Chapter 1

BACKGROUND

1.1 Historical importance of the opium poppy

The opium poppy has been a part of the culture of man dating back to 3400 B.C. It was cultivated in lower Mesopotamia by the Sumerians, who referred to the plant as "Hul Gil", the joy plant. In 1300 B.C. the opium trade began to flourish during Egyptian rule, and it was traded as far as Europe. In 460 B.C. Hippocrates, the father of modern medicine, acknowledged the usefulness of opium as a styptic and a narcotic (Booth, 1998). Linnaeus classified the opium poppy in 1753, as *Papaver somniferum,* meaning sleep inducer. Opium is the dried exudate from the capsule of an opium poppy (Bryant, 1988). Fredrich Sertuerner discovered the first active ingredient of opium to be the alkaloid morphine in 1803 (Kutchan, 1995; Booth, 1998). Other important phenanthrene alkaloids that are produced by opium poppy are codeine and its precursor thebaine. These poppy alkaloids are associated with specialized cells called laticifers (Fahn, 1979).

Today, the pharmaceutical poppy alkaloids are extracted from the dried laticifer cytoplasm, dried capsules, or poppy straw (Bryant, 1988). Semi-synthetic compounds, such as the illicit drug heroin, are also produced from opium alkaloids (Bryant, 1988). Only certain countries are allowed to legally grow and harvest opium poppy for pharmaceutical purposes, seed or oil, and these countries have been designated by the United Nations (Nessler, 1990).

1.2 Laticifers

Laticifers are a group of specialized cells that contain an emulsion known as latex, Figure 1.1 (Fahn, 1997). Latex harbors the morphiniane alkaloids of pharmacological relevance. Laticifers occur in leaves, stems, roots, and flowers, but are limited to vascular tissue associated with the phloem (Nessler and Mahlberg, 1976). Though present in many species, laticifers are restricted to approximately 20 plant families (Fahn, 1997). They are found in families with little taxonomic relationship suggesting that the adaptation of latex production has occurred more than once in plant evolution (Fahn, 1997).

There are two major classes of laticifers: the articulated and the non-articulated, Figure 1.2 (Esau, 1965). Articulated laticifers are made up of a series of cells that become laticiferous vessels when cell walls become porous and break down. Non-articulated laticifers develop from a single cell which elongates and branches (Fahn, 1997). The opium poppy has articulated laticifers (Esau, 1965). Laticifer cells can be seen in young *Papaver somniferum* seedlings when the root has emerged about 2 mm, and these cells differentiate soon after the appearance of sieve elements but before the tracheary elements of the xylem (Fahn, 1979). The complex system of latex vessels in the poppy capsule has been studied, and has been found to become denser as the capsule develops (Fahn, 1997). These metabolically active latex vesicles not only contain alkaloids, but proteins such as the major latex proteins, that are associated with the laticifers of opium poppy (Nessler *et al.*, 1985).

1.3 Secondary metabolites

Secondary metabolites are organic compounds that do not seem to be associated directly with plant growth and development. Secondary metabolites are also known as natural products, due to the indistinct classification between primary and secondary compounds (Buchanan *et al.*, 2000). Another approach of describing natural (secondary) products is that these compounds influence ecological interactions between a plant and its environment (Verpoorte *et al.*, 2000). Many of these metabolites have been shown to have an important role in plant protection from herbivory and microbial infections (Buchanan *et al.*, 2000).

Secondary compounds have been of interest to humans form many years due to their many uses such as; medicines, dyes, and their effect on the quality of our food (Verpoorte *et al.*, 2000). These secondary metabolites can be classified into three groups; the terpenoids, the phenylpropanoids/allied phenolics, and the alkaloids. Terpenoids are derived from the five carbon precursor isopentenyl diphosphate. The phenylpropanoids/allied phenolics are formed in the shikimic acid and mevalonate/acetate pathways. Alkaloids contain one or more nitrogen atoms, and are biosynthesized from amino acids (Buchanan *et al.*, 2000).

1.4 Alkaloids

Alkaloids are physiologically active, nitrogen containing compounds with low molecular weights. These compounds are found in approximately 20% of all flowering plants and can also be found in other organisms such as insects, frogs, and fungi (Kutchan, 1995; Buchanan *et al.*, 2000). There is a large amount of information about the pharmacological effects of many alkaloids, but very little is known about the regulation of alkaloid synthesis (Kutchan, 1995; De Luca and St Pierre, 2000). Morphine is one of the most widely known alkaloids. It is interesting to note that the structure of morphine was not determined until 1952 due to its stereochemical complexity, containing five asymmetric centers (Facchini, 2001). The role of alkaloids in plants is not well understood, although they are thought to be used as a chemical defense against insects, herbivores, and pathogens (Kutchan, 1995; Buchanan *et al.*, 2000; Facchini, 2001).

P. somniferum contains two classes of alkaloids: the benzylisoquinolines and the morphinane, also known as the phenanthrene alkaloids. The benzylisoquinolines, such as papaverine, are used as muscle relaxants (Bryant, 1988). Morphine, codeine, and thebaine (a useful precursor for semisynthesis of more potent alkaloids) all belong to the morphinane group. Figure 1.3 shows the partial biosynthetic pathway for morphine alkaloids. Many synthetic derivatives have been synthesized from opium alkaloids such as: hydromorphone, hydrocodone, and oxycodone (Nessler, 1990).

Numerous studies have demonstrated that alkaloid biosynthesis is a highly organized process in relationship to plant development; alkaloids can be highly tissue specific and can be found in specific organs, cells or within the organelles of cells (De Luca and St Pierre, 2000; Verpoorte *et al.*, 2000). Alkaloid accumulation in specific tissues, such as the cytoplasm of opium poppy, may be exploited to aid in the engineering of increased concentration of compounds. This seems especially evident in the case of morphine and codeine which occur exclusively in latex.

Recent studies have been aimed at identifying the site of biosynthesis of the benzylisoquinoline alkaloids. Researchers have shown that two other cell types associated with the phloem, sieve elements and companion cells, were found to be the site of the biosynthesis of certain key enzymes in the biosynthesis of morphine (Bird *et al.*, 2003). Elucidating the cell metabolic pathways active in sieve elements, companion

cells, and laticifers that associate with the phloem, may aid in understanding the complex pathway of alkaloid biosynthesis. Enhancing production of general alkaloid precursors will promote increased yield for pharmaceutical uses.

1.5 Promoters

Promoters are a key element of gene expression and regulation in a cell. Without this gene control a cell would not express a desired protein at the correct time, location or level necessary for proper growth and development. Promoters are genetic switches that recruit protein factors to initiate gene transcription. Regulatory elements found on the same DNA strand as the gene coding region are known as *cis*-elements (Buchanan *et al.*, 2000). The promoter and its regulatory elements are found primarily in the 5' upstream region of the transcription start site. A potential method for enhancing alkaloid production involves using the natural latex-specific promoters found in opium poppy to drive expression of biosynthetic genes.

A common tool in plant molecular biology is the constitutive 35S cauliflower mosaic virus (CaMV 35S) promoter. The CaMV 35S promoter is claimed to be always active in all cell types (Baulcombe *et al.*, 1986). This promoter has advanced the understanding of what the gene function is, but it may not necessarily be an accurate representation of the gene expression *in planta*. The expression of genes can be tissue specific, for example the biosynthesis of vinblastine and vincristine in *Catharanthus roseus*, in which only certain metabolites are associated with specific cell types (Verpoorte *et al.*, 2000).

Determining tissue specificity of gene expression can provide clues for understanding a gene's function and can lead to identification of important promoters for specific bioengineering objectives. Two examples of tissue specific promoters are the pENOD12A, nodulation-specific promoter, and the γ- kafarin seed-specific promoter (Leite *et al.*, 2000; Wisniewski and Brewin, 2000). The γ-kafarin seed-specific promoter was used to express a human growth hormone in tobacco seed so that the protein would be stable for efficient extraction (Leite *et al.*, 2000). The concept of tissue-specific promoters can be exploited for increasing alkaloid biosynthesis in *P. somniferum*. The isolation of latex-specific promoters, LSPs, would be especially helpful in metabolically engineering the latex-specific enzymes required for morphinane alkaloid production. The

possibility of over-expressing or over-producing these important compounds and precursors would increase its overall yield per hectare.

1.6 Metabolic engineering of poppy alkaloid biosynthesis

Many attempts have been made to manufacture large-scale amounts of alkaloids in plant suspension cultures. Such efforts have not been successful for morphinane alkaloid production, and the pharmaceutical industry still relies mainly on the organic synthesis of morphine from raw opium and from precursors such as thebaine, noscapine and codeine (Bryant, 1988). There has been success using California poppy cell cultures to produce benzophenanthridine alkaloids (Park *et al.*, 2002). The increase in abundance of the essential compounds such as morphine and codeine *in planta* using transgenic approaches would increase the usefulness of organic synthesis from raw materials. In order to use molecular biology tools to incorporate desirable gene expression, one must first understand how to regenerate and transform the crop of interest.

1.7 Tissue culture regeneration

The concept of totipotency, the ability of one cell possessing the potential to reproduce an entire organism, was not confirmed until the 1960s (Hartmann *et al.*, 1997). The concept was proven in tobacco using specific nutritional and physiological conditions (Vasil and Hildebrandt, 1965). Regeneration of whole plants from tissue culture of opium poppy is not as easy as with other plant models such as tobacco or carrot. Nevertheless, the regeneration process for opium poppy was developed and can be used to regenerate whole plants from hypocotyl callus (Nessler, 1982).

1.7.1 Transformation techniques for opium poppy

Building on the knowledge of plant regeneration, techniques of transformation were developed to insert desired genes into opium poppy. The most successful transformation methods used successfully are biolistics and *Agrobacterium*-mediated (Nessler, 1998; Park and Facchini, 2000).

Somatic embryogenesis is the development of embryos from cells *in vitro* (Hartmann *et al.*, 1997). Somatic embryos in suspension cultures are tissues that will readily accept microparticle bombardment (Nessler, 1998). These cell clusters are ideal because they are continuously dividing to form new embryoids. The biolistic methodology was pioneered in 1987 by Sanford who used DNA coated tungsten particles

to penetrate tobacco cells (Klein *et al.*, 1988). The process allows DNA coated on a microprojectile to desorb from its surface and incorporate into the plant chromosome DNA (Klein *et al.*, 1988). The exact process of DNA incorporation into the genome is not well understood. Biolistic transformation has been optimized in various aspects for opium poppy somatic embryos (Nessler, 1998). These general parameters include: microparticle type and size, distance of DNA to the target, and amount of DNA, among other factors. The particle bombardment method yields transgenic plants in approximately nine to twelve months.

Agrobacterium-mediated transformation methods for opium poppy have been developed by multiple groups using either *A. tumefaciens* or *A. rhizogenes* (Nessler, 1998; Park and Facchini, 2000; Yoshimatsu and Shimomura, 2001). Transformations have been performed by standard regeneration protocols using hypocotyls or root cultures. *Agrobacterium*-mediated methods provide a way to insert genes at a low copy number, but regeneration time can take up to two years.

1.8 Summary

The metabolic engineering of opiate pathways is now possible with the knowledge of the biosynthetic pathways and cloning of many enzymes in poppy alkaloid biosynthesis. Reaching the goal of increasing precursor production requires the consideration of various aspects of pathway regulation. One important parameter is flux through the entire pathway. This can be assessed through measurement of intermediate or product degradation, competitive pathways, rate-limiting enzymes in the reaction, over-expression of certain enzymes and precursors, and silencing of genes encoding regulatory enzymes (Verpoorte *et al.*, 2000). These parameters are important when designing a metabolic engineering project. The morphinane and benzylisoquinoline pathways are fairly well understood in contrast to many other alkaloid pathways and represent an ideal target for manipulation. Biochemical pathways in eukaryotic organisms are becoming better understood, and with this knowledge the concept of a linear pathway is becoming increasingly challenged. A complex network of pathways is the most likely model for poppy alkaloid production. LSPs will promote the elucidation of complex interactions between/among players in the morphinane alkaloid pathway,

while providing tools for tissue-specific expression of genes that encode pathway enzymes.

Figure 1.1: Latex ‘bleeding’ from a capsule that has been wounded by scoring.

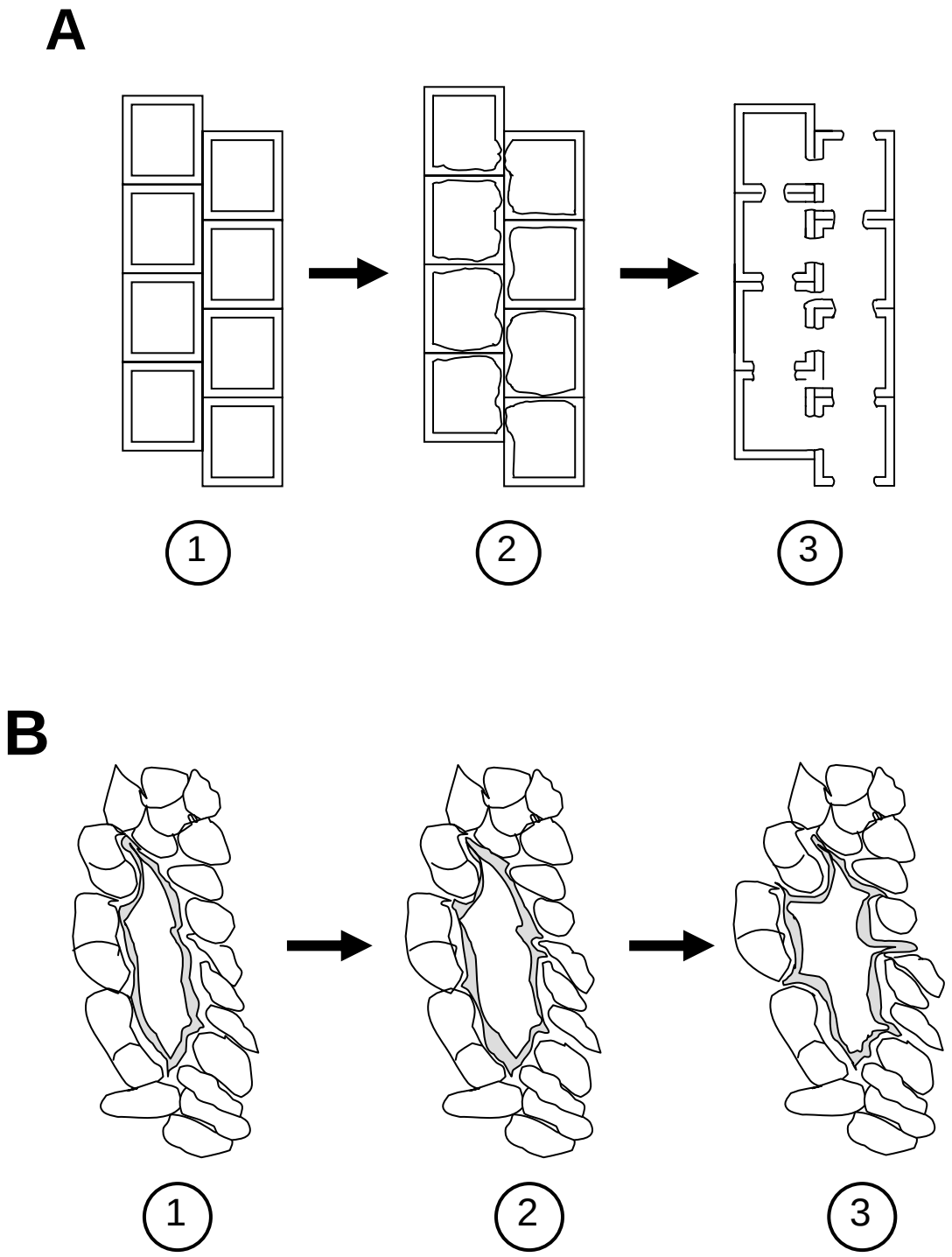

Figure 1.2: Diagram of two laticifer cell types
A: Example of an articulated laticifer and development of latex vessels. A1, Laticifer elements. A2, Walls between cells begin thinning and pores begin to form. A3, Cells connect through degraded end walls forming a latex vessel. B: Example of a non-articulated laticifer and branching of the initial. B1, Non-articulated initial. B2, Intrusive nature of initial between cells. B3, Intrusive branching of laticifer initial between cells.

