridge 1991, Higo et al. 1999) databases to look at the

Table 5.2: The locations of the MeJA response motifs identified in the upstream regions of *ARGAH1* and *ARGAH2*.

Motif	Consensus	*ARGAH1*	*ARGAH2*
MeJA response element	TGACG	-2208	-1501
		-2311	
		-2412	
		-3948	
		-4184	
		-4681	
T/G box	AACGTG	-636	-974
		-642	-1301
		-2417	

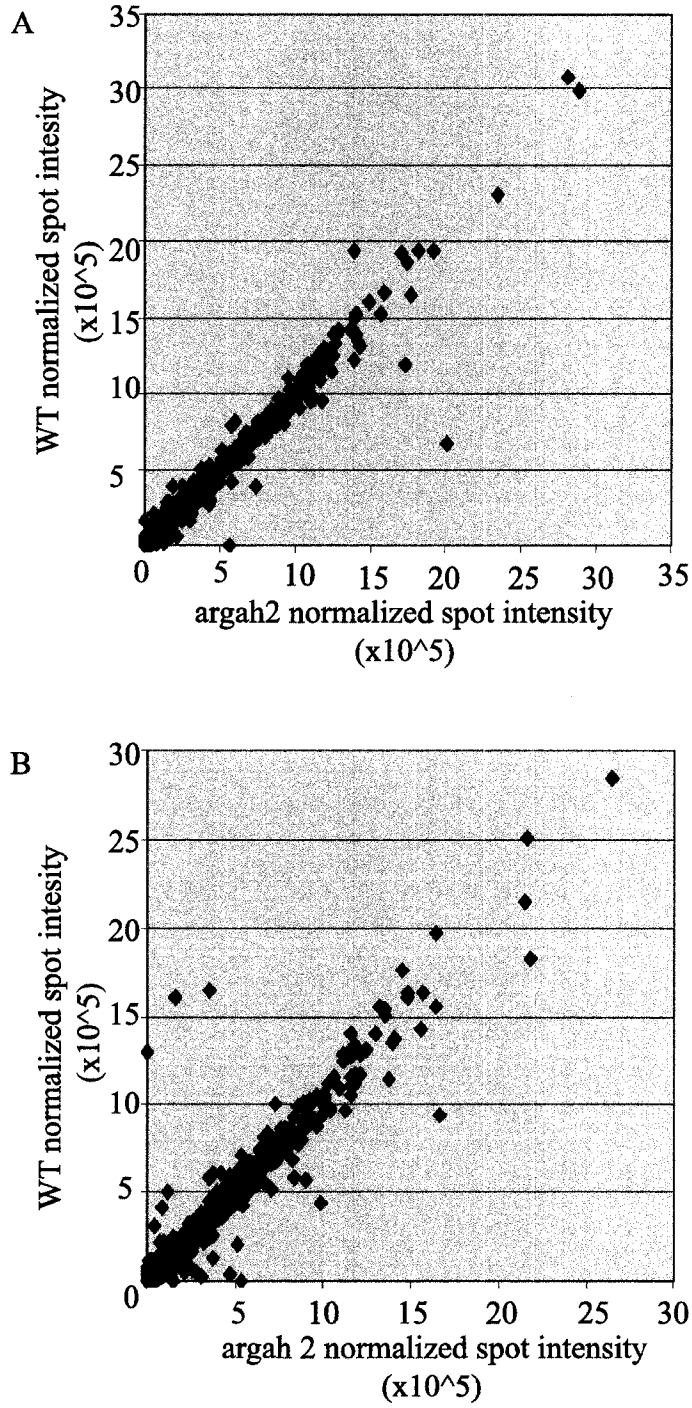


Figure 5.5: Normalized average spot intensities from microarrays. A. Spot intensities of micorarrays comparing argah 2 and wild-type plants treated with a methyl jasmonate control solution. B. Spot intensities of microarrays comparing argah2 and wild-type plants treated with a methyl jasmonate solution.

cis-acting regulatory elements in these regions. PLACE identified one type of MeJA responsive cis-acting element present in our constructs (Table 5.2). A T/G box previously identified in a leucine aminopeptidase gene in tomato (Boter et al. 2004) was present once in the *ARGAH2*-B fragment and twice in *ARGAH1*-B fragment. In addition, an additional copy of this element was present further upstream of the cloned regions of both *ARGAH1*-B and *ARGAH2*-B.

PLACE also identified several copies of another MeJA responsive element in genomic DNA that was not included in our constructs. This motif was previously identified in the promoter of a MeJA responsive barley lipoxygenase gene (Rouster et al. 1997). The length of upstream sequence from the start codon of *ARGAH1* and *ARGAH2* to the start codon of the next gene was 4883bp and 2428bp, respectively. The largest upstream regions present in our promoter-GUS fusions were 1481bp and 1117bp for *ARGAH1* and *ARGAH2,* respectively. There were six putative MeJA response elements present upstream of *ARGAH1* and one upstream of *ARGAH2* (Table 5.2). The *ARGAH1* elements were located 2208bp, 2311bp, 2412bp, 3948bp, 4184bp and 4681bp upstream of the start codon while the single *ARGAH2* MeJA response element was located 1501bp upstream of the start codon. Thus, due to the presence of the MeJA response elements farther upstream it was possible that the arginases of Arabidopsis were responsive to MeJA but the regions required for this were not present in the constructs we used for GUS staining.

5.3.3 The identification of differentially expressed genes through microarrays

To further characterize a possible role for *ARGAH2* in MeJA-mediated defense responses, we conducted a microarray experiment to identify MeJA-inducible transcripts that were differentially expressed in the presence or absence of normal *ARAGAH2* function (Figure 5.5). While we recognized that it was unlikely that arginase was a major determinant of transcript abundance within a regulatory network, we wanted to determine which (if any) genes would be affected by the loss of *ARGAH2*, to provide insight on the role of this enzyme in the cellular response to MeJA treatment. Each microarray slide contained oligonucleotide probes for over 27,000 predicted Arabidopsis genes. In each hybridization experiment involving these slides, we compared mutant (*argah2*) and wild-type plants that were either treated with MeJA or the control solution. We completed two biological replicates with a dye reversal for each replicate. The dye reversal was peformed to reduce the possibility of bias due to dye specific labeling efficiency or fluorescent dye measurement. Upon completion of the microarrays we normalized the data (Figure 5.1) and conducted SAM (significance analysis for microarrays, Tusher et al. 2001) analysis using the MeV program (Saeed et al. 2003) to determine which genes were differentially expressed. A one-class SAM test on the MeJA treated plants indicated that four genes were significantly differentially expressed in the mutant compared to the wild-type in response to MeJA (Table 5.3). Of the genes identified, two had decreased expression in the mutant relative to the wild-type in the presence of MeJA: At4g08870 (*ARGAH2*) and At2g14610. At2g14610 is a pathogenesis related gene

Table 5.3: Microarray and qPCR data of wild-type and *argah 2* plants treated with either a methyl jasmonate solution or a control solution. Data is expressed as a log 2 ratio of mutant (argah2) versus wild-type transcript abundance. A. Genes (10) were identified in the t-test because they were the most negatively differentially expressed. They had the lowest level of transcript abundance in *argah 2* plants compared to wild-type plants both treated with methyl jasmonate. B. Genes (9) were arbitrarily chosen to validate the microarray. C. Genes (3) were identified in the t-test because they were the most positively differentially expressed. They had the highest level of transcript abundance in argah 2 plants compared to wild-type plants both treated with methyl jasmonate.

		Microarray	Microarray	qPCR	qPCR
A		**Control**	**MJ**	**Control**	**MJ**
At2g14610	pathogenesis-related protein 1 (PR-1)	-0.59	-1.24	-0.33	-1.06
At2g32680	disease resistance protein family	-0.37	-1.16	-0.70	-0.74
At4g35770	senescence-associated protein sen1	-0.44	-0.98	-0.54	-1.14
At2g29350	putative short-chain dehydrogenase/reductase	0.03	-0.91	0.13	-0.47
At5g20230	plastocyanin-like domain containing protein	-0.58	-0.79	-0.41	-0.49
At4g27850	proline-rich protein family	-0.62	-0.87	0.14	0.20
At2g31880	leucine-rich repeat transmembrane protein	-0.40	-0.74	-0.30	-0.60
At1g33960	AIG1	-1.12	-0.73	-1.32	-1.26
At1g14880	expressed protein	-0.64	-0.72	-0.79	-0.78
At5g03350	expressed protein	-0.64	-0.69	-0.95	-0.77
B					
At1g74710	isochorismate synthase 1	-0.55	-0.57	-0.47	-0.39
At2g42690	putative lipase	-0.10	-0.41	0.27	-0.20
At4g12490	protease inhibitor/seed storage	-0.15	-0.40	-0.17	-1.13
At2g47730	putative glutathione transferase	-0.30	-0.38	-0.09	-0.01
At2g33150	putative acetyl-CoA C-acyltransferase	-0.11	-0.26	0.20	-0.12
At2g22780	putative malate dehydrogenase	-0.07	0.35	-0.05	-0.08
At3g29970	germination related protein	0.40	0.54	0.95	0.67
At5g49460	ATP citrate lyase	0.41	0.61	0.22	0.16
At5g64080	protease inhibitor/seed storage	0.29	0.62	-0.39	0.65
C					
At1g10640	putative polygalacturonase	0.48	0.85	0.63	1.27
At2g38770	expressed protein	0.15	0.93	0.25	0.27
At5g03090	hypothetical protein	1.10	1.24	1.91	3.94

(PR-1) whose function is unknown. SAM also identified two genes that had increased expression in the mutant relative to the wild-type: At5g03090 and At3g24790. At5g03090 has an unknown function and At324790 is predicted to be a kinase and has sequence similarity to a serine threonine kinase from rice (*Oryza sativa*) (TAIR, www.arabidopsis.org).

We had completed microarrays comparing mutant (*argah2*) and wild-type plants in order to determine if there were genes that under non-inducing conditions were affected by the loss of *ARGAH2.* A second one-class class SAM test was completed on micorarrays comparing mutant and wild-type plants, and only one gene was found to bedifferentially expressed at a significant level. At2g38350 is predicted to be a mitochondrial protein; however, its function is also unknown. In a third SAM analysis a two-class unpaired SAM was completed, this would identify any genes that were differentially expressed (between the mutant and wild-type plants) in both the control and MeJA treated conditions. *ARGAH2* (At4g08870) was the only gene identified as being differentially expressed (mutant versus wild-type) in both treated and control plants.

Based on criteria including the possible relevance to MeJA, we performed quantitative PCR on transcripts from two of four of the genes identified from the SAM analysis to confirm the pattern of transcript expression observed. In the microarrays it was observed that At2g14610 had lower expression in the mutant as compared to wild-type, in both the presence and absence of MeJA, although the difference between mutant and wild-type expression was greatest when plants were treated with MeJA (Figure 5.6). This pattern was also observed in the quantitative

PCR of At2g14610. Furthermore, it was clear from the qPCR that in the mutant, transcript abundance of this gene decreased following MeJA treatment. For At5g03090, both microarray and qPCR showed higher transcript abundance in mutants as compared to wild-type, under both control and methyl jasmonate conditions (Figure 5.7).