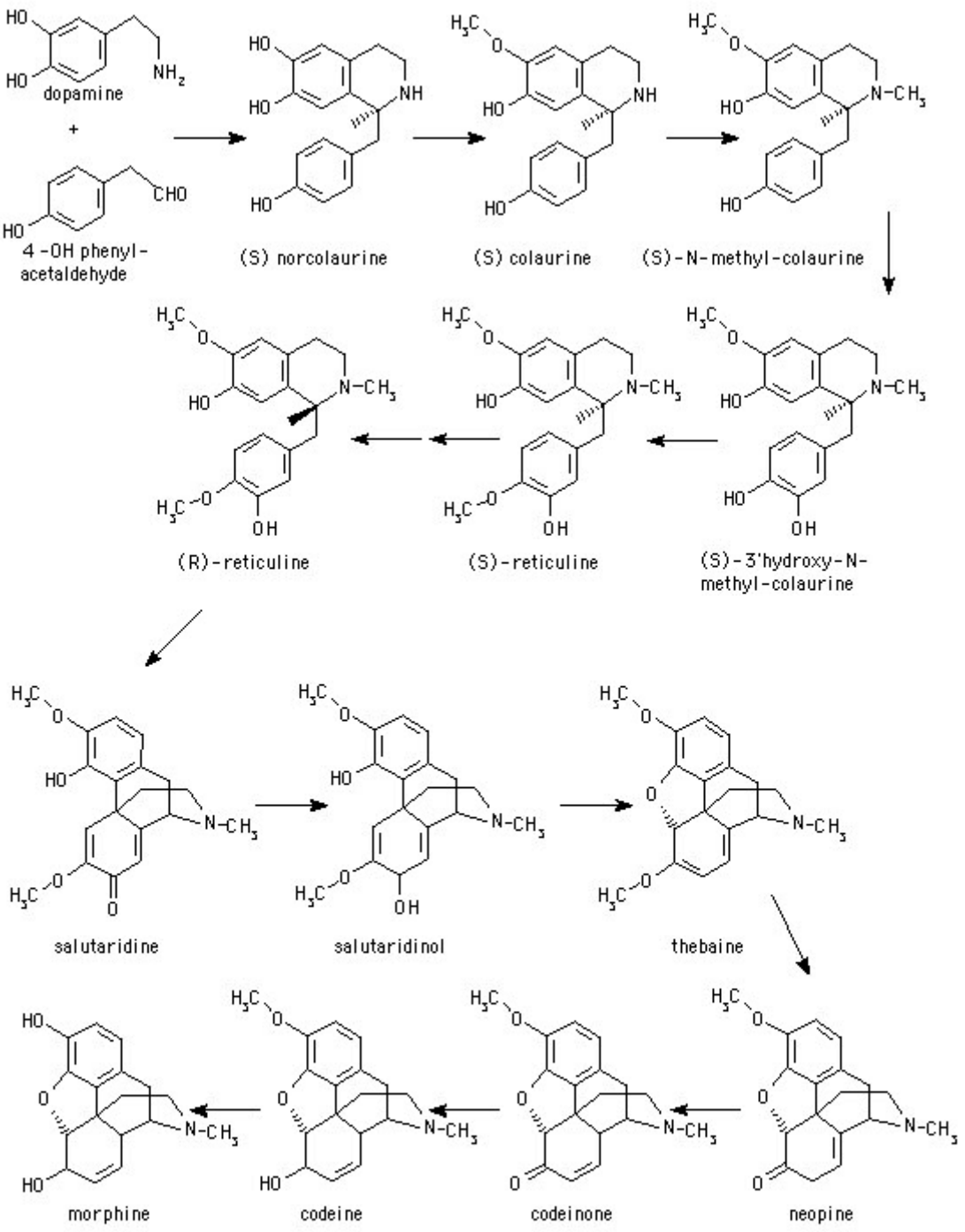

Figure 1.3 Morphinane alkaloid biosynthetic reactions

1.9 REFERENCES

Baulcombe DC, Saunders GR, Bevan MW, Mayo MA, Harrison B (1986) Expression of biologically active viral satellite RNA from the nuclear genome of transformed plants. Nature **321:** 446-449

Bird DA, Franceschi VR, Facchini PJ (2003) A tale of three cell types: alkaloid biosynthesis is localized to sieve elements in opium poppy. Plant Cell **15:** 2656-2635

Booth M (1998) Opium A History, Ed 1. St. Martin's Press, New York

Bryant RJ (1988) The manufacture of medicinal alkaloids from the opium poppy-a review of a traditional biotechnology. Chem Ind **5:** 146-153

Buchanan B, Gruissem W, Jones RL (2000) Biochemistry & Molecular Biology of Plants. American Society of Plant Physiologists, Rockville

De Luca V, St Pierre B (2000) The cell and developmental biology of alkaloid biosynthesis. Trends in Plant Sci **5:** 168-173

Esau K (1965) Plant Anatomy, Ed 2. John Wiley & Sons, New York

Facchini PJ (2001) Alkaloid biosynthesis in plants: biochemistry, cell biology, molecular regulation, and metabolic engineering applications. Annu Rev Plant Physiol Plant Mol Biol **52:** 29-66

Fahn A (1979) Secretory tissues in plants, Ed 1. Academic Press, New York

Fahn A (1997) Plant Anatomy, Ed 4. Aditya Books (P) Ltd., New Delhi

Hartmann HT, Kester DE, Davies FT, Geneve RL (1997) Plant Propagation: Principles and Practices, Ed 6. Simon and Schuster, Upper Saddle River

Klein TM, Harper EC, Svab Z, Sanford JC, Fromm ME, Maliga P (1988) Stable genetic transformation of intact *Nicotiana* cells by the particle bombardment process. Proc Natl Acad Sci USA **85:** 8502-8505

Kutchan T (1995) Alkaloid biosynthesis- the basis for metabolic engineering of medicinal plants. Plant Cell **7:** 1059-1070

Kutchan T (1998) The alkaloids: chemistry and pharmacology. *In* G Cordell, ed, The Alkaloids Chemistry and Biology, Vol 50. Academic Press, New York, pp 257-316

Leite A, Kemper EL, de Silva M, Luchessi A, Siloto RMP, Bonaccorsi ED, El-Dorry HF, Arruda P (2000) Expression of correctly processed human growth hormone in seeds of transgenic tobacco plants. Mol Breed **6:** 47-53

Nessler CL (1982) Somatic embryogenesis in the opium poppy, *Papaver somniferum*. Physiol Plant **55:** 453-458

Nessler CL (1990) Poppies. *In* P Ammirato, DR Evans, WR Sharp, YPS Bajaj, eds, Handbook of Plant Tissue Culture Ornamental Species, Vol 5. Mc-Graw Hill Publishing Company, New York, pp 693-715

Nessler CL (1998) In Vitro Culture Technologies. *In* J Bernáth, ed, Poppy the genus *Papaver*, Vol 3. Harwood Academic Publishers, Luxembourg, pp 209-217

Nessler CL, Allen RD, Galewsky S (1985) Identification and characterization of latex-specific proteins in opium poppy. Plant Physiol **79:** 499-504

Nessler CL, Mahlberg PG (1976) Laticifers in stamens of *Papaver somniferum* L. Planta **129:** 83-85

Park S-U, Facchini PJ (2000) *Agrobacterium rhizogenes*-mediated transformation of opium poppy, *Papaver somniferum* L., and California poppy, *Eschscholzia californica* Cham., root cultures. J Exp Bot **51:** 1005-1016

Park S-U, Yu M, Facchini PJ (2002) Antisense RNA-mediated suppression of benzophenanthridine alkaloid biosynthesis in transgenic cell cultures of California poppy. Plant Physiol **128:** 696-706

Vasil V, Hildebrandt AC (1965) Differentiation of tobacco plants from single, isolated calles in microcultures. Science **150:** 889-892

Verpoorte R, van der Heijden R, Memelink J (2000) Engineering the plant cell factory for secondary metabolite production. Transgenic Res **9:** 323-343

Wisniewski JP, Brewin NJ (2000) Construction of transgenic pea lines with modified expression of diamine oxidase and modified nodulation responses with exogenous putrescine. Mol Plant-Microbe Interact **13:** 922-928

Yoshimatsu K, Shimomura K (2001) Genetic transformation of *Papaver somniferum* (Opium Poppy) for Production of Isoquinoline Alkaloids. *In* YPS Bajaj, ed, Transgenic Medicinal Plants. Springer, New York, pp 178-192

Chapter 2

LATEX-SPECIFIC PROMOTER CONSTRUCTS IN *PAPAVER SOMNIFERUM, ARABIDOPSIS THALIANA* AND *NICOTIANA TABACUM*

2.1 INTRODUCTION

2.1.1 Dioxygenases

2-Oxoglutarate and Fe (II)-dependent dioxygenases are widespread in eukaryotes and bacteria, and typically catalyze reactions involving oxidation of an organic substrate using a dioxygen molecule (Aravind and Koonin, 2001). Dioxygenases are non-heme iron containing enzymes, and are involved in the biosynthesis of plant signaling compounds such as abscisic acid, gibberellins, and ethylene, as well as secondary metabolites. Some of these important secondary metabolites include flavonoids and alkaloids (Prescott and John, 1996). There are two classes of plant dioxygenases: lipoxygenases and 2-oxoglutarate dependent oxygenases. 2-Oxoglutarate dioxygenases perform hydroxylation, desaturation and epoxidation reactions, and most of them require a 2-oxoacid as a substrate. This class of enzymes also requires Fe^{2+} and ascorbate for optimal activity.

One of the first studied 2-oxoglutarate-dioxygenases was prolyl-4-hydroxylase (P4H EC 1.14.11.2), which is involved in post-translational processing of collagen (De Carolis and Deluca, 1994). Prolyl-4-hydroxylase has a requirement for Fe^{2+} ions as well as ascorbate for optimal activity. The P4H works similarly in plants by introducing a hydroxyl group into prolyl residues that are incorporated into cell wall proteins of higher plants and algae (Prescott and John, 1996). 2-Oxoglutarate-dioxygenases have been shown to be involved in synthesis of plant growth regulators, notably gibberellins (GA) and ethylene. One of the best characterized gibberellin dioxygenases is gibberellin-20 oxidase, which catalyzes the latter steps of GA biosynthesis (Prescott and John, 1996). Recently a plant growth retardant has been developed, prohexadione-calcium, that acts as a structural mimic of 2-oxoglutarate, thereby inhibiting dioxygenase activities that catalyze distinct steps in the biosynthesis of the growth hormone gibberellin (Roemmelt *et al.*, 2003). When apple (*Malus x domestica*) was treated with prohexadione calcium, novel flavonoids were formed and induced resistance to fire blight (*Erwinia amylovora*) and other plant diseases was observed (Roemmelt *et al.*, 2003).

The structure of a microbial 2-oxoglutarate dioxygenase, isopenicillin N synthase (IPNS) has been published, and shares a number of conserved residues/motifs and structural elements with plant 2-oxoglutarate dioxygenases. IPNS has been proposed as a model for plant 2-oxoglutarate dioxygenases (Prescott and John, 1996). The general structure contains the jellyroll motif, a comparatively rare structure in enzymes, but commonly found in viral capsid proteins (Prescott and John, 1996). Another structure that has been putatively identified from IPNS is a leucine zipper motif that has also been found in a number of plant dioxygenases (Prescott and John, 1996).

A 2-oxoglutarate-dependent oxygenase, hyoscyamine 6-β-hydroxylase (H6H), has been used to increase a pharmaceutically important alkaloid, scopolamine, in the plant *Atropa belladonna* (Yun *et al.*, 1992). This provided the first example of how medicinal plants can be successfully altered using transgenic techniques to produce increased quantities of a medicinally important alkaloid (Kutchan, 1998). Another 2- oxoglutarate-dependent oxygenase has been cloned and characterized, the enzyme deacteovindoline 4 hydroxylase (D4H), which is involved in the biosynthesis of the alkaloid vindoline in *Catharanthus roseus* (Vazquez-Flota *et al.*, 1997). Vindoline monomers and catharanthine found in leaves of *C. roseus* make up the dimeric alkaloids vinblastine and vincristine used in treatment of Hodgkin's disease and certain types of leukemia (Vazquez-Flota *et al.*, 1997). The enzyme D4H shares a conserved amino acid motif with the *e8* protein, D6H, and ethylene forming enzyme (EFE) (Vazquez-Flota *et al.*, 1997).

A latex cDNA clone, 2L1, was found to have homology to a 2-oxoglutarate-dioxygenase by previous research in the Nessler laboratory using Expressed Sequence Tags (EST) analysis (Pilatzke, 1999). As mentioned previously, 2-oxoglutarate-dioxygenases have been found to function in ethylene formation (Prescott and John, 1996), but no EFE activity was determined for 2L1 (Pilatzke, 1999). RNA gel blot analysis has shown the latex-specificity of the opium poppy 2-oxoglutarate-dioxygenase, Figure 2.1. The latex-specific 2-oxoglutarate dioxygenase (DIOX) from opium poppy has a high homology with the, senescence related growth, SRG1 protein (59% identity and 78% similarity) involved in organ senescence (Pilatzke, 1999). The *EFE*, *H6H* and

SRG genes have three conserved intron positions and indicate a high degree of homology among 2-oxoglutarate-dioxygenases (Prescott and John, 1996).

Overall, 2-oxoglutarate dioxygenases play an important role in plant metabolism, and overlapping functions are shared among the 2-oxoglutarate dioxygenases, lipoxygenases and the P450 monooxygenases (Prescott and John, 1996). The overlapping and extensive functional complementarities are intriguing and have been demonstrated specifically in alkaloid biosynthesis (Yun *et al.*, 1992; Rosco *et al.*, 1997).

2.1.2 Plant model systems

The model systems used in this research were opium poppy, for its laticifer cell-type, *Arabidopsis thaliana* for its ease of transformation and tobacco, another alkaloid producing plant that does not contain laticifers. Knowledge of the regulation and gene expression in various plant systems may help determine the role and functions of genes of interest. Knowledge of enzymes and how they work in various plant species is critical to understanding processes and biosynthetic pathways.

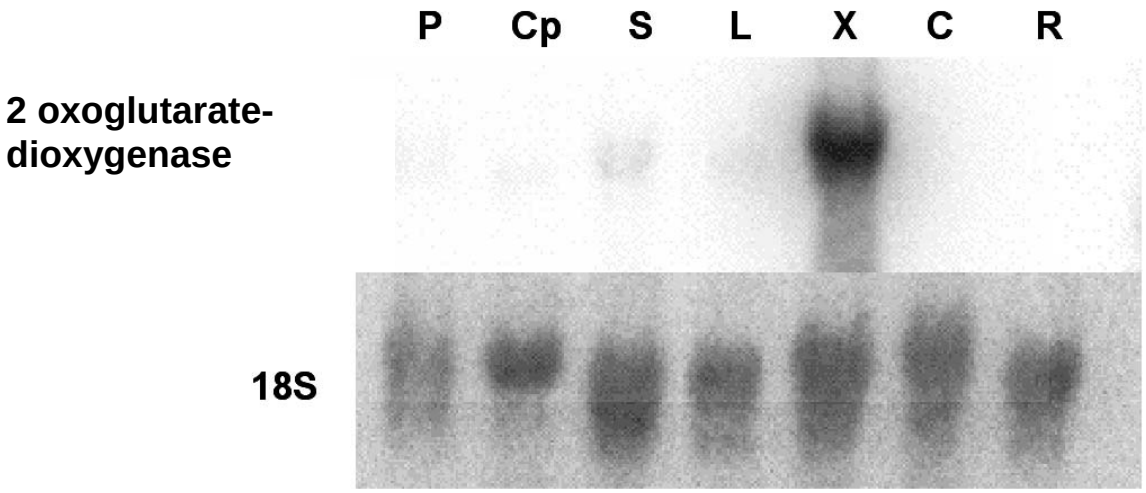

Figure 2.1 RNA gel blot analysis of 2-oxoglutarate-dioxygenase homolog in various poppy tissues. 10 µg of total RNA from poppy tissues were loaded in each lane with 18S ribosomal band below to demonstrate equal loading. Samples were hybridized with the homolog from a latex EST library. P, petal; Cp, capsule; S, stem; L, leaf; X, latex; C, callus; R, root (Pilatzke, 1999).

2.2 MATERIALS AND METHODS

2.2.1 Isolation of LSP: 2-oxoglutarate-dioxygenase promoter (*DIOX)*

Previous work in the Nessler laboratory led to the isolation and the characterization of a latex-specific cDNA in poppy, 2L1, with sequence similarity to 2-oxoglutarate-dioxygenases (DIOX) (Pilatzke, 1999). The *DIOX* cDNA clone was used to screen a *P. somniferum* λ EMBL3 genomic DNA library, in order to clone the 5' upstream region and for identification of promoter elements. A 3.0 kb portion of the 5' upstream region (UR) was isolated and subcloned it into pBluescript KS (Stratagene, La Jolla, CA) using *Xba I* and *Sal I* restriction sites (Nessler, unpublished results).