Because SAM identified very few genes as significantly differentially expressed in *argah2* mutants as compared to wild-type, we utilized a second type of statistical test to select other genes whose transcript abundance we could then also measure by qPCR, to provide a further evaluation of the accuracy of our microarray data. A T-test of MeJA treated data using an α of 0.05 was used to identify genes candidate genes for this purpose. From this list we chose 20 genes, based either on their annotations or on the magnitude of their mutant versus wild-type transcript abundance ratios, in order to confirm the pattern of expression we observed in the microarray using qPCR (Table 5.3). We chose an additional 9 genes that had the largest degree of decreased expression in the mutant relative to the wild-type (in MJ treated plants) and an additional two genes that had the greatest increase in expression in the mutant relative to the wild-type. In addition, we also arbitrarily chose nine genes that were less significantly expressed (positively and negatively) in MeJA mutant versus wild-type plants. Figures 5.6 and 5.7 show several of these additional genes that had the largest degree of differential expression, both positive and negative, when comparing methyl jasmonate treated mutant plants to wild-type plants. In total, we completed quantitative PCR for 22 genes. Table 5.3 shows microarray and qPCR data that was the average of two biological replicates. When

comparing the values 12 of 22 genes were consistent with the microarray data. That is, when comparing control and MeJA expression ratios (expressed as mutant versus wild-type) they had similar patterns. Of the 12 genes whose patterns of microarray and qPCR data was not consistent five genes had at least one value that did not agree with the others. For example, we observed negative ratios for both MeJA and control slides in the microrarrays for At4g27850; that is the wild-type had higher expression compared to the mutant. However, the qPCR data showed the opposite trend, having positive ratios which indicated higher expression in the mutant relative to the wild-type. This situation was also observed for four other genes.

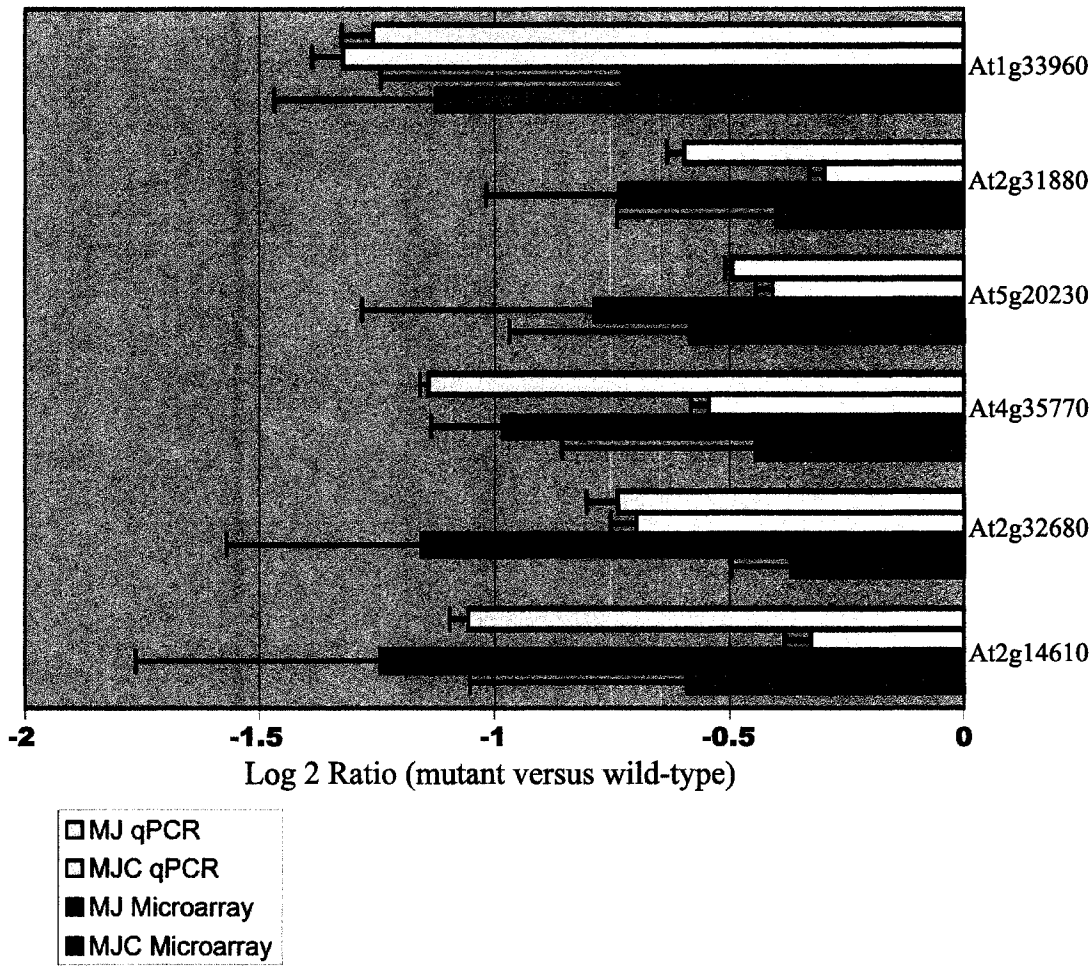


Figure 5.6: Comparison of microarray and quantitative PCR data for genes which are expressed relatively less in argah2 plants than wild-type plants when treated with methyl jasmonate (MJ) or the control solution (MJC). All data is expressed as a log 2 ratio of mutant vs wild-type transcript abundance. The error bars represent the standard error of at least 2 biological replicates.