2.2.1.1 Sequencing

The UR *DIOX* was sequenced to identify restriction sites within this region. Initial sequences were obtained using the T3 and T7 primers. Multiple primers were then designed to sequence the entire fragment by primer walking; Appendix A. Sequencing was performed using BigDye Terminator v.3.1 cycle sequencing kit (Applied Biosystems, Foster City, CA) on an ABI Prism 3100 at the Virginia Bioinformatics Institute core laboratory facility (VBI CLF) at Virginia Tech. The 5' region of interest was determined to be 2,534 bp long and was analyzed for restriction endonuclease recognition sites (Figure 2.2) using Sequencher (Gene Codes, Ann Arbor, MI) and Mac Vector 7.2.2 (Accelrys, San Diego, CA).

2.2.1.2 PCR Cloning

The 5' UR was amplified via polymerase chain reaction (PCR) using the T7 reverse primer 5' GTAATACGACTCACTATAGGG 3' in order to preserve cloning restriction sites and a 2L1ProNco forward primer 5'TCCATGGACTCTATATTTGATGAAC 3' to make a functional translational fusion. PfuTurbo® (Stratagene, LaJolla, CA) high fidelity polymerase 0.1 µl was used to amplify the sequence of interest along with PCR beads puRE Taq Ready to Go Beads™ (Amersham Life Sciences/GE Healthcare, Waukesha, WI). PCR conditions included an initial denaturation step of 1 minute at 95°C, 30 cycles of 95°C for 1 minute, 50°C for 1 minute, and 72°C for 3 minutes, and a final extension of 72°C for 10 minutes. The T7 and 2L1ProNco 3.0 kb PCR product was isolated from a 1% agarose gel using Qiaex II kit (Qiagen, Valencia, CA), ligated into the pGEM-T easy vector (Promega, Madison,

WI), and transformed into *E. coli* DH5α chemically competent cells (Invitrogen, Carlsbad, CA).

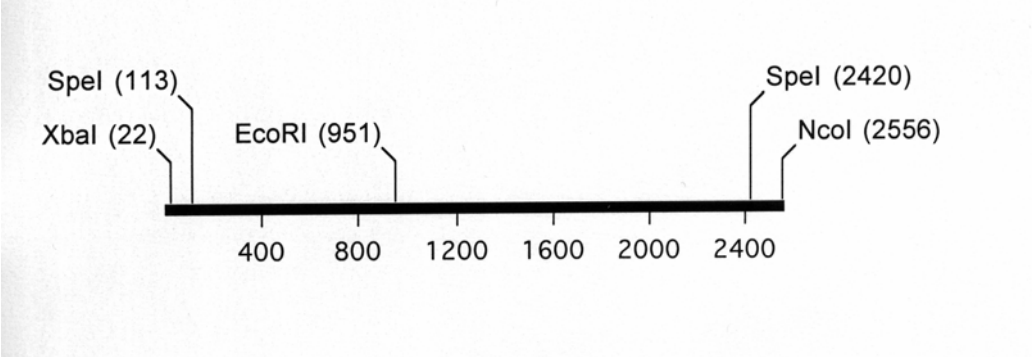

Figure 2.2. Restriction map of the 5'UR of *DIOX*.

Colonies were selected and screened on Luria Broth (LB) medium containing 100 mg/L ampicillin (Sigma, St Louis, MO), 100 mM IPTG (isopropyl-beta-D-thioglactoside Gold BioTechnology St. Louis, MO), and 50 mg/L of X-gal (5-bromo-4-chloro-3-indoxyl-beta D-glactopyranoside, Gold BioTechnology St. Louis, MO). White colonies were grown in 3 mL of liquid LB and 100 mg/L of ampicillin. Plasmid DNA from a bacterial culture was isolated using the QIAprep miniprep kit (Qiagen, Valencia, CA). The *DIOX*:pGEM-T easy plasmids were digested with restriction enzymes *Xba I* and *Nco I* to confirm the presence of the insert. The DIOX insert was also sequenced to insure that no mutations had been introduced during the PCR amplification.

2.2.1.3 Subcloning of 5'upstream region *DIOX*

An *Xba I*/*Nco I* fragment was excised from the *DIOX*:pGEM-T easy plasmid and ligated into the modified vector *Nco*:GUS pUC18 (Nessler, unpublished), containing the β-glucuronidase gene, GUS (Jefferson *et al.*, 1987) and the nopaline synthase terminator, NosA+, from binary vector pBI-121 (Accession number AF 485783). The resulting vector was a DIOX promoter translational reporter gene construct or DIOX:*Nco*GUS pUC18.

DIOX:*Nco*:GUS pUC18 was then subcloned into the binary vector pCAMBIA 1300 (Roberts *et al.*, 1998) (Accession number AF234296). This produced the

DIOX:GUS pCAMBIA 1300 (DIOX:GUS pCAM) plasmid (Figure 2.3). To assess the functional sites within the promoter, deletions were made at the *EcoRI*951 bp and the *SpeI* 2420 bp restriction sites, Figure 2.2. The resulting promoter deletion constructs were then designated *EcoRI*:*DIOX*:GUS pCAMBIA 1300 and *SpeI*:*DIOX*:GUS pCAMBIA1300. The *EcoRI:*GUS and NosA+ fragment was also subcloned into pBIN19, Figure 2.4. The two different binary vectors were used to determine if there was read through in the pCAMBIA1300 vector. Read through of the plasmid could cause expression of the GUS gene that would not be due to the *DIOX* promoter, but to the constitutive promoter driving antibiotic selection in pCAMBIA1300.

2.2.2 Poppy transformation methods

2.2.2.1 Generation of somatic embryos for biolistics

In order to develop somatic embryos usable for particle bombardment, 1-2 week old seedling hypocotyls were placed on sterile H medium (Nessler, 1990), Table 2.1. After approximately 2 weeks, callus formed on the cut hypocotyls. Meristemoids developed on the callus after 2 to 3 months, with monthly subculture of callus onto fresh H medium. Somatic embryos were then transferred into a liquid 0.25D medium (Table 2.1), where they were allowed to continuously divide until subjected to biolistic transformation.

2.2.2.2 Biolistic transformation of poppy somatic embryos

The plasmid constructs were linearized by restriction digestion with the enzyme *XbaI* and inserted into embryo cultures via particle bombardment. When used for particle bombardment, linearized DNA is more stable, produces lower copy number, and has fewer rearrangements in contrast to supercoiled plasmids (Fu *et al.*, 2000). Opium poppy somatic embryos will readily accept particle bombardment because they are continuously dividing (Nessler, 1998). Bombardment was performed using a PDS-100 Helium Gene Gun (Bio-Rad Laboratories, Hercules, CA). Embryos were shot with 1.0 μM gold particles coated with the desired plasmid construct and 20 μL of 0.1 M spermidine (Sigma, St Louis, MO), at a DNA concentration of 5 μg/ml, distance of 16 cm, and a pressure of 1550 psi (Nessler, 1998).

Bombarded somatic embryos were grown on 0.25D solid medium containing the antibiotic hygromycin (25 mg/mL) in order to select for transformants (Nessler, 1998).

After several rounds of antibiotic selection, embryos were subjected to the regeneration protocol for somatic embryos, Figure 2.5. The embryos differentiated into shoots or entire plants on ½ MS solid medium lacking hormones with a photoperiod of 12 hours at 22°C in a Percival incubator (Percival Scientific, Perry, IA) (Nessler, 1990). Plantlets exhibiting a well developed leaf and root system were transplanted into Magenta boxes containing poppy soil mix. Poppy soil mix contains 4 parts of BACCTO® soil (Michigan Peat Company, Houston, TX), 2 parts of coarse vermiculite and 1 part perlite.

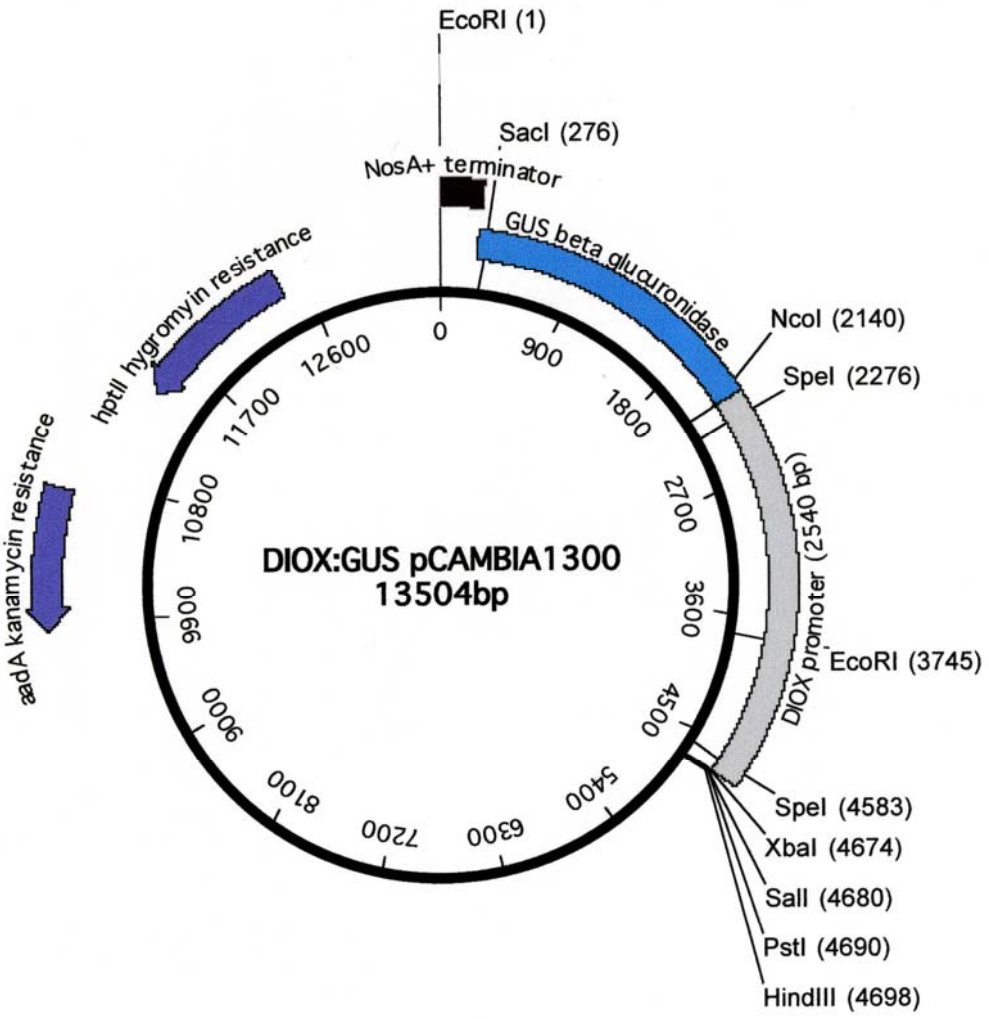

Figure 2.3. Plasmid map of the DIOX:GUS fusion in binary vector pCAMBIA-1300

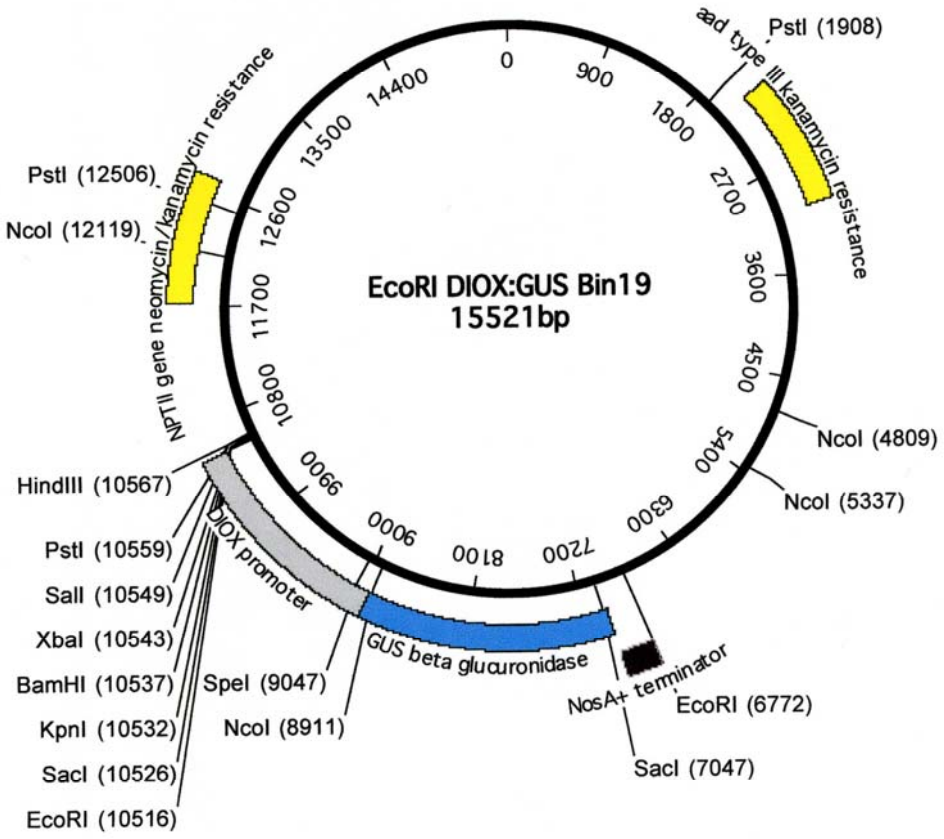

Figure 2.4 Plasmid map of *EcoRI* DIOX:GUS plasmid in binary vector pBIN19.

Ingredient	H solid	.25D liquid	.25D solid	1/2 MS liquid	1/2 MS solid	MS solid	(TR)[1]
MS salts and vitamins2 (Murashige and Skoog 1962)	4.4 g	4.4 g	4.4 g	2.2g	2.2g	4.4	4.4 g
Sucrose[3]	30 g	30 g	30 g	15 g	15 g	30 g	30 g
Magnesium sulfate $MgSO_4$ $7H_20$	0.4g		0.4 g		0.4 g	0.4 g	0.4 g
2,4 D (2,4-Dichlorophenoxyacetic acid[4] stock solution 10mg/100ml	2.2 ml	2.5 ml					
Kinetin (6-Fufurylaminopurine)[4] stock solution 10mg/100ml	11ml						
NAA (1-Napthalenacetic aicd) [4] stock solution 10mg/100ml							0.2 ml
BAP (6-Benzyl-aminopurine)[4] stock solution 10mg/100ml							4 ml
Phytagel [4]	3.5 g		3.5 g		3.5 g	3.5 g	3.5 g
pH	5.8	5.8	5.8	5.8	5.8	5.8	5.8
H_2O	add to 1 L	add to 1 L	add to 1 L	add to 1 L	add to 1 L	add to 1 L	add to 1 L

1 TR= Tobacco Regeneration

2 Cassion Laboratories Sugar City, ID

3. Fisher Scientific, Pittsburgh, PA

4 Sigma St. Louis, MO

Table 2.1 Plant media used for growth and regeneration of various plant species

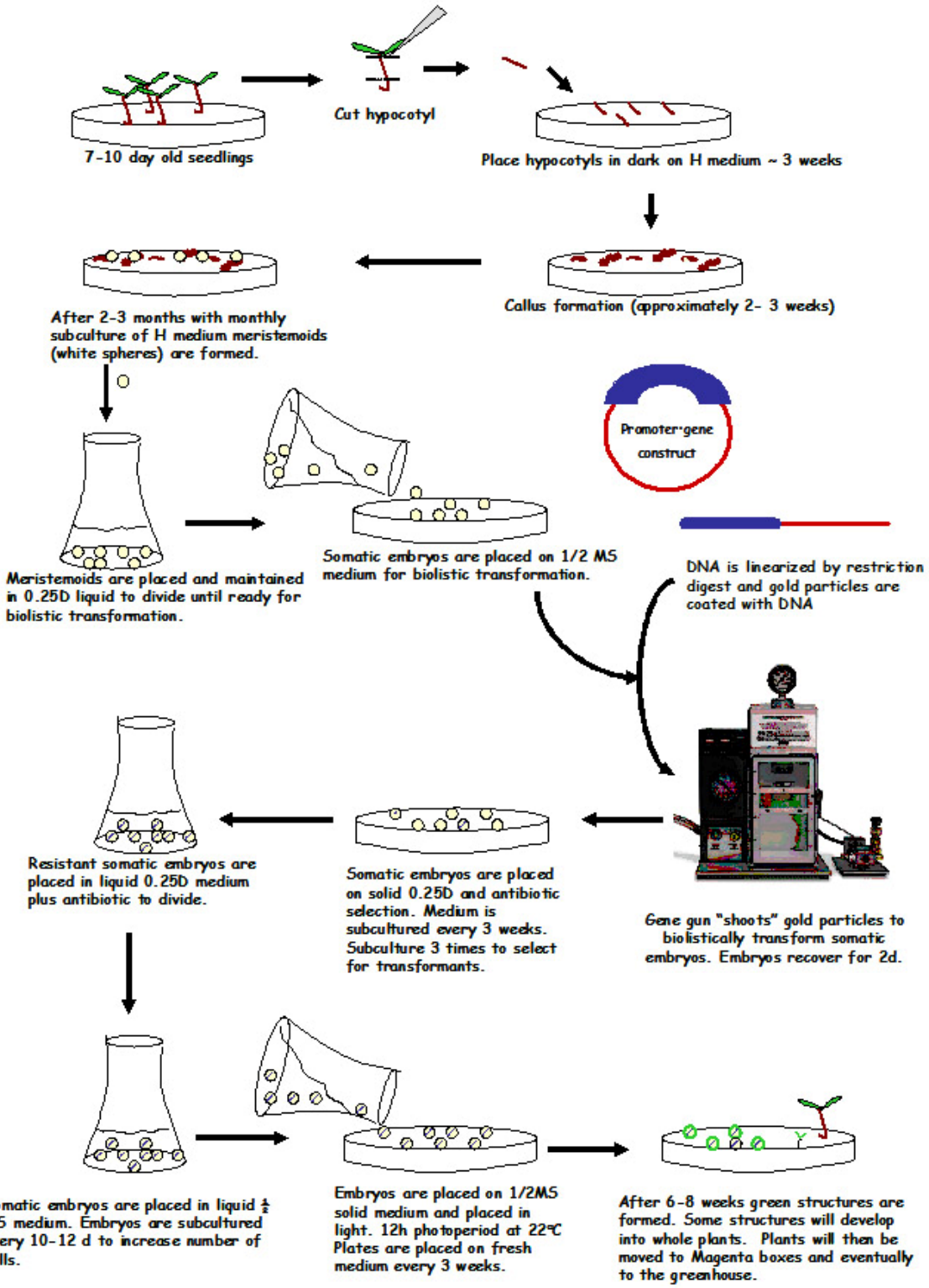

Figure 2.5 Generation of somatic embryos for biolistic transformation and poppy regeneration procedure.

2.2.3 *Agrobacterium* mediated transformation of *A. thaliana* ecotype Columbia

Agrobacterium tumefaciens strain GV3101 was transformed with the DIOX:GUS pCAM, EcoRI:DIOX:GUS pCAM1300, SpeI:DIOX:GUS pCAM1300 or EcoRI:DIOX:pBIN19 plasmids by the heat shock method of transformation (Gelvin and Schliperoort, 1988). To prepare competent cells, a 5 mL culture of the GV3101 strain was grown in YEP broth containing 50 mg/L of gentamycin (Sigma, St. Louis, MO) and 100 mg/L of rifampicin (Sigma, St. Louis, MO) for selection of the strain. YEP broth contained 2% (w/v) Bacto peptone, 1% (w/v) yeast extract and 85.5 mM NaCl, pH 7.0. Cells were grown for 24 h at 28° C and shaking at 250 rpm. Cultures were placed on ice and then pelleted at 3,000 x g for 5 min. Pellets were resuspended in 100 μL ice-cold 20 mM calcium chloride. Approximately 500 ng of plasmid DNA was added to the solution and then frozen in liquid nitrogen. Cells were thawed at 37°C for 5 min and then 400 μl of YEP medium was added for recovery. Cells recovered for 4.5 h at 28°C in a shaking incubator and plated on YEP medium containing 50 mg/L of gentamycin, 100 mg/L of rifampicin and 50 mg/L kanamycin (Sigma, St. Louis, MO) to select for transformed cells.

A. thaliana plants were grown in Fafard #2 soil mix (Griffin Greenhouse Supply Richmond, VA) under 24 h light and 25°C in a Percival growth chamber (Percival Scientific, Perry, IA). Plants approximately 4 weeks old, were cut back to induce the formation of more floral buds for transformation two days later. The transformation of *Arabidopsis* plants was performed using the floral dip method (Clough and Bent, 1998). The transformed cells containing the promoter translational fusion plasmid were grown in 500 mL YEP for 48 h. Cells were centrifuged at 3,000 x g for 15 min and resuspended in approximately 500 mL of 5% sucrose (Fisher Scientific, Pittsburgh, PA) and 0.02% (v/v) Silwet L-77 (Lehle Seed, Round Rock, TX). Plants were dipped in the *Agrobacterium* suspension for 15 min at room temperature. The dipped plants were placed on their sides and covered for 24 h, and then placed upright and grown under 24 h light and 25°C until maturity.

Seeds were collected from primary *Arabidopsis* transformants and were plated on MS agar medium containing 25 mg/L of hygromycin (AgriBio, N. Miami, Fl) and 500 mg/L carbenicillin (AgriBio N, Miami, FL). Hygromycin is the selectable marker in

plants and carbenicillin prevents the growth of residual *Agrobacterium*. Plants that grew under selection were transferred to soil after 3 to 4 weeks and grown for seed. Plants that tested positive for the GUS gene by histochemical staining continued to be grown for seed. The next generations were planted and placed on selective medium until homozygous resistance was observed. Independent homozygous lines were identified so that histochemical staining and localization could be observed.

2.2.4 *Agrobacterium* mediated transformation of *Nicotiana tabacum* cv. Xanthi

Agrobacterium tumefaciens strain LBA4404 was transformed using the heat shock method previously described with the exception that 50 mg/L of streptomycin (Sigma, St. Louis, MO) was used instead of gentamycin for proper growth of the strain. The promoter constructs transformed into LBA4404 were EcoRI:DIOX:GUS pCAM1300 or EcoRIDIOX:GUS pBIN19, Figure 2.4. *Nicotiana tabacum* cv. Xanthi was transformed using two different methods: the petiole scoring method (Medina-Bolivar *et al.*, 2003) and the leaf disc method (Horsch *et al.*, 1986). For the former small, fully developed Xanthi leaves were excised from the plant with approximately 0.5 cm of the petiole remaining on the leaf. A clean scalpel tip was dipped in a 2 d old *A. tumefaciens* colony and was then used to score the petiole parallel to the veins.

In the leaf disc method Xanthi leaves are cut into strips, incubated for 30 min in 2 d old culture, and resuspended in 5 mL MS liquid medium. Leaf strips were blotted on sterile Whatman paper. The scored leaves and cut leaves were placed on MS media for 2-3 d. The leaves were placed on tobacco regeneration media, Table 2.1, MS media containing 1-napthalene acetic acid (NAA), 6-benzlaminopurine (BAP), 200 mg/L kanamycin for plant selection and 500 mg/L of carbenicillin to kill the *Agrobacterium*. After 2-3 weeks, shoots and calli formed at the wounding sites. Shoots were placed on MS medium that contained 200 mg/L kanamycin and 500 mg/L carbenicillin to maintain selection. Rooted plants were genotyped via PCR and stained for GUS activity.

2.2.5 Plant histochemical staining and fixation

2.2.5.1 Histochemical staining

Various plant tissues such as leaves, veins and capsules were stained with the chromogenic substrate 5-bromo-4-chloro-3-indolyl glucuronide acid cyclohexyl ammomium salt (X-Gluc) (Gold BioTechnology, St. Louis, MO) 1 mg/mL in 50mM

sodium phosphate buffer pH 7.0 for up to 48 h at 37°C (Jefferson *et al.*, 1987). The reaction was stopped by the addition of 70% ethanol. Tissue was first observed under a Leica G27 stereomicroscope, Zeiss Stemi SV APO stereoscope, or the compound Zeiss Axioskop 2 (Thornbrook, NY), and photographed using the SPOT Advanced program (Diagnostic Instruments, Sterling Heights, MI).

2.2.5.2 Fixation of plant tissue

Selected tissues were post-fixed in 3% (v/v) glutaraldehyde Grade I (Sigma, St. Louis, MO) solution in 0.05M PIPES [Piperazine-1,4-bis(2-ethanesulfonic acid), Sigma, St. Louis, MO)] buffer pH 7.4 at room temperature for 2 h. The samples were washed in two exchanges of PIPES buffer, 15 min each and dehydrated in an ethanol series at 4°C. The ethanol series was 25%, 50%, 75%, 90%, and 100% with 30 min washes, at each concentration. A final 30 minute wash of 100% ethanol was used. After a final exchange of 4°C absolute ethanol, tissues were embedded in L.R. White resin (SPI supplies, Westchester, PA). The resin concentration was increased by 25% increments each day, until samples reached 100%. At 100% L.R. White resin, samples were left for 2 d at -20°C and then were allowed to slowly warm to room temperature by placing on a rocking shaker for 24 h. Plant tissues were transferred to BEEM capsules (SPI scientific, Westchester, PA), which were sealed and placed at 62°C under vacuum to polymerize the LR white resin.

Sectioning was done on a Lecia microtome. Samples were sectioned and heat fixed onto a glass slide. Sections were stained with a 0.5% safranin-O aqueous solution and glass coverslips were fixed with Permount (Fisher Scientific, Pittsburgh, PA).

2.2.6 Genomic DNA Extraction

DNA was extracted from all available putative transformants. One g of poppy tissue was used for extraction, and 0.1 g of tobacco and *Arabidopsis* tissues were used. Genomic DNA was extracted using the Nucleon Phytopure kit (Amersham Life Sciences, now GE Healthcare, Waukesha, WI.)

2.2.7 PCR amplification of DNA from putative transformants

PCR beads puRE Taq Ready to Go Beads™ (Amersham Life Sciences/GE Healthcare Waukesha, WI) were used to amplify DNA from putative transformants. One hundred and fifty ng of genomic DNA was used, as well as 1 μL of 10 μM of forward

primer, 1 µl of 10 µM of reverse primer and sterile molecular biology grade water to a final volume of 25 µL. Primers used to identify DNA containing the DIOX:GUS promoter fusion in poppy were 2L1-GUS forward 5' CCCGGCAGGTTAGATCGTAGA 3' and the 2L1-GUS reverse 5' TCTGCCAGTTCAGTTCGTTC 3', and the expected product size is 651 bp. The primers used to identify the GUS gene in tobacco were the GUS forward 5' GCAAAGTGTGGGTCAATAAT 3' and the GUS reverse 5' TACAGTTCTTTCGGCTTGTT 3', with an expected product size of 841 bp. Genomic poppy DNA was sonicated for 3 min before use in PCR experiments. PCR conditions for poppy consisted of an initial denaturation step of 4 min at 94°C. Cycling parameters were: 35 cycles of: 1 min denaturation at 94°C, 1 min of annealing temperature at 58°C, 1 min 72°C extension, and a final extension of 72°C for 10 min. PCR conditions for tobacco consisted of an initial denaturation step of 1 min at 94°C. Cycling parameters consisted of 30 cycles of: 1 min denaturation 94°C, 1 min annealing temperature at 57°C, 1 min extension of 72°C and a final extension of 72° for 10 min.

2.3 RESULTS

2.3.1 Plant transformation

Not all transformations attempted were successful. Table 2.2 lists the results and attempts at transformation in the plant species studied. In opium poppy 13 plants were regenerated following biolistic transformation procedures. Six embryo cell lines of each variation of DIOX:pCAMBIA were maintained, but on average only one embryo line would successfully regenerate whole plants, others generated green meristemoids. The most complete set of promoter:reporter gene fusions and transformants were attained for poppy. Figure 2.6 shows examples of successfully regenerated poppy plants at the greenhouse. In *Arabidopsis* a large number of EcoRI:DIOX:GUS pCAM lines were generated and one line of the SpeI:DIOX:GUS pCAM was generated. Tobacco transformations generated only the EcoRI:DIOX:GUS pBIN19 construct among the three plant species studied. Two tobacco plants/lines were generated (Figure 2.13). The attempt at transforming tobacco with EcoRI:DIOX:GUS pCAM was unsuccessful.

Construct	Opium poppy	*Arabidopsis*	Tobacco
DIOX:GUS pCAM	8 plants 1 line	unable to regenerate transformants	Not performed
EcoRI DIOX:GUS pCAM	2 plants 1 line	13 lines 6 weak, 7 strong	unable to regenerate transformants
EcoRI DIOX:GUS pBIN19	Not performed	unable to regenerate transformants	2 plants 2 lines
SpeI DIOX:GUS pCAM	3 plants 2 lines	1 line	Not performed
pCAMBIA 1300	4 plants 1 line	used previously generated lines	used previously generated lines

Table 2.2 Plants regenerated from transformation procedures.

2.3.2 Poppy histochemical staining

Opium poppy DIOX:GUS pCAM transformants demonstrated vascular-specific expression in leaf veins (Figure 2.7 A). Leaf veins were the main tissue showing expression in the DIOX:GUS pCAM transformants. Cutting the tissue (wounding) before staining increased the amount of GUS expression observed (data not shown). Typically a 48 h incubation was necessary to detect a color change in tissues, indicating gene expression. Leaves were also analyzed in the EcoRI:DIOX:GUS transformants and a weak expression of GUS was observed. More mature and laticifer-rich tissues showed the highest levels of GUS expression. For example, GUS expression was observed in the capsule (Figure 2.8) and the petiole (Figure 2.9). The cortex region of the pedicel also shows GUS activity (Figure 2.9 C and D). The SpeI:DIOX:GUS pCAM leaves of transformants were analyzed and showed no GUS activity (Figure 2.10 A).

2.3.3 Fixation of GUS-stained poppy tissue

Some selected transformant tissues were stained for GUS activity, fixed, embedded in plastic resin, and stained with safranin-o to observe cell walls. Overall, GUS staining could be observed following fixation, but anatomy was not well preserved and the cell-types associated with GUS could not be unambiguously identified (Figure 2.10 B, C).

2.3.4 *Arabidopsis* histochemical staining

Seedling and leaf tissues from *A. thaliana* transformants were stained with X-Gluc for 36 to 48 h. The EcoRI:DIOX:GUS pCAM seedlings showed two distinct patterns of expression, classified as strong and weak. An example of the strong

expression was found primarily in the developing vascular regions of the seedling (Figure 2.11 C). A weak pattern of GUS expression was found at the cotyledon tips (Figure 2.11 E). The SpeI:DIOX:GUS pCAM GUS expression pattern occurs in the shoot apical meristem and the root-hypocotyl axis (Figure 2.11 F). The leaf tissues of the EcoRI:DIOX:GUS pCAM were also analyzed and a variety of expression patterns were observed (Figure 2.12 C, D and E). A structure that showed GUS expression in all of the transformants was the trichome (Figure 2.12).

2.3.5 Tobacco histochemical staining

Tobacco EcoRI:DIOX:GUS pBIN19 seedlings, leaves and petioles were tested for GUS expression. Seedlings showed no GUS expression (Figure 2.14). A leaf from EcoRI:DIOX:GUS pBIN19 P2 stained lightly in the vasculature, but the stain was lost during tissue fixation. The EcoRI:DIOX:GUS pBIN19 P2 did show GUS gene expression in the vascular tissue of the petiole after 48 h of staining, although specific cell types could not be determined (Figure 2.15).

2.3.6 PCR of poppy transformants

Multiple attempts were made using PCR to verify the DIOX:GUS pCAM construct was present in poppy genomic DNA. One gel containing positive results was obtained (Figure 2.16). Bands were obtained at approximately 650 bp in the DIOX:GUS pCAM P1, P4, and P5 plants. This correlates with GUS expression observed in poppy transformants.

2.3.7 PCR of tobacco transformants

Only two putative transformants of *EcoRI:DIOX*:GUS pCAM were generated. The bands present for the two plants at approximately 840 bp confirm the presence of the GUS gene (Figure 2.17).

Figure 2.6 Transformed and regenerated opium poppies at Virginia Tech greenhouse. The wild type control poppy is on the far left. All DIOX:GUS pCAM and DIOX:GUS promoter deletion constructs are represented.