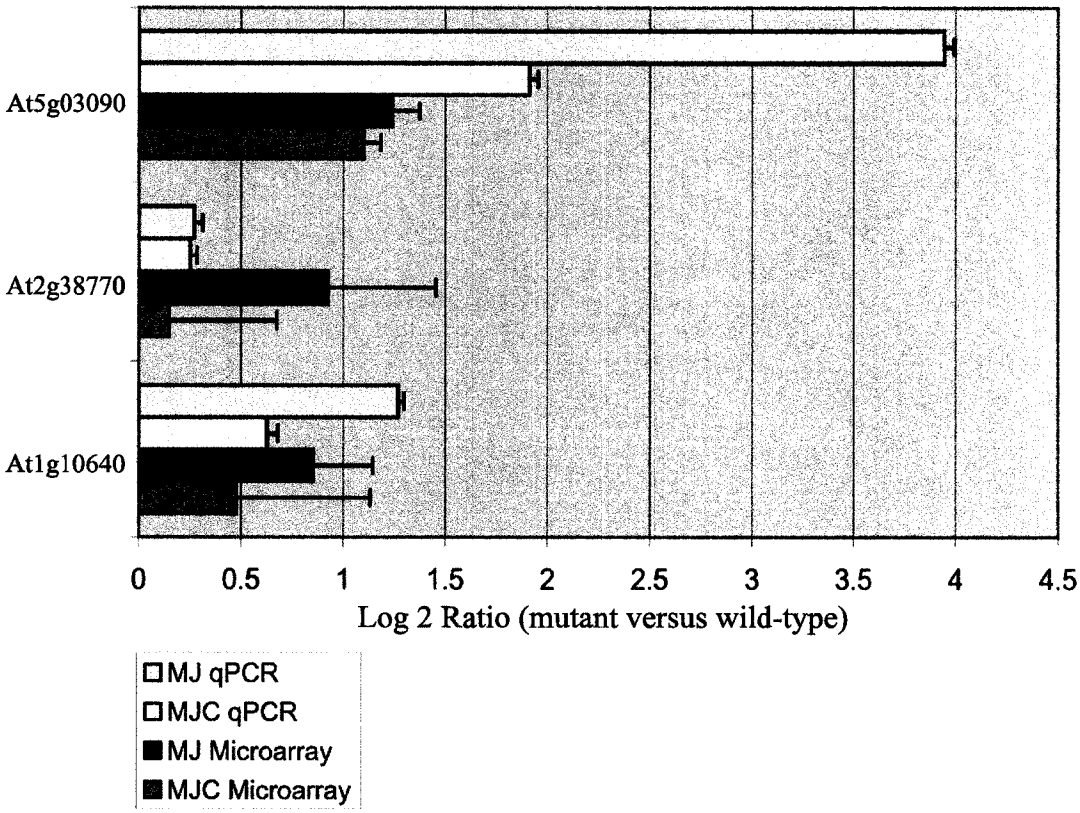


Figure 5.7: Comparison of microarray and quantitative PCR data for genes which are more higly expressed on a relative basis in argah 2 plants than wild-type plants when treated with methyl jasmonate (MJ) or the control solution (MJC). All data is expressed as the log 2 ratio of mutant vs wild-type transcript abundance. The error bars represent the standard error of at least two biological replicates.

5.4 Discussion

5.4.1 Cis-acting MeJA response elements

The lack of MeJA-induced staining in all *ARGAH*-GUS lines, in spite of the observed increase in transcript abundance as measured by qPCR, was explained by the lack of MeJA response elements in the promoter fragments used. Using the PlantCare database (Lescot et al. 2002) two types of MeJA response element was identified in the upstream of the the *ARGAH1* and *ARGAH2* genes (Table 5.2). The first, a motif (TGACG) was previously identified in the promoter of a barley lipoxygenase gene (Rouster et al. 1997). Promoter deletions indicated that a short 35bp region was necessary for the MeJA response; this region contained a pair of complementary sequences CGTCA and TGACG. Site directed mutagenesis (TGACG to TTCCT) of either copy or both copies of this motif resulted in the loss of the MeJA response (Rouster et al. 1997). However, in the upstream sequences of the Arabidopsis arginases none of the MeJA response elements were present as inverted repeats. Inverted repeats appear to be an essential feature of this motif as inserting a DNA fragment between the motifs strongly decreased the MeJA inducibility (Rouster et al. 1997). Thus, it appears that these motifs, which are believed to be bZIP binding motifs, must be present as inverted repeats to be functional, making it unlikely that they are involved in the MeJA response of *ARGAH2*. In addition, the presence of six copies of this element in the upstream region of the non-inducible *ARGAH1* but only

one copy in the upstream region of the MeJA inducible *ARGAH2* made it even less likely that this element was responsible for the MeJA response of *ARGAH2*.

The second type of MeJA response element identified using the PLACE database (Prestridge 1991, Higo et al. 1999) was a T/G box previously identified in a leucine aminopeptidase (LAP) gene in tomato (Boter et al. 2004). This element is bound by bHLH-leucine zipper proteins JAMYC2 and JAMYC10 and mutation of this element decreases LAP expression (Boter et al. 2004). However, similar to the previously discussed motif (TGACG) there is a discrepancy between element copy number and MeJA induced transcript accumulation. The *ARGAH1* upstream contains three copies of this element, while the MeJA responsive arginase (*ARGAH2*) only contains two. Thus, the relative numbers of each type of element in *ARGAH1* and *ARGAH2* does not explain the difference in MeJA induction.

5.4.2 Induction of arginase

We have shown that *ARGAH2* transcript accumulates in response to MeJA (Figures 5.2 and 5.3) and this was supported by the GENEVESTIGATOR response viewer (Zimmermann et al. 2004) and our microarrays (see below). In addition, GENEVESTIGATOR indicated that *ARGAH2* responds not only to wounding (see Chapter 4) and MeJA but also to *Pseudomonas syringe,* the fungus *Botrytis cinerea* and to a lesser degree to the fungus *Erysiphe orontii*. It would be interesting to confirm the induction of arginase in response to these pathogens.

5.4.3 JA signaling

In tomato, Chen et al. (2004) showed that the MeJA induction of *LeARG2* was dependant on the JA (jasmonic acid) signaling pathway. The authors showed that a mutant deficient in *COI1* (*jai1*) was not able to elicit an induction in response to wounding or MeJA treatment in *LeARG2*. In addition, less arginase activity was present in *jai1* plants compared to wild-type plants and the authors concluded that basal arginase activity in tomato leaves is regulated by this pathway (Chen et al. 2004). *COI1* is important in JA signaling, it is predicted to be a functional ubquitin ligase and to target JA repressors (Devoto and Turner 2003, Devoto et al. 2005). In Arabidopsis, *COI1* was shown to be an important signaling molecule, essential for the expression of approximately 84% of the genes induced by JA (Devoto et al. 2005). As *COI1* appears to have similar functions in tomato and Arabidopsis (Boter et al. 2004) we may expect that the arginases would behave in the same manner as those of tomato; however, this would be interesting to investigate.

5.4.4 The identification of differentially expressed genes through microarrays

In our microarray experiments we attempted to identify genes which were differentially expressed in mutant plants (*argah2*) compared to wild-type plants. We looked for these differentially expressed genes under both control conditions and in plants that had been treated with MeJA. Microarray analysis revealed that there were relatively few genes that were differentially expressed in the *argah2* mutant relative

to the wild-type in either in MeJA treated plants or untreated controls. This suggests that arginase may be a relatively far downstream component of the MeJA response pathway. In comparison, *COI1* is required for virtually all JA-dependant responses (Devoto and Turner 2005) and is believed relatively close to the top of the regulatory hierarchy in the JA signaling pathway. Microarrays using *coi1* and wild-type plants identified approximately 170 genes that required *COI* and each of these genes was differentially expressed by at least 2.5 fold (Devoto et al. 2005).

SAM analysis of our microarrays identified a total of five differentially expressed genes, one of which was *ARGAH2*. We confirmed that At5g03090 (Figure 5.6) had higher expression in the mutant compared to the wild-type in both MeJA treated and control plants, but there was a larger difference in the expression in MeJA treated plants. An even greater difference between mutant and wild-type plants was observed in MeJA plants in the qPCR data. This difference resulted because, upon treatment with MeJA, At5g03090 decreased expression in wild-type plants but there was little change in *argah2* plants. Thus, it appears that in response to MeJA, At5g03090 normally decreases in expression, but this does not occur when *ARGAH2* is not functional. At present there is no known function of At5g03090 (www.arabidopsis.org). BLAST analysis indicated that the cDNA was similar to another protein of unknown function (At1g53480, E value 7e-96) and to a predicted DNA binding protein (At1g53490, E value 7e-96). The predicted protein was 97 amino acids in length and has a predicted pI of 7.25 (www.arabidopsis.org). The protein sequence did not contain a chloroplast transit peptide, and the probability of the At5g03090 protein to be exported to the mitochondria was 0.11

(www.expasy.org). A Kyte and Doolittle (1982) hydrophobicity plot was generated (www.expasy.org) and 2 hydrophobic regions (>+1.8) at the C terminus of the protein were identified. Thus, it is possible that At5g03090 is an integral membrane protein with two potential transmembrane domains.

The second gene we identified by microarray and qPCR analysis as being dependent on *ARGAH2* function was At2g14610 (Figure 5.5). At2g14610 is annotated as a pathogenesis related 1 protein (PR-1). In both our microarray and qPCR data we observed that MeJA did not significantly alter At2g14610 transcript abundance (compared to the control) in wild-type plants. Mutant (*argah2*) control plants also did not have significantly altered transcript abundance compared to the wild-type control and wild-type MeJA plants. Mutant plants treated with MeJA however, had a significantly decreased level of At2g14610 transcript. Thus, it appears that the loss of *argah2* only has an effect on PR1 when the plants are treated with MeJA.

PR proteins are frequently induced as part of the hypersensitive response (HR) and systemic acquired resistance (SAR) (Klessig and Malmy 1994). In tobacco, there are several PR gene families which have increased transcript accumulation in response to the tobacco mosaic virus (TMV) and to salicyclic acid (SA) (Ward et al. 1991). PR-1 proteins exist in two isoforms: acidic and basic (Santamaria et al. 2001). In tobacco, the acidic form of PR-1 is induced by SA and inhibited by JA whereas the basic form is induced by JA and inhibited by SA (Niki et al. 1998). The Arabidopsis protein (At2g14610) is an acidic PR-1 isoform as the proteins were originally classified by their relative electrophoretic mobility under non-denaturing conditions

(Jamet and Fritig 1986) and not by their predicted pI values. Our results indicate that its transcript accumulation is not induced by MeJA in wild-type plants. In addition, Arabidopsis PR-1 has been shown to be induced response to SA (Lebel et al. 1998). Thus, the Arabidopsis and tobacco the acidic PR-1 isoforms appear to be regulated in a similar manner. However, the role of *ARGAH2* in At2g14610 expression is unclear; it would be interesting to look at transcript levels of At2g14610 in both wild-type and *argah2* plants treated with SA. In addition, it would be interesting to determine if SA has an effect on *ARGAH2* expression alone and in combination with MeJA.