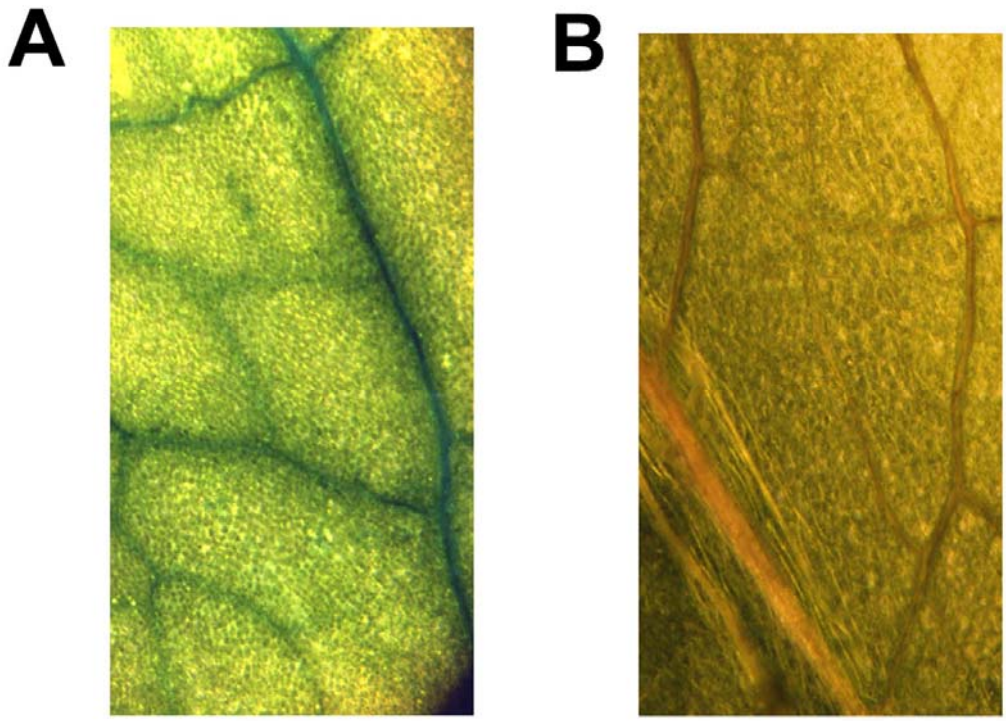

Figure 2.7 Histochemical staining of opium poppy leaves.

A: DIOX:GUS P5 demonstrating vascular specific expression after staining with X-Gluc solution. Photograph taken at 5.0X magnification

B: PsTll, opium poppy T11 field variety showing no staining pattern after staining with X-Gluc. Photograph taken at 5.0X magnification.

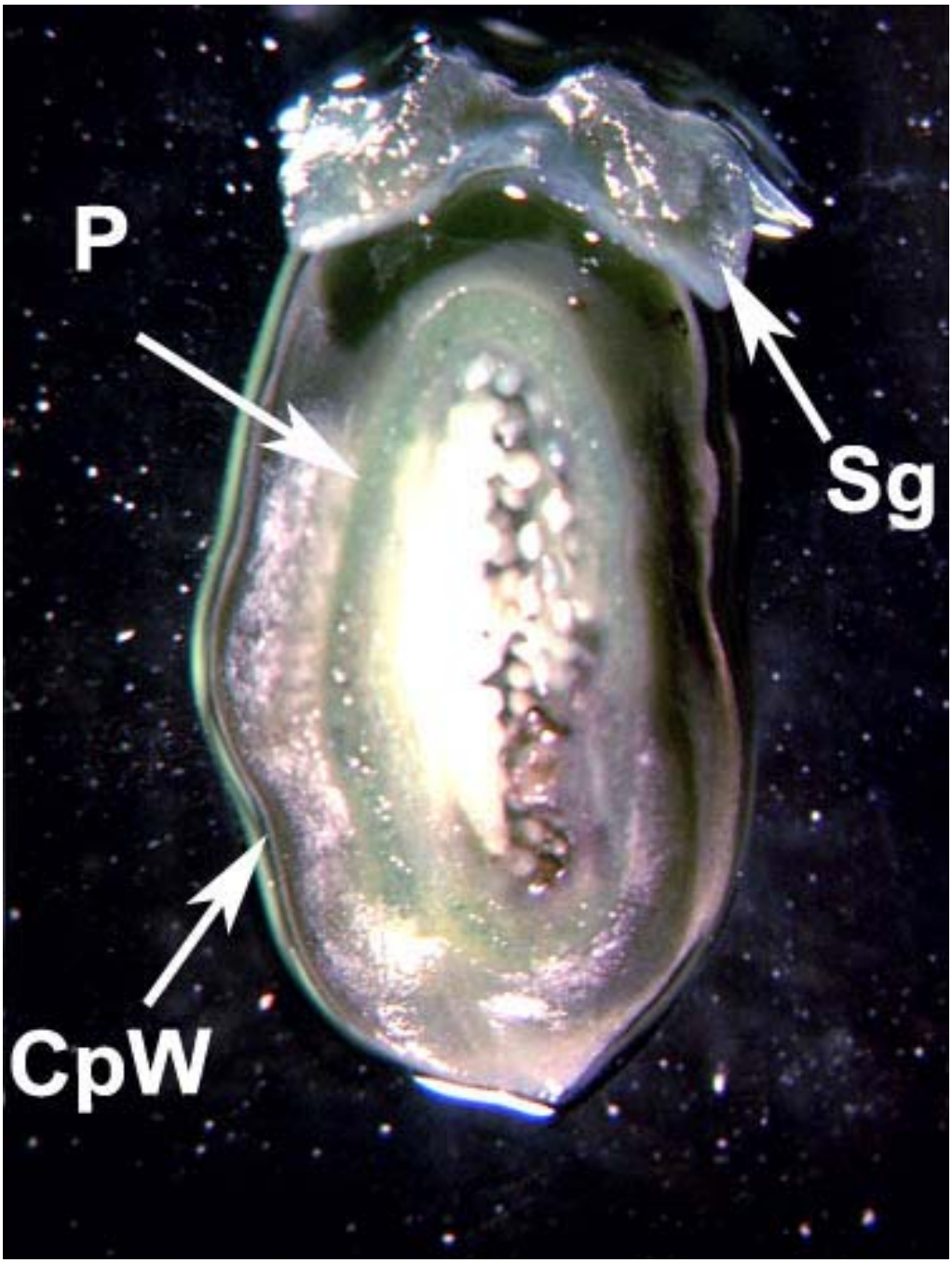

Figure 2.8 Histochemical staining of a longitudinal section of EcoRI:DIOX:GUS P1 opium poppy capsule. Photograph was taken on a stereomicroscope. Staining was observed along the capsule wall (CpW), the placenta (P) and on the stigma (Sg).

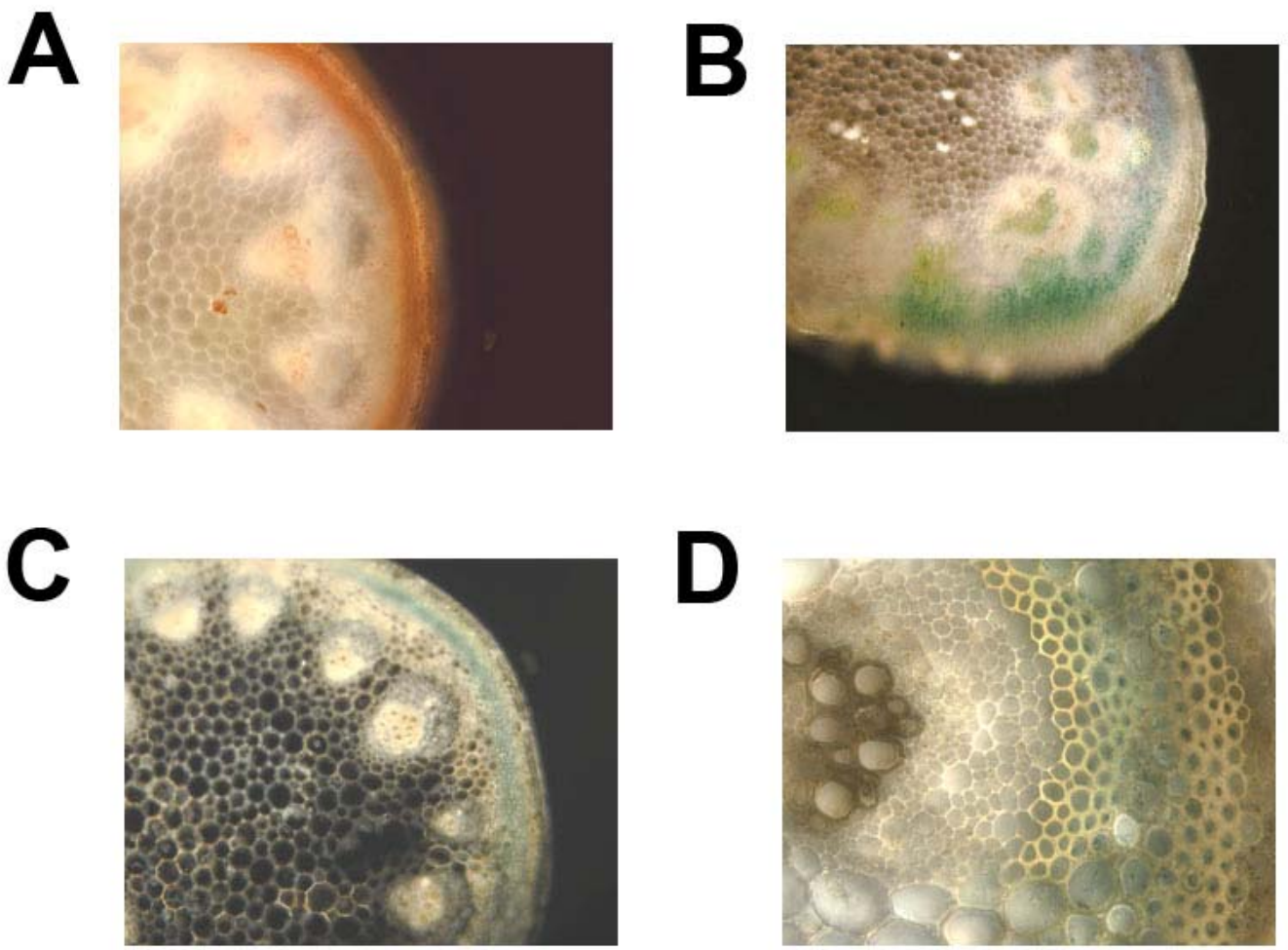

Figure 2.9 Histochemical staining of PsT11 and EcoRI:DIOX:GUS pCAM P1 opium poppy pedicels.
A. PsTll pedicel cross section, no GUS staining was observed. 10X magnification.
B. EcoRI:DIOX:GUS pCAM P2 pedicel 1 cross section, GUS staining was observed in the cortex and surrounding phloem. 10X magnification
C. EcoRI:DIOX:GUS pCAM P2 pedicel 2, GUS staining observed in cortex. 10X magnification
D. EcoRI:DIOX:GUS pCAM P2 pedicel 2, GUS staining observed in cortex near phloem fibers. 20X magnification

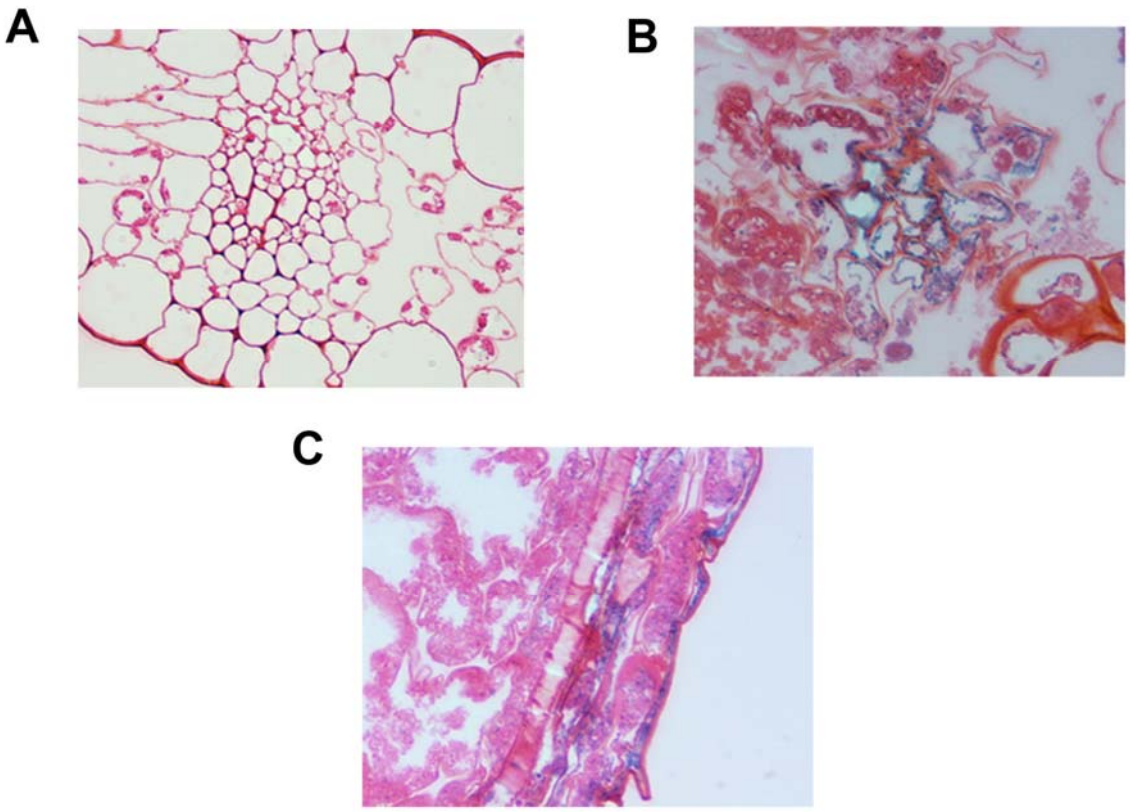

Figure 2.10 Histochemical staining in fixed opium poppy leaf tissues.
A: SpeI:DIOX:GUS pCAM Line 10 P1 demonstrating no GUS staining. Photo taken 40X magnification
B: DIOX:GUS pCAM P7 demonstrating GUS staining in the leaf vein. Photo taken at 100X magnification (oil immersion)
C: DIOX:GUS pCAM P7 demonstrating GUS staining in the leaf epidermis and adjacent to xylem vessels. Photo taken at 100X magnification (oil immersion)

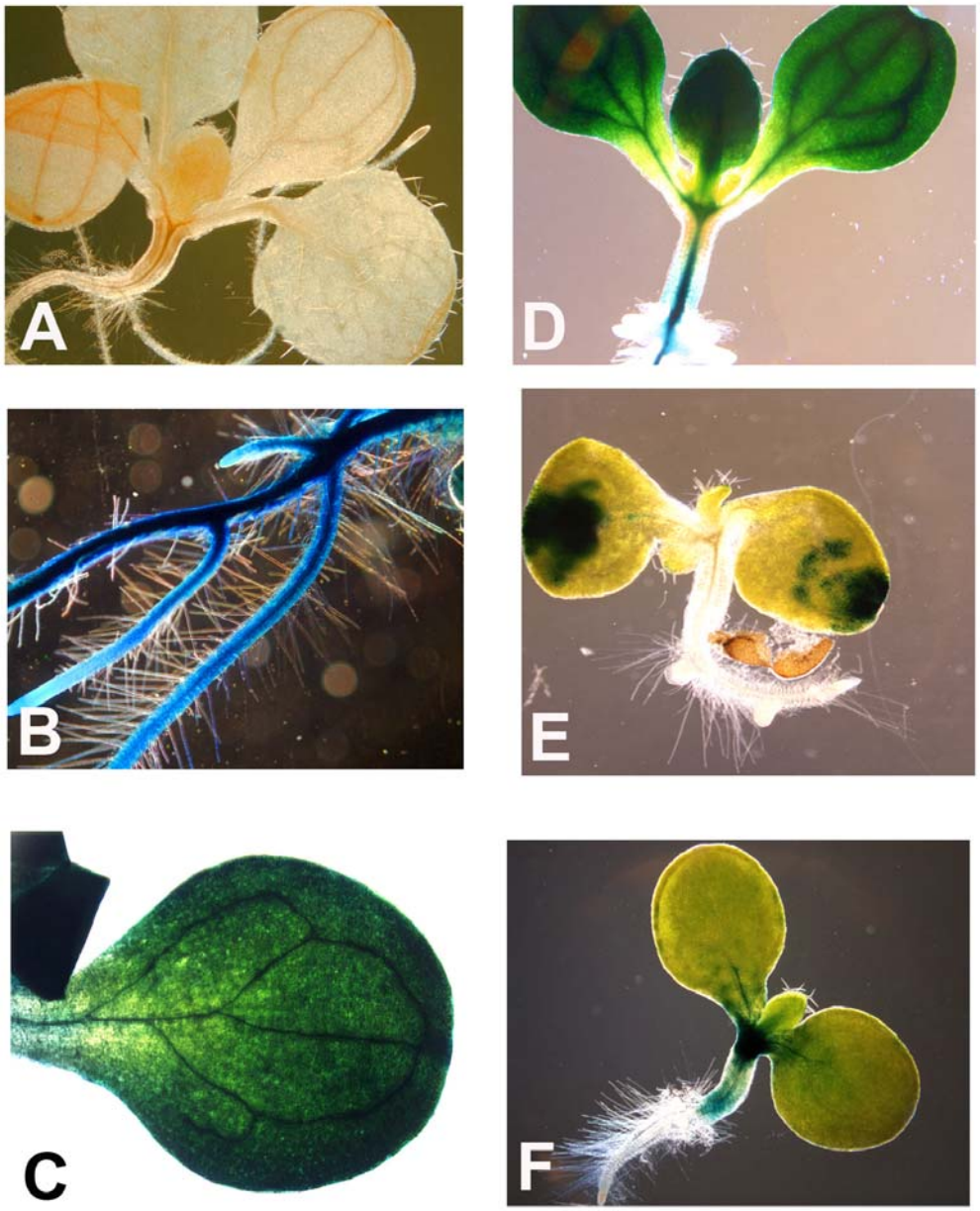

Figure 2.11 Histochemical staining of *Arabidopsis thaliana* seedlings.
All seedlings were stained 10 d after planting on MS media.
A: Wild type *A. thaliana* seedling. 2.5X magnification. No GUS staining observed
B: pBI-121 35S:GUS positive control *A. thaliana* seedling root. 2.5X magnification. GUS staining observed throughout root.
C: pBI-121 35S:GUS positive control *A. thaliana* seedling leaf. 2.5X magnification. GUS staining observed throughout plant and leaf.
D: EcoRI:DIOX:GUS pCAM seedling, example of a strong line 3-11-4 2.5X magnification. GUS staining observed in vascular specific areas. 2.5X
E: EcoRI:DIOX:GUS pCAM seedling, example of a weak line, 2-2-3 GUS staining observed in the tips of cotyledons. 2.5X
F: SpeI:DIOX:GUS pCAM seedling line 2-3-1,GUS staining observed in apical meristem and hypocotyl root axis. 2.5X.