We have demonstrated that MeJA induces accumulation of *ARGAH2* transcripts in Arabidopsis. This leads to the question of the function of arginase in the plant response to MeJA. Chen et al. (2004) proposed that arginase may be involved in producing polyamines in JA treated tissues. However, in Arabidopsis, an *ODC* gene has not been identified, and its existence remains controversial (Ferier at al. 1997, Hanfey et al. 2001, Tassoni et al. 2003). Thus, in the absence of *ODC*, the relevance of arginase to polyamine synthesis is unclear. In a recent report Chen et al. (2005) showed a more direct role for arginase in plant defense; in tobacco plants that induce *LeARG2* in response to wounding the arginase enzyme is active in the mid-guts of tobacco hornworms who feed on these plants. This increased mid-gut arginase activity decreases the amount of free Arg available and acts in an antinutritive capacity on the larvae (Chen et al. 2005). Thus, the arginase produced by the plant in response to the herbivore appears to have a direct effect on the herbivore.

Other possible roles for *ARGAH2* in MeJA-mediated defense responses include providing substrates for proline synthesis. In mammals, proline derived from

ornithine is important in the production of collagen in response to wounding (Witte and Barbul 2003). In plants, proline does accumulate in response to various stresses (Delauney and Verma 1993) and proline rich proteins have been shown to accumulate in response to wounding (Suzuki et al. 1993). However, to our knowledge the direct importance of proline in response to wounding has not been demonstrated.

We have shown that MeJA does affect arginase transcript accumulation, specifically in *ARGAH2*. *ARGAH2* in turn appears to have an effect the level of transcript accumulation of several genes. Microarrays identified PR-1 (At2g14610) as being differentially expressed in response to the loss of *ARGAH2* when plants were treated with MeJA. However, there was no significant difference in wild-type treated and control plants. Thus, it appears that PR-1 transcript accumulation is mediated at least in part by *ARGAH2* in MeJA treated plants.

5.5 References

Alabadí D, Agüero MS, Pérez-Amador MA, Carbonell J. 1996. Arginase, arginine decarboxylase, ornithine decarboxylase, and polyamines in tomato ovaries. Plant Physiology 112: 1237-1244.

Boter M, Ruiz-Rivero O, Abdeen A, Prat S. 2004. Conserved MYC transcription factors play a key role in jasmonate signaling both in tomato and Arabidopsis. Genes & Development 18: 1577-1591.

Boutin JP. 1982. Purification, properties, and subunit structure of the arginase from iris bulbs. European Journal Biochemistry 127: 237-243.

Brzozowski T, Konturek SJ, Majka J, Dembinski A, Drozdowicz D. 1993. Epidermal growth factor, polyamines, and prostaglandins in healing of stress-induced gastric lesions in rats. Digestive diseases and Sciences 38: 276-283.

Chen H, McCaig BC, Melotto M, He SY, Howe GA. 2004. Regulation of plant arginase by wounding, jasmonate, and the phytotoxin coronatine. Journal of Biological Chemistry 279: 45998-46007.

Chen H, Wilkerson CG, Kuchar JA, Phinney BS, Howe GA. 2005. Jasmonate-inducible plant enzymes degrade essential amino acids in the herbivore midgut. Proceedings of the National Academy of Science 102: 19237-19242.

Cohen Y, Gysi U, Niderman T. 1993. Local and systemic protection against *Phytopthora infestans* induced in potato and tomato plants by jasmonic acid and jasmonic methyl ester. Phytopathology 83: 1054-1062.

Delauney AJ and Verma DPS. 1993. Proline biosynthesis and osmoregulation in plants. The Plant Journal 4: 215-223.

De Ruiter H and Kollöffel C. 1983. Arginine catabolism in the cotyledons of developing and germinating pea seeds. Plant Physiology 73: 525-528.

Devoto A and Turner JG. 2003. Regulation of jasmonate-mediated plant responses in Arabidopsis. Annals of Botany 92: 329-337.

Devoto A and Turner JG. 2005. Jasmonate-regulated Arabidopsis stress signaling network. Physiologia Plantarum 123: 161-172.

Devoto A, Ellis C, Magusin A, Chang HS, Chilcott C, Zhu T, Turner JG. 2005. Expression profiling reveals COI1 to be a key regulator of genes involved in wound- and methyl jasmonate-induced secondary metabolism, defense, and hormone interactions. Plant Molecular Biology 58: 497-513.

Downum KR, Rosenthal GA, Cohen WS. 1983. L-arginine and L-canavanine metabolism in jack bean, *Canavalia ensiformis* (L.) DC. and soybean, *Glycine max* (L.) Merr. Plant Physiology 73: 965-968.

Feirer RP, Hocking KL, Woods PJ. 1997. Involvement of arginine decarboxylase in the response of *Arabidopsis thaliana* to osmotic stress. Journal of Plant Physiology 153: 733-738.

Hanfrey C, Sommer S, Mayer MJ, Burtin D, Michael AJ. 2001. *Arabidopsis* polyamine biosynthesis: absence of ornithine decarboxylase and the mechanism of arginine decarboxylase activity. Plant Journal 27: 551-560.

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

Jamet E and Fritig B. 1986. Purification and characterization of 8 of the pathogenesis-related proteins in tobacoo leaves reacting hypersensitively to tobacco mosaic virus. Plant Molecular Biology 6: 69-80.

King JE and Gifford DJ. 1997. Amino acid utilization in seeds of loblolly pine during germination and early seedling growth. I. Arginine and arginase activity. Plant Physiology 113: 1125-1135.

Klessig DF and Malmy J. 1994. The salicyclic acid signal in plants. Plant Molecular Biology 26: 1439-1458.

Kollöfeel C and van Duke HD. 1975. Mitochondrial arginase activity from cotyledons of developing and germinating seeds of *Vicia faba* L. Plant Physiology 55: 507-510.

Krumpelman PM, Freyermuth SK, Cannon JF, Fink GR, Polacco JC. 1995. Nucleotide Sequence of *Arabidopsis thaliana* Arginase Expressed in Yeast. Plant Physiology 107: 1479-1480.

Kyte J and Doolittle RF. 1982. A simple method for displaying the hydropathic character of a protein. Journal of Molecular Biology 157:105-132.

Lebel E, Heifetz P, Thorne L, Uknes S, Ryals J, Ward E. 1998. Functional analysis of regulatory sequences controlling PR-1 gene expression in Arabidopsis. The Plant Journal 16: 223-233.

Lescot M, Déhais P, Moreau Y, De Moor B, Rouzé P, Rombauts S. 2002. PlantCare: a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Research 30: 325-327.

Livak KJ and Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C}{}_T$ method. Methods 25: 402-408.

Micallef BJ and Shelp BJ. 1989. Arginine metabolism in developing soybean cotyledons. I. Relationship to nitrogen nutrition. Plant Physiology 90: 624-630.

Niki T, Mitsuhara I, Seo S, Ohtsubo N, Ohashi Y. Antagonistic effect of salicyclic acid and jasmonic acid on the expression of pathogenesis-related (PR) protein genes in wounding mature tobacco leaves. Plant and Cell Physiology 39: 500-507.

Pauw B and Memelink J. 2005. Jasmonate-responsive gene expression. Journal of Plant Growth and Regulation 23: 200-210.

Perez-Amador MA, Leon J, Green PJ, Carbonell J. 2002. Induction of the arginine decarboxylase *ADC2* gene provides evidence for the involvement of polyamined in the wound response in Arabidopsis. Plant Physiology 130: 1454-1463.

Prestridge DS. 1991. SIGNAL SCAN: A computer program that scans DNA sequences for eukaryotic transcriptional elements CABIOS 7: 203-206.

Rouster J, Leah R, Mundy J, Cameron-Mills V. 1997. Identification of a methyl jasmonate-responsive region in the promoter of a lipoxygenase 1 gene expressed in barley grain. Plant Journal 11:513-23.

Saeed AI., Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, Sturn A, Snuffin M, Rezantsev A, Popov D, Ryltsov A, Kostukovich E, Borisovsky I, Liu Z, Vinsavich A, Trush V, Quackenbush J. 2003. TM4: A free, open-source system for microarray data management and analysis. Biotechniques 34: 374-378.

Santamaria M, Thompson CJ, Read ND, Loake GJ. 2001. The promoter of a basic PR-1 like gene, *AtPTB1*, from Arabidopsis establishes an organ-specific expression pattern and responsiveness to ethylene and methyl jasmonate. Plant Molecular Biology 47: 641-652.

Splittoesser WE. 1969. The appearance of arginine and arginase in pumpkin cotyledons. Characterization of arginase. Phytochemistry 8: 753-758.

Suzuki H, Wagner T, Tierney ML. 1993. Differential expression of two soybean (*Glycine max* L.) proline-rich proteins after wounding. Plant Physiology 101: 1283-1287.

Tassoni A, Fornale S, Bagni N. 2003. Putative ornithine decarboxylase activity in Arabidopsis thaliana: inhibition and intracellular localization. Plant Physiology and Biochemistry 41: 871-875.

Taylor AA and Stewart GR. 1981. Tissue and subcellular localization of enzymes of arginine metabolism in *Pisum sativum*. Biochemical and Biophysical research communications 101: 1281-1289.

Tusher VG, Tibshiranit R, Chu G. 2001. Significance analysis of microarrays applied to the ionizing radiation response. Proceedings of the National Academy of Sciences 98: 5116-5121.

Ward ER, Uknes SJ. Williams SC, Dincher SS, Wiederhold DL, Alexander DC, Ahl-Goy P, Métrauz JP, Ryals JA. 1991. Coordinate gene activity in response to agents that induce systemic acquired resistance. The Plant Cell 3: 1085-1094.

Witte MB and Barbul A. 2003 Arginine physiology and its implications for wound healing. Wound repair regulation 11: 419-423.

Wright LC, Brady CJ, Hinde RW. 1981. Purification and properties of the arginase from Jerusalem artichoke tubers. Phytochemistry 20: 2641-2645.

Zhao FG and Qin P. 2004. Protective effect of exogenous polyamines on root tonoplast function against salt stress in barley seedlings. Plant Growth Regulation 42: 97-103.

Zimmermann P, Hirsch-Hoffmann M, Hennig L, and W Gruissem. 2004. GENEVESTIGATOR: Arabidopsis Microarray Database and Analysis Toolbox Plant Physiology 136: 2621-2632

Chapter 6: Conclusions

In this study we primarily investigated arginase in both Arabidopsis (*Arabidopsis thaliana*) and loblolly pine (*Pinus taeda* L.). In pine, arginase acts to degrade storage proteins to provide the seedling with nutrients until it is photosynthetic. Storage proteins are 1 of the 2 major types of storage reserves found in the pine seed, the other being triacylglycerols (TAGs). Somatic embryos are fundamentally different from their zygotic counterparts, lacking the maternally derived haploid megagametophyte which is where the majority of the seed storage reserves are located (Stone and Gifford 1997, Stone and Gifford 1999). In an analysis of a line of loblolly pine somatic embryos we found that they contained significantly lower levels of TAGs but higher levels of storage proteins relative to zygotic embryos. With respect to the storage proteins we found that the accumulation of storage proteins of the somatic embryo was biased towards soluble proteins. The somatic embryos contained 5 fold more soluble than insoluble protein whereas mature zygotic embryos contained approximately 1.5 fold more soluble than insoluble. Thus, the somatic embryo does not contain the same storage reserves as the zygotic embryo and this may affect the viability of the seedlings.