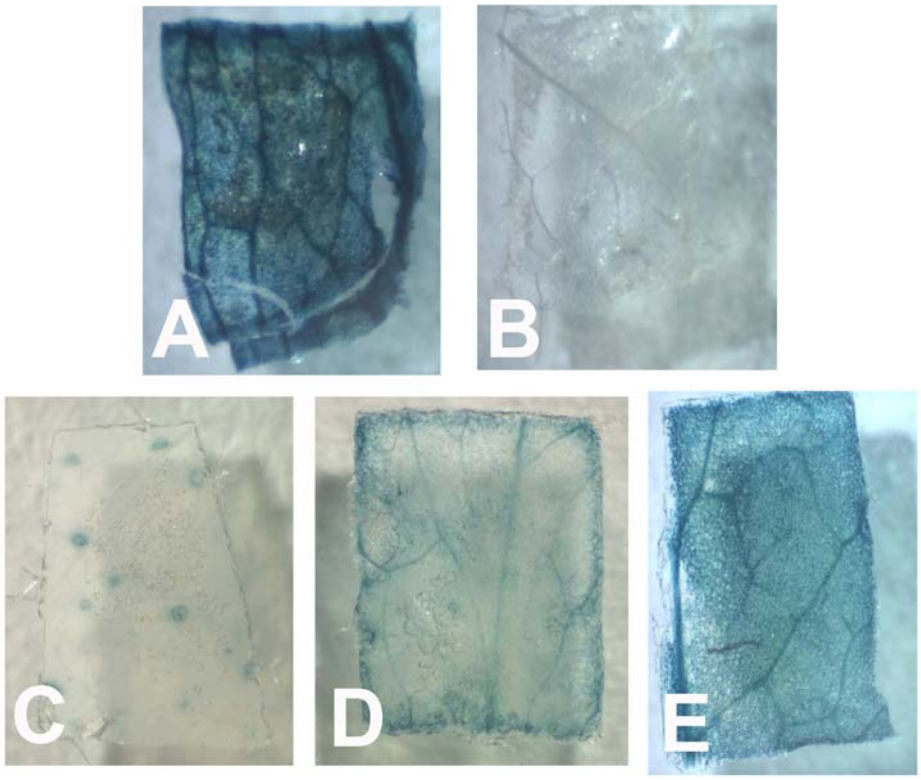

Figure 2.12 Histochemical staining of *A. thaliana* leaf tissues
All tissues were observed 36 h after addition of X-Gluc and observed under a stereomicroscope.
A: pBI-121 leaf tissue showing expression throughout leaf sample.
B: Wild type leaf tissue showing no expression in leaf sample.
C: EcoRI:DIOX:GUS Line 3-7 showing GUS expression in trichomes.
D: EcoRI:DIOX:GUS Line 3-11 showing GUS expression in trichomes and vasculature.
E: EcoRI:DIOX:GUS Line 2-7 showing strong GUS expression throughout leaf sample.

Figure 2.13 EcoRI:DIOX:GUS pBIN19 tobacco growing at the Virginia Tech greenhouses.

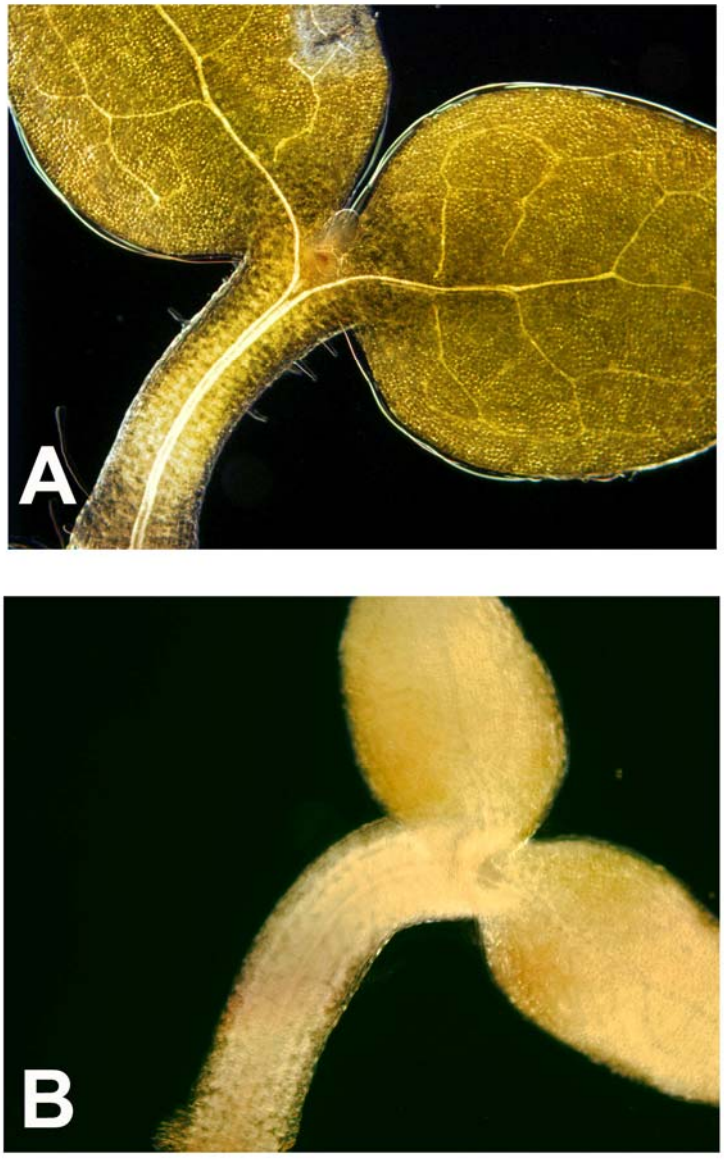

Figure 2.14 Histochemical staining of tobacco seedlings.
A: *Xanthi* wild type tobacco seedling. No GUS expression was observed. 2.5X magnification
B: EcoRI:DIOX:GUS P2 seedling. No GUS expression was observed. 2.5X magnification

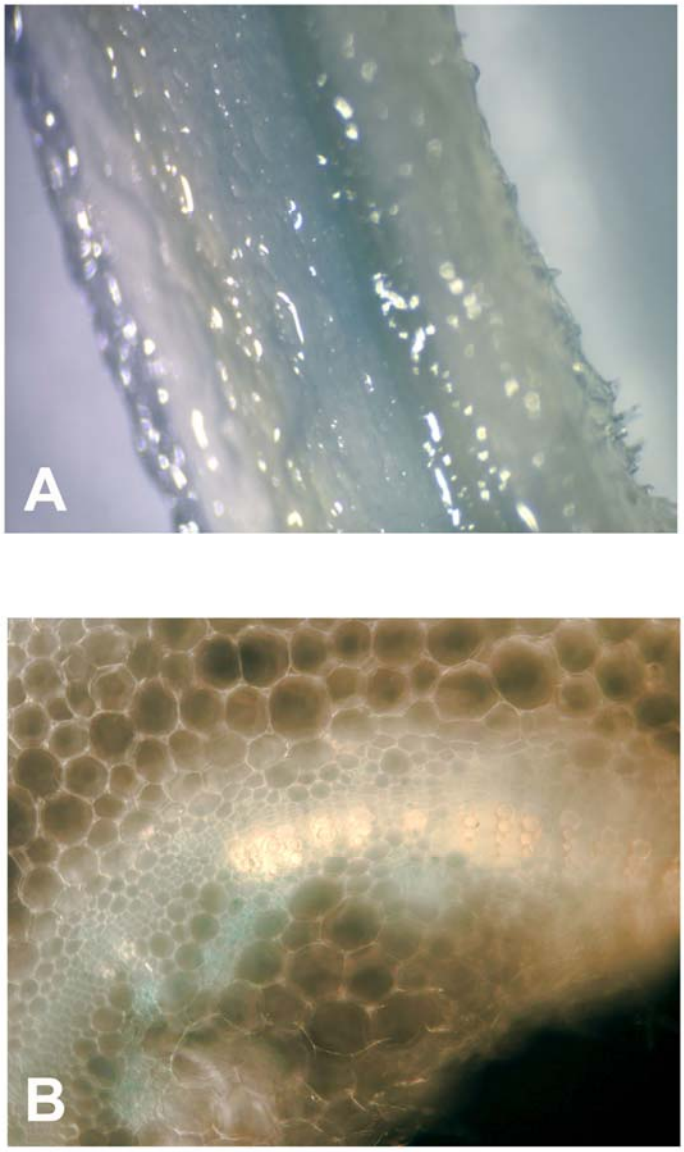

Figure 2.15 Histochemical staining of tobacco EcoRI:DIOX:GUS pBIN19 P2 petiole.
A: Longitudinal section of the petiole. GUS expression is localized to the vasculature. Photo taken using stereomicroscope.
B: Cross-section of petiole, GUS expression is adjacent to the phloem. 20X magnification

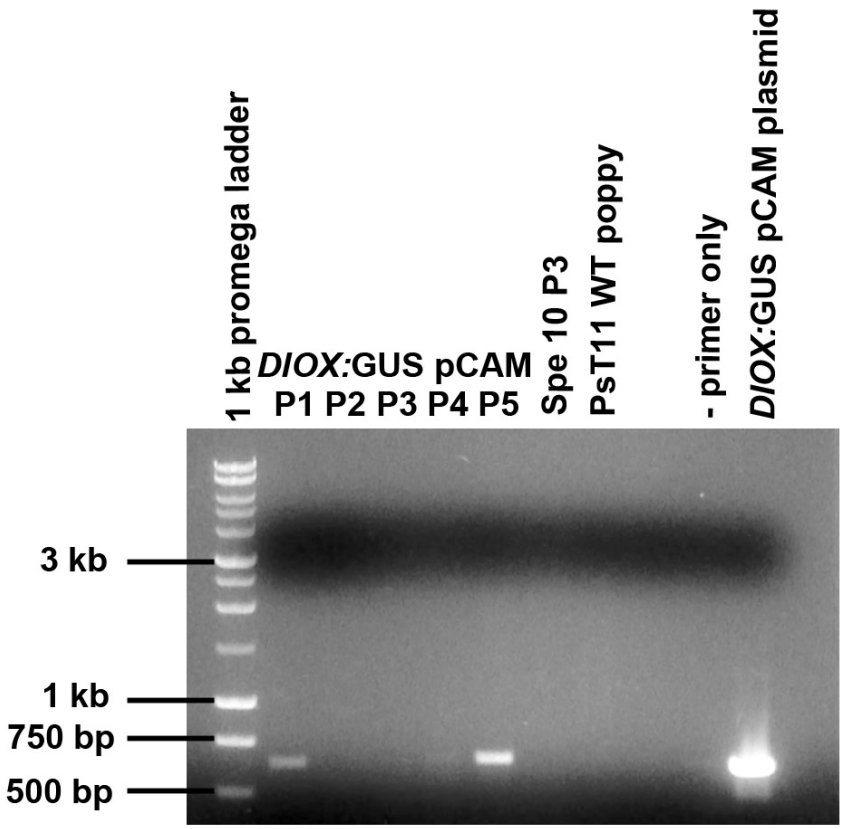

Figure 2.16 PCR amplification of DIOX:GUS fragment in opium poppy transformants.
Expected product size is 651 bp. PsTll wild type genomic DNA was used as a negative control. The DIOX:GUS pCAM plasmid was used as a positive control.

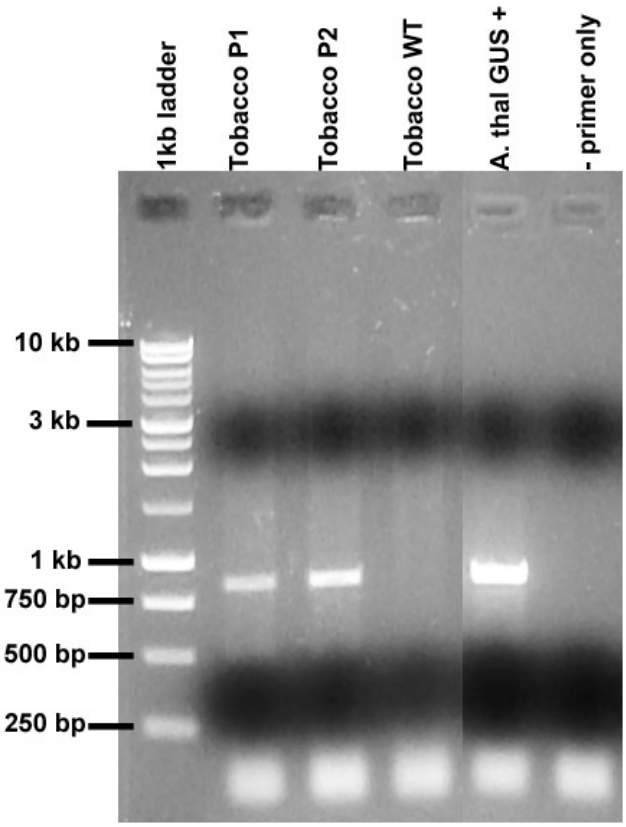

Figure 2.17 PCR amplification of the GUS gene in EcoRI:DIOX:GUS pBIN19 tobacco plants. Expected product size is 841 bp. Untransformed tobacco cv. Xanthi was used as a negative control. *A. thaliana* transformed with pBI-121 GUS was used as a positive control

2.4 DISCUSSION

2.4.1 Plant transformations

Opium poppy biolistic transformation and regeneration has low transformation efficiency in contrast to the *Agrobacterium* methods used for *Arabidopsis* and tobacco, Table 2.2. Transformed plants were generated mostly from one "strong" line of embryos. The EcoRI:DIOX:GUS pBIN19 construct was not used to transform poppy embryos because poppies are naturally resistant to kanamycin (Nessler, unpublished results). Plant regeneration procedures after biolistic transformation from embryos to produce seed took approximately 14-15 months.

Arabidopsis transformations were successful when the promoter:reporter gene constructs were initially transformed into *Agrobacterium.* Overall, the DIOX:GUS

promoter constructs were not easily transformed into *Agrobacterium*. Various methods were attempted before successful bacterial transformation. Due to time constraints, tobacco transformation was only performed once.

2.4.2 Poppy histochemical staining

Vascular-specific expression was clearly observed in the leaf of the DIOX:GUS pCAM transformants. The expression pattern of EcoRI:DIOX:GUS pCAM was not reliably vascular-specific, but it is highly expressed in the laticifer-rich regions of the capsule. GUS expression was found in mature tissues in contrast to young tissues (data not shown). Transformed seedlings were tested for GUS activity and no expression was observed. The SpeI:DIOX:GUS pCAM constructs showed no expression in vascular tissues indicating that the minimum promoter fragment needed for vascular-specific expression in opium poppy is located between the *Eco RI* 951 bp site and the *Spe I* 2420 bp site. It is also important to note that the constructs were linearized before biolistic transformation, and therefore it is unlikely that GUS expression was due to the selectable marker promoter activity. The vascular-specific expression of the DIOX: GUS transformants reinforces other research that shows that alkaloid biosynthesis enzymes are specifically located in the sieve elements which associate with phloem of the vascular system (El-Ahmady and Nessler, 2001; Bird *et al.*, 2003).