In order to understand how the seed storage reserves are mobilized and what is involved in the regulation we focused on the arginase enzyme. Arginine (Arg) is abundant in loblolly seed storage proteins contributing almost 50% of the total nitrogen (N) to the buffer insoluble fraction (King and Gifford 1997). Thus, the mobilization of Arg provides the seedling with a significant proportion of N during early seedling growth. Arginase begins to increase in transcript abundance, protein

level and specific activity following germination and all peak before $12DAI_{30}$ (12 days after imbibition at 30°C) when the radicle is between 12-68mm in length (King and Gifford 1997, Todd et al. 2001). In our attempt to isolate the region upstream of the arginase gene we isolated 2 products and an intron in the 5' untranslated region (UTR) of the arginase gene. We completed PCR on individual megagametophytes in order to attempt to determine if the regions isolated represented individual genes or alleles of the same gene. We determined that these regions either represent 2 individual genes that were tightly linked (0% recombination frequency) or were alleles of a single gene. Other plant species, including Arabidopsis, have multiple copies of the arginase gene; however, we were unable to determine if this was also true for loblolly pine. We also identified several repeated regions that were present in the upstream regions. Unfortunately, transient expression assays using each of the regions isolated showed only very low levels of reporter gene (GUS) in seedlings which arginase is actively being transcribed (Todd et al. 2001).

We also isolated the regions upstream of both the Arabidopsis arginase genes. Transgenic Arabidopsis plants containing either 250bp or approximately 1kb of upstream sequence from the Arabidopsis arginases were generated. GUS staining of $7DIC_{22}$ (7 days in culture at 22°C) seedlings indicated that the 250bp regions were sufficient to confer a similar pattern of staining. We observed staining in the vasculature of the roots and cotyledons of both *ARGAH1*-GUS and *ARGAH2*-GUS plants (all 4 construct sizes). In addition, we also observed staining in young leaves. Staining of young leaves was also observed in rosettes and in plants that had bolted: in young leaves of shoots emerging from axils of both the rosette and cauline leaves.

These patterns of staining were observed for the 250bp upstream fragments of both *ARGAH1* and *ARGAH2*.

We also observed a difference in the flower staining pattern conferred by the different sized upstream regions. Very little flower staining was present in plants containing the smallest *ARGAH1* fragment, but there was increased staining in plants containing the longer *ARGAH1* upstream fragment. This included *ARGAH1* staining in the pollen. *ARGAH1* transcript but not *ARGAH2* transcript was present in the pollen of wild-type Arabidopsis plants. Arginase is methyl jasmonate (MeJA) inducible and as MeJA plays a role in pollen development it was possible that MeJA mediated arginase in pollen. However, we observed that while *ARGAH1* had a high level of transcript in pollen it was *ARGAH2* that increased in transcript abundance in response to MeJA. Thus, it is unlikely that arginase expression in pollen is mediated through MeJA.

In order to understand what the effect the loss of *ARGAH2* had on other genes, especially in the response to MeJA, we conducted microarray analysis. We identified a total of 4 genes that were differentially expressed when comparing wild-type and *argah2* plants. We focused on 2 of these genes At2g14610 and At5g03090 and completed quantitative PCR to verify their expression patterns. At5g03090 expression was higher in the mutant when compared to the wild-type both in MeJA and control plants. At5g03090 was a hypothetical protein with an unknown function but a hydrophobicity plot indicated that it had 2 potential transmembrane domains but likely was not targeted to mitochondria or the chloroplast. At2g14610 was also identified as differentially expressed in the mutant compared to the wild-type in

response to MeJA. It appeared that the loss of *argah2* only had an effect on PR-1 when the plants were treated with MeJA. At2g14610 was annotated as a PR-1 protein and was predicted to be a basic isoform. In tobacco, the acidic form of PR-1 is induced by SA and inhibited by JA whereas the basic form is induced by JA and inhibited by SA (Niki et al 1998). However, our data show that PR-1 is negatively affected by MeJA. Other experiments on At2g14610 in wild-type and *argah2* plants including SA treatment may increase our understanding of this situation.

Combined, we have attempted to investigate arginase through its upstream regions in both Arabidopsis and loblolly pine. We have shown that arginase is actively transcribed in pollen and in response to wounding and MeJA. Thus, arginase may have a role in pollen development and in the pathogen response. The role of arginase in the pathogen response pathway was further supported by the differential expression of PR-1 in response to the loss of *ARGAH2* in MeJA treated seedlings. Future experiments could determine if there is increased arginase activity and protein accumulation in these processes and attempt to identify the role of arginase. It would also be interesting to determine if arginase is inducible in response to SA both in wild-type and in some of the signaling mutants. Thus, there are still many unanswered questions pertaining to the roles of arginase in plant development.

6.1 References

King JE and Gifford DJ. 1997. Amino acid utilization in seeds of loblolly pine during germination and early seedling growth. I. Arginine and arginase activity. Plant Physiology 113: 1125-1135.

Niki T, Mitsuhara I, Seo S, Ohtsubo N, Ohashi Y. Antagonistic effect of salicyclic acid and jasmonic acid on the expression of pathogenesis-related (PR) protein genes in wounding mature tobacco leaves. Plant and Cell Physiology 39: 500-507.

Stone SL and Gifford DJ. 1997. Structural and biochemical changes in loblolly pine (*Pinus taeda* L.) seeds during germination and early-seedling growth. I. Storage protein reserves. International Journal of Plant Science 158: 727-737.

Stone SL and Gifford DJ. 1999. Structural and biochemical changes in loblolly pine (*Pinus taeda* L.) seeds during germination and early-seedling growth. II. Storage triacylglycerols and carbohydrates. International Journal of Plant Science 160: 663-671.

Todd CD, Cooke JEK, Mullen RT, Gifford DJ. 2001. Regulation of loblolly pine (*Pinus taeda L.*) arginase in developing seedling tissue during germination and post-germinative growth. Plant Molecular Biology 45: 555-565.