2.4.3 Fixation of poppy tissues

Only the SpeI:DIOX:GUS pCAM slide provided images that clearly identified cellular regions. It is likely that using smaller sample sizes when fixing tissues and gentle treatment of tissue would improve the visualization and overall preservation of the cells. Larger sample sizes do not allow the fixative to fully penetrate through the tissue.

2.4.4 *Arabidopsis* histochemical staining

Vascular-specific GUS expression patterns were observed in both *A. thaliana* and opium poppy plant transformants. GUS expression was observed in the SpeI:DIOX:GUS pCAM transformants in contrast to the lack of the poppy expression with the gene construct. This could be due to GUS expression due to the selectable marker promoter of the pCAMBIA 1300 vector (35S) since the vector is not linearized, or to a position effect. Position effect also may be the likely cause of the observed variation in GUS expression patterns in the EcoRI:DIOX:GUS pCAM plant transformants.

2.4.5 Tobacco histochemical staining

Expression of the EcoRI:DIOX:GUS pBIN19 indicates that the GUS expression observed in other species is not dependent upon expression caused by the 35S promoter driving antibiotic selection of the pCAMBIA vector. The vascular expression pattern also indicated that the DIOX: reporter gene fusion functions similarly in different binary vectors.

2.4.6 PCR identification of poppy and tobacco transformants

Identification of poppy transformants by PCR, southern blots, and by gene expression would provide stronger evidence that the DIOX:GUS construct was present in the transformed plant tissues. It was difficult to get consistent results from the PCR experiments and the best results were obtained when poppy DNA was sonicated before use in the PCR experiment. This indicates that the method for extracting the DNA is likely to be problematic for the PCR experiment. PCR was also a useful tool for demonstrating the presence of the GUS gene construct in tobacco, in combination with GUS expression.

2.5 Conclusions and future research

Reporter gene expression driven by the latex-specific *DIOX* promoter expression is consistent with other reports showing that vascular elements are associated with laticifers. It is interesting to note that more visible GUS expression is found in mature tissues, since DIOX is highly homologous to a senescence related protein, SRG1. Future experiments to identify the location of the cell-types associated with latex-specific expression will need to be performed. Fluorescent microscopy could be potentially useful for determining the cell type-specific expression of the *DIOX* promoter. This method has been used for low GUS activity in vascular tissues to determine the minimal cinnamyl alcohol dehydrogenase (CAD) promoter region needed for expression in phloem fibers, but not in secondary xylem (Rech *et al.*, 2003). This method could directly compare research that has localized alkaloid biosynthetic genes to sieve elements of the phloem. Researchers have known that laticifers are associated with the phloem (Nessler and Mahlberg, 1976), but the exact locations of laticifer-specific gene expression within the phloem cell types are not well characterized. The transgenic lines

produced during the present research could serve as a tool for more detailed understanding of laticifers and associated gene expression.

Other quantitative methods of measuring GUS activity, such as the MUG (4-methylumbelliferyl-b-D-glucuronic acid dehydrate) assay (Martin *et al.*, 1992), could be used when measuring activity of new promoter deletion constructs that define potentially important *cis*-regulatory elements.

Much of the research on alkaloid promoters and alkaloid enzyme localization has been performed in transient cell culture systems (Hauschild *et al.*, 1998; Park and Facchini, 2000; Bock *et al.*, 2002). The Nessler laboratory has performed promoter experiments *in planta* (Maldonado-Mendoza *et al.*, 1996; El-Ahmady and Nessler, 2001). The present research was also done *in planta* on multiple species therefore demonstrating more direct physiological relevance. Much of the research in the area of opium poppy focuses on the metabolism of alkaloids, and not the laticifer cell-type that harbors the pharmacologically important alkaloids. Understanding more about the laticifer-specific genes and their expression in this cell type may provide another route for metabolically engineering the accumulation of specific alkaloids of interest.

2.6 REFERENCES

Aravind L, Koonin EV (2001) The DNA-repair protein AlkB, EGL-9 and leprechaun define new families of 2-oxoglutarate- and iron-dependent dioxygenases. Genome Biology **2:** research0007.0001-0007.0008

Bird DA, Franceschi VR, Facchini PJ (2003) A tale of three cell types: Alkaloid biosynthesis is localized to sieve elements in opium poppy. Plant Cell **15:** 2626-2635

Bock A, Wanner G, Zenk MH (2002) Immunocytological localization of two enzymes involved in berberine biosynthesis. Planta **216:** 57-63

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

De Carolis E, Deluca V (1994) 2-Oxoglutarate-dependent dioxygenase and related enzymes: Biochemical characterization. Phytochemistry **36:** 1093-1107

El-Ahmady SH, Nessler CL (2001) Cellular localization of tyrosine decarboxylase expression in transgenic opium poppy and tobacco. Plant Cell Reports **20:** 313-317

Fu XD, Duc LT, Fontana S, Bong BB, Tinjuangjun P, Sudhakar D, Twyman RM, Christou P, Kohli A (2000) Linear transgene constructs lacking vector backbone sequences generate low-copy-number transgenic plants with simple integration patterns. Transgenic Res **9:** 11-19

Gelvin S, Schliperoort R (1988) Plant Molecular Biology Manual. Kluwer Academic Press, New York

Hauschild K, Pauli HH, Kutchan TM (1998) Isolation and analysis of a gene bbe1 encoding the berberine bridge enzyme from the California poppy *Eschscholzia californica*. Plant Mol Biol **36:** 473-478

Horsch RB, Klee HJ, Stachel S, Winans SC, Nester EW, Rogers SG, Fraley RT (1986) Analysis of *Agrobacterium tumefaciens* Virulence mutants in leaf-disks. Proc Natl Acad Sci USA **83:** 2571-2575

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

Kutchan T (1998) The alkaloids: chemistry and pharmacology. *In* G Cordell, ed, The Alkaloids Chemistry and Biology, Vol 50. Academic Press, New York, pp 257-316

Maldonado-Mendoza IE, Lopez-Meyer M, Galef JR, Burnett RJ, Nessler CL (1996) Molecular analysis of a new member of the opium poppy tyrosine 3,4-dihydroxyphenylalanine decarboxylase gene family. Plant Physiol **110:** 43-49

Martin T, Wöhner R-V, Hummel S, Willmitzer L, Frommer W (1992) The GUS reporter system as a tool to study plant gene expression. *In* S Gallagher, ed, GUS Protocols Using the GUS Gene as a Reporter of Gene Expression. Academic Press, London, pp 23-99

Medina-Bolivar F, Wright R, Funk V, Sentz D, Barroso L, Wilkins TD, Petri W, Cramer CL (2003) A non-toxic lectin for antigen delivery of plant-based mucosal vaccines. Vaccine **21:** 997-1005

Nessler CL (1990) Poppies. *In* P Ammirato, DR Evans, WR Sharp, YPS Bajaj, eds, Handbook of Plant Tissue Culture Ornamental Species, Vol 5. Mc-Graw Hill Publishing Company, New York, pp 693-715

Nessler CL (1998) In Vitro Culture Technologies. *In* J Bernáth, ed, Poppy the Genus *Papaver*, Vol 3. Harwood Academic Publishers, Luxembourg, pp 209-217

Nessler CL, Mahlberg PG (1976) Laticifers in stamens of *Papaver somniferum* L. Planta **129:** 83-85

Park S-U, Facchini PJ (2000) *Agrobacterium rhizogenes*-mediated transformation of opium poppy, *Papaver somniferum* L., and California poppy, *Eschscholzia californica* Cham., root cultures. J Exp Bot **51:** 1005-1016

Pilatzke IFC (1999) Gene expression on enzyme activity of cell-wall degrading enzymes in the latex of opium poppy *Papaver somniferum* L. Master's Thesis. Texas A&M University

Prescott A, John P (1996) Dioxygenases: molecular structure and role in plant metabolism. Annu Rev Plant Physiol Plant Mol Biol **47:** 245-271

Rech P, Grima-Pettenati J, Jauneau A (2003) Fluorescence microscopy: a powerful technique to detect low GUS activity in vascular tissues. Plant J **33:** 205-209

Roberts C, Rajagopal S, Smith L, Nguygen T, Yanf W, Nugroho S, Ravi KS, Cao M-L, Vijayachandra K, Patell V, Harcourt R, Dransfield L, Desamero N, Slamet I, Kleese P, Kilian A, Jefferson R (1998) A comprehensive set of modular vectors for advanced manipulations and efficient transformation of plants by both *Agrobacterium* and direct DNA uptake methods. pCAMBIA release manual version 3.05**:** 6

Roemmelt S, Zimmermann N, Rademacher W, Treutter D (2003) Formation of novel flavonoids in apple (*Malus x domestica*) treated with the 2-oxoglutarate-dependent dioxygenase inhibitor prohexadione-Ca. Phytochemistry **64:** 709-716

Rosco A, Pauli HH, Priesner W, Kutchan TM (1997) Cloning and heterologous expression of NADPH-cytochrome P450 reductases from the *Papaveraceae*. Arch. Biochem. Biophys. **348:** 369-377

Vazquez-Flota F, De Carolis E, Alarco AM, DeLuca V (1997) Molecular cloning and characterization of desacetoxyvindoline-4-hydroxylase, a 2-oxoglutarate dependent dioxygenase involved in the biosynthesis of vindoline in *Catharanthus roseus* (L) G. Don. Plant Mol Biol **34:** 935-948

Yun D-J, Hashimoto T, Yamada Y (1992) Metabolic engineering of medicinal plants: Transgenic *Atropa belladonna* with an improved alkaloid composition. Proc Natl Acad Sci USA **89:** 11799-11809

Chapter 3

ANALYSIS OF LATEX-SPECIFIC PROMOTERS

3.1 INTRODUCTION

The major latex proteins (MLPs) of opium poppy were first identified and characterized in 1985 (Nessler, *et al.)*. These proteins were named MLPs because they are the most abundant class of proteins found in the latex (Nessler *et al.*, 1985). MLPs are low weight polypeptides with a molecular mass of approximately 20 kDa and compose up to 50% of soluble latex protein (Nessler *et al.*, 1990). Protein gel blots and immunolabeling studies have shown that the MLPs are only found in latex (Nessler *et al.*, 1985; Nessler and Griffing, 1989). RNA gel blots have shown that MLP messages are very abundant in the latex serum (Nessler and Vonder Harr, 1990). The capsule, root and leaves contain laticifers and show MLP expression that directly correlates to laticifer distribution (Nessler and Vonder Harr, 1990). Several members of the *MLP* family have been cloned, and were classified into two subfamilies based on their sequence (Nessler and Burnett, 1992; Nessler, 1994). The genomic clones of interest for this project were designated *MLP146* and *MLP149*, and are physically linked (Nessler, 1994). This linkage facilitated the isolation of the 5' upstream region of two known latex-specific genes.

MLP-like proteins (MLP-LPs) are not exclusive to plants that contain laticifers, and homologs are found in *Arabidopsis thaliana*, melon, soybean, tobacco, and wild strawberry (Osmark *et al.*, 1998). A gene expressed during fruit ripening of melon (*Cucumis melo* L.) shares sequence homology and expression pattern similar to that of the *MLPs* in opium poppy (Aggelis *et al.*, 1997). Also, mRNAs isolated during the fruit ripening of wild strawberry (*Fragaria vesca L.*) have homology to the MLPs of opium poppy (Nam *et al.*, 1999). Another MLP-LP, Pp-MLP1, identified during peach (*Prunus persica*) fruitlet abscission, shows significant similarity to the MLP homologs, and its gene transcript was highly accumulated in the cells of the fruit pedicel and to a lesser extent in epicotyls, stems and roots (Ruperti *et al.*, 2002). The suggested function of these homologous proteins found during fruit ripening is to participate in wound healing and the storage of secondary metabolites (Nam *et al.*, 1999). This is similar to the

function of latex in opium poppy, which is under positive osmotic pressure so that it covers wound surfaces when injured (Nessler *et al.*, 1990). MLPs have been shown to share sequential and structural homology with the family of the intracellular pathogenesis related proteins (IPR/PR10) (Osmark *et al.*, 1998). IPR protein sequences are only 25% homologous to MLPs, but they have similar structures, sizes, and pIs (Osmark *et al.*, 1998). The gene expression of soybean gene *Msg*, another MLP-LP, shows a complex developmental pattern that parallels laticifer distribution in opium poppy (Stromvik *et al.*, 1999). The model system *Arabidopsis thaliana* also has homologous proteins and its large MLP family has been used for phylogenic studies. The study of the MLP family demonstrated the rapid evolution of the family due to polyploidy (Cannon and Young, 2003). This rapid evolution and distribution of MLP homologs in various species suggests a still undiscovered function in plant tissues.

3.2 MATERIALS AND METHODS

3.2.1 RNA isolation

RNA was isolated from the capsules of United Nations line 186 (UNL186) of *Papaver somniferum*. The capsule contains the largest amount of laticifers, and therefore is enriched for latex. One hundred mg of UNL186 capsule was frozen in liquid nitrogen and thoroughly ground using a mortar and pestle. A 5:1 (v:w) ratio of TRI reagent (Sigma, St. Louis, MO) to sample material was added and mixed by inversion. An equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) was added to the sample and mixed vigorously. The mixture was centrifuged at 10,000 x g for 10 min at 4°C. The top, aqueous layer, was saved and an equal volume of phenol:chloroform:isoamyl alcohol was added to it and centrifuged at 10,000 x g for 10 min at 4°C. This step was repeated until the top layer was clear. The aqueous supernatant was removed without disturbing the interface and an equal volume of chloroform was added. The solution was centrifuged at 10,000 x g for 10 min at 4°C. Again the top layer was removed without touching the interphase and placed in a new tube. A 3:1 (v:v) of isopropanol to RNA solution was added to the extract, mixed and the sample was placed at -20°C for 30 min. The tube was then centrifuged at 12,000 x g for 10 min at 4°C, the isopropanol was removed, and sample was air dried. The pellet was resuspended in 35 µl of 1 mM sodium citrate, pH 6.4 (RNA storage solution, Ambion, Austin, TX). The RNA was analyzed by

spectrophotometric analysis at A260/A280 to determine purity. Approximately 1.5 μg of RNA sample was prepared for electrophoresis by addition of 3 volumes of NorthernMax™ Formaldehyde Load Dye (Ambion, Austin, TX) to RNA solution. The mixture was then incubated at 65°C for 15 min and centrifuged and placed on ice. Ethidum bromide at a final concentration of 10 μg/ml was added in order to visualize RNA after electrophoresis. The RNA sample was run on a 1.2% formaldehyde MOPS (morpholinopropanesulfonic acid) gel at 90 V for 1 h and 45 min. The MOPS formaldehyde gel was prepared according to (Sambrook and Russell, 2000). If distinct ribosomal bands were observed in the RNA samples, then the preparation was used for reverse transcriptase-polymerase chain reaction (RT-PCR).

3.2.2 Semi-quantitative RT-PCR

DNA contaminating the RNA solution was removed by DNase I treatment with DNA-*free*™ (Ambion, Austin, TX). A 0.1 volume of 10X DNase I buffer and 1 μl DNase I was added to the RNA and incubated at 37°C for 30 min. Then 5 μl of DNase I inactivation reagent was added and incubated at room temperature for two min. The mixture was centrifuged at 10,000 x g for 1 min to pellet the DNase I inactivation reagent.

The RETROscript™ (Ambion, Austin, TX) first strand synthesis kit for RT-PCR was used to make cDNA as per the manufacturer's protocol using the oligo$(dT)_{18}$ primers (Ambion, Austin, TX). The same amount of starting cDNA, 3.0 μl, was then added to a PCR reaction using one of the cMLP146 or cMLP149 primers (Table 3.1). The Robocycler® gradient temperature cycler (Stratagene, La Jolla, CA) allowed for simultaneous amplification even though primers anneal at different temperatures. PCR conditions included an initial denaturation step of 2 min at 95°C. Cycling conditions consisted of 30 cycles of: 1 min denaturation at 95°C, 1 min annealing temperature specific to each set of primers (Table 3.1), 1 min 72°C extension and a final extension of 72°C for 10 min. Samples were removed at the 15, 20, 25 and 30 cycle time points and immediately placed at 4°C. The cMLP146 and cMLP149 primers flank an intron and were designed to discriminate between genomic DNA and synthesized cDNA (GenBank Accession numbers L06467 and L06469). A genomic contaminated cMLP146 PCR

product would be l21bp larger than the RT-PCR product and the cMLP149 PCR product would be 109 bp larger than the RT-PCR product. DNA contamination was also tested by using the RNA without the RT step as a template for the primers. The primers were designed specifically to differentiate between the two closely related sequences. Genomic wild type DNA (UNL186) and plasmid DNA of the MLP DNA (gDNA clone in pBluescript) were used as controls to show the size difference between gDNA and RT-PCR products. The experiment was done two times to verify the results.

Primer Name	Sequence	Annealing temperature	Product size
cMLP146 forward	CCA AGC GTA ATC CCT CAT ATT TA	59	
cMLP146 reverse	GTC TTC GGT GAT CTG TTG GTA G	59	375 bp
cMLP149 forward	CCA AGC GTA ATC CCT CAT ATT GT	66	
cMLP149 reverse	GTC TTC GGT GGC TTG ATT GCA C	66	375 bp

Table 3.1- Primer sequences, annealing temperatures, and PCR product size used to semi-quantitate the expression levels of MLP146 and MLP149.

3.2.3 Sequencing of promoters

Three latex-specific 5' upstream regions were sequenced in order to analyze the promoters for any conserved regions or domains that may contribute to the understanding of tissue-specific expression. The promoter regions analyzed include the *2L1* or *DIOX*, a gene homologous to *SRG1* and a 2- oxoglutarate dioxygenase, and the *MLP146* and *MLP149* promoters that were previously isolated (Nessler *et al.*, 1990). The 5' upstream regions were in the plasmid pBluescript KS and were sequenced upstream of the coding region by primer walking. (See details in appendix A). Sequencing was performed using the BigDye Terminator v.3.1 cycle sequencing (Applied Biosystems, Foster City, CA) on an ABI Prism 3100 at the VBI Core Laboratory facility at Virginia Tech. Analysis of enzyme restriction sites and sequence alignments were performed with the Mac Vector 7.2.2 software (Accelrys, San Diego, CA) using the CLUSTALW program (Thompson *et al.*, 1994).

3.2.4 *Cis*-element analysis

The 5' upstream regions of these promoters were analyzed using the PLACE signal scan program: http://www.dna.affrc.go.jp/PLACE (Higo *et al.*, 1999). The scan was done on 2540 bp of the DIOX 5'upstream region, 2081 bp of the MLP146 promoter and 1332 bp of the MLP 149 promoter (Appendix B). Output data in html format was

transferred to plain text in Microsoft Excel (Pullman, WA) spreadsheets. Elements shared by two or more sequences were identified and classified into groups. Classes were designed on the basis of keywords located in the PLACE database and references.

3.3 RESULTS

3.3.1 Semi-quantitative RT-PCR

After 25 cycles of PCR there was a visible difference in the abundance of the RT-PCR product between *MLP146* and *MLP149*. The *MLP149* RT-PCR product has a band of approximately 375 bp at 25 cycles whereas *MLP146* does not produce a band until 30 cycles (Figure 3.1). This suggests that the *MLP149* gene is expressed higher than MLP146 in capsule tissues. No bands from DNA contamination were observed.

3.3.2 *Cis*-element identification

Sixty-five conserved elements were identified between the three latex-specific upstream regions (Table 3.2). There are basic conserved regulation elements such as TATA and CAAT boxes, and other elements involved in critical functions such as carbon metabolism. Other relevant *cis* elements shared among the three promoters include the MYB (MYB are transcription factors that include signaling factors involved in many processes, including secondary metabolism and named after their similarity to the avian myeloblastosis proto-oncogene), light, stress/defense and vascular classes. The location of all shared elements among the three promoters is included in Appendix C.

The *DIOX* promoter contains a vascular-specific element previously characterized in tobacco (Baumann *et al.*, 1999). The expression pattern of the *EcoDIOX*:GUS pCAM and *SpeDIOX*:GUS pCAM in all species tested confirms the regulation by the vascular-specific *cis* element, described in Chapter 2. This element NTBFFARROLB (*Nicotiana tabacum* domain B binding factor of *rolB*) associates with the NtBBF protein found in the meristematic cells in the cortex and vascular tissue (Figure 3.2) (Baumann *et al.*, 1999). The *DIOX* promoter also contains *cis* elements needed for the activation by WRKY transcription factors associated with pathogen response. The WRKY proteins are a family of plant-specific zinc-finger type factors implicated in the regulation of genes associated with pathogen defense (Eulgem *et al.*, 1999). The WRKY proteins are named after their conserved N-terminal amino acid sequence WRKYGQK (Eulgem *et al.*, 2000).

There are shared *cis* elements between the *DIOX* and *MLP149* upstream regions. These elements are the HDZIP2ATAHB2, a *cis* element that regulates a homeodomain leucine zipper II subfamily (HD-Zip II) (Ohgishi *et al.*, 2001). Also, another *cis* element that shares its binding site among DIOX and MLP146 is the element MYBPZM associated with a MYB transcription factor from *Zea mays* protein P that is associated with secondary metabolism (Grotewold *et al.*, 1994).

Both *MLP* upstream regions share elements. These include the SEBFCONSSTPR10A element that represses the *PR10a* defense gene and the TGTCACACMCUCUMISIN fruit-specific element (Figure 3.2) (Boyle and Brisson, 2001; Yamagata *et al.*, 2002). They contain the same WRKY *cis* elements found in the *DIOX* promoter.

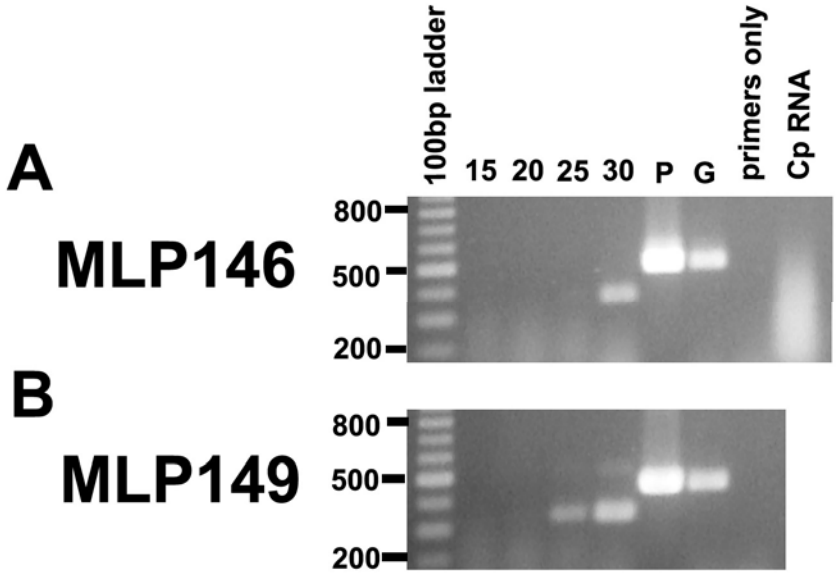

Figure 3.1. MLP146 vs. MLP149 semi-quantitative RT-PCR
A. MLP146 expression at different PCR cycle numbers.
PCR cycle numbers are: 15 cycles, 20 cycles, 25 cycles, and 30 cycles.
P, plasmid; G, UNL186 genomic DNA; Cp RNA, capsule RNA.
B. MLP149 expression at different PCR cycle time points.
PCR cycle time points are: 15 cycles, 20 cycles, 25 cycles, and 30 cycles.
P, plasmid; G, UNL186 genomic DNA

Table 3.2 *Cis* elements common among DIOX and MLP promoters

CLASS	*CIS* ELEMENT	SEQUENCE	DIOX (2540bp)	MLP146 (2081bp)	MLP 149	REFERENCE
AP2 related	RAV1AAT	CAACA	x	x	x	Kagaya *et al.* Nucleic Acids Res 27:470-478 (1998)
Basic regulation	CAATBOX1	CAAT	x	x	x	Shirsat *et al.* Mol Gen Genet 215:326-331 (1989)
	POLASIG1	AATAAA	x	x	x	Heidecker and Messing Annu Rev Plant Physiol 37:439-466 (1986)
	POLASIG2	AATTAAA	x	x	x	O'Neill *et al.* Mol Gen Genet 221:235-244 (1990)
	POLASIG3	AATAAT	x	x	x	Heidecker and Messing Annu Rev Plant Physiol 37:439-466 (1986)
	TATABOX2	TATAAAT	x	x	NP	Shirsat *et al.* Mol Gen Genet 215:326-331 (1989)
	TATABOX3	TATTAAT	x	NP	x	Wang *et al.* Plant Physiol 108:829-830 (1995)
	TATABOX5	TTATTT	x	x	x	Tjaden *et al.* Plant Physiol 108:1109-1117 (1995)
	TRANSINITMONOCOTS	RMNAUGGC	x	x	NP	Joshi et al. Plant Mol Biol 35:993-1001 (1997)
C metabolism	AMYBOX1	TAACARA	NP	x	x	Huang *et al.* Plant Mol Biol 14:655:668 (1990)
	PYRIMIDINEBOXOSRAMY1A	CCTTTT	x	x	x	Morita *et al.* FEBS Lett 423:81-85 (1998)
	TAAAGSTKST1	TAAAG	x	x	x	Plesch *et al.* Plant J 28:455-464 (2001)
	DOFCOREZM	AAAG	x	x	x	Yanagisawa *et al.* Plant J 21:281-288 (2000)
Flowering	CARGCW8GAT	CWWWWWWWWG	NP	x	x	Tang and Perry J Biol Chem 278:28154-28159 (2003)
Fruit	TGTCACACMCUCUMISIN	TGTCACA	NP	x	x	Yamagata *et al.* J Biol Chem 277:11582-11590 (2002)
Gibberellin	GAREAT	TAACAAR	NP	x	x	Ogawa *et al.* Plant Cell 15:1591-1604 (2003)
Heat shock	CCAATBOX1	CCAAT	xx	x	x	Rieping *et al.* Mol Gen Genet 231:226-232 (1992)
Homeodomain	HDZIP2ATATHB2	TAATMATTA	x	NP	x	Ohgishi *et al.* Plant J 25:389-398 (2001)
Homologous pectate lyase	GTGANTG10	GTTA	x	x	x	Rogers *et al.* Plant Mol Biol 45:577-585 (2001)
Light	-10PEHVPSBD	TATTCT	x	x	x	Thum *et al.* Plant Mol Biol 47:353-366 (2001)
	GATABOX	GATA	x	x	x	Lam and Chua Plant Cell 1:1147-1156 (1989)
	GT1CONSENSUS	GRWAAW	x	x	x	Terzaghi and Cashmore Annu Rev Plant Physiol Plant Mol Biol 46:445-474 (1995)
	GT1CORE	GGTTAA	x	NP	x	Green *et al.* EMBO J 7:4035-4044 (1988)
	IBOXCORE	GATAA	x	x	x	Terzaghi and Cashmore Annu Rev Plant Physiol Plant Mol Biol 46:445-474 (1995)
	INRNTPSADB	YTCANTYY	x	x	x	Nakamura *et al.* Plant J 29:1-10 (2002)
	REALPHALGLHCB21	AACCAA	x	x	x	Degenhardt and Tobin Plant Cell 8:31-41 (1996)
	TBOXATGAPB	ACTTTG	x	NP	x	Chan *et al.* Plant Mol Biol 46:131-141 (2001)
	ZDNAFORMINGATCAB1	ATACGTGT	x	x	NP	Yadav *et al.* Plant J 31:741-753 (2002)
MYB	MYBST1	GGATA	x	x	x	Baranowskij *et al.* EMBO J 13:5383-5382 (1994)
MYB/ABA	MYB1AT	WAACCA	x	x	x	Abe *et al.* Plant Cell 15:63-78 (2003)
MYB/c metabolism	MYBGAHV	TAACAAA	NP	x	x	Morita *et al.* FEBS Lett 423:81-85 (1998)
	TATCCAOSAMY	TATCCA	x	x	x	Lu *et al.* Plant Cell 14:1963-1980 (2002)
MYB/secondary metabolism	MYBZM	CCWACC	x	NP	x	Grotewald *et al.* Cell 76:543-553 (1994)
	MYBCORE	CNGTTR	x	x	x	Lusher and Eiseman Genes Dev 4:2235-2241 (1990)
	MYBPLANT	MACCWAMC	x	NP	x	Sablowski *et al.* EMBO J 13:128-137 (1994)
MYB/Stress	MYB2AT	TAACTG	x	x	NP	Urao *et al.* Plant Cell 5:1529-1539 (1993)
MYB/Stress/ABA	MYB2CONSENSUSAT	YAACKG	x	x	x	Abe et al. Plant Cell 15:63-78 (2003)
MYC/Stress	MYCATERD1	CATGTG	x	x	x	Simpson *et al.* Plant J 33:259-270 (2003)
MYC/Stress/ABA	MYCATRD22	CACATG	x	x	x	Abe *et al.* Plant Cell 9:1859-1868 (1997)
	MYCCONSENSUSAT	CANNTG	x	x	x	Abe *et al* . Plant Cell 15:63-78 (2003)
Plastid	BOXIINNTPATPB	ATAGAA	x	x	x	Kapoor and Sugiura Plant Cell 11:1799-1810 (1999)
	S1FBOXSORPS1L21	ATGGTA	x	x	NP	Zhou *et al.* J Biol Chem 267:23515-23519 (1992)
Pollen	POLLENLELAT52	AGAAA	x	x	x	Bate and Twell Plant Mol Biol 37:859-869 (1998)
	QELEMENTZMZM13	AGGTCA	NP	x	x	Hamilton *et al.* Plant Mol Biol 38z:663-669 (1998)
Root	ROOTMOTIFTATPOX1	ATATT	x	x	x	Elmayan and Tepfer *et al.* Trangenic Res 4:388-396 (1995)
Secondary metabolism	PALBOXAPC	CCGTCC	NP	x	x	Logeman *et al.* Proc Natl Acad Sci USA 92:5905-5909 (1995)
Secondary metabolism/light/defense	CACGTGMOTIF	CACGTG	x	x	NP	Siberil *et al.* Plant Mol Biol 45:477-488 (2001)
Seed	-300ELEMENT	TGHAAARK	x	x	x	Thomas and Flavell Plant Cell 2:1171-1180 (1990)
	AACACOREOSGLUB1	AACAAAC	x	NP	x	Wu *et al.* Plant J 23:415-421 (2000)
	CANBNNAPA	CNAACAC	x	x	NP	Ellerstrom *et al.* Plant Mol Biol 32:1019-1027 (1996)
	CICADIANELHC	CAANNNNATC	x	x	NP	Piechulla *et al.* Plant Mol Biol 38:655-662 (1998)
	EBOXBNNAPA	CANNTG	x	x	x	Stalberg *et al.* Planta 199:515-519 (1996)
	SEF4MOTIFGM7S	RTTTTTR	x	x	x	Allen *et al.* Plant Cell 1:623-631 (1989)
Seed/ABA	DPBFCOREDCDC3	ACACNNG	x	x	NP	Kim *et al.* Plant J 11:1237-1251 (1997)
	RYREPEATBNNAPA	CATGCA	x	NP	x	Ezcurra *et al.* Plant Mol Biol 40:699-709 (2000)
Stress	ABRELATERD1	ACGTG	x	x	x	Simpson *et al.* Plant J 33:259-270 (2003)
	ACGTATERD1	ACGT	x	x	x	Simpson *et al.* Plant J 33:259-270 (2003)
	DRECRTCOREAT	RCCGAC	x	NP	x	Dubouzet *et al.* Plant J 33:751-763 (2003)
	LTRECOREATCOR15	CCGAC	x	x	x	Baker *et al.* Plant Mol Biol 24:701-713 (1994)
Stress/defense	ASF1MOTIFCAMV	TGACG	x	x	NP	Redman *et al.* Plant Cell Rep 21:180-185 (2002)
	ELRECOREPCRP1	TTGACC	x	x	x	Rushton *et al.* EMBO J 15:5690-5700 (1996)
	SEBFCONSSTPR10A	YTGTCWC	NP	x	x	Boyle and Brisson Plant Cell 13:2525-2537 (2001)
	WBBOXPCWRKY1	TTTGACT	x	x	x	Eulgem *et al.* EMBO J 18:4689-4699 (1999)
	WBOXATNPR1	TTGAC	x	x	x	Yu *et al.* Plant Cell 13:1527-1540 (2001)
	WBOXHVISO1	TGACT	x	x	x	Sun *et al.* Plant Cell 15:2076-2092 (2003)
Vascular	NTBBF1ARROLB	ACTTTA	x	x	NP	Baumann *et al.* Plant Cell 11:323-333 (1999)

Table 3.2- Table of shared cis elements among latex specific promoter regions. x= element present NP= element is not present

ABA= Abscisic Acid
AP2=APETELA2 transcription factor
MYB= nuclear transcription factor originally found as a protooncogene
MYC= nuclear transcription factor originally found as a protooncogene

W= A or T
Y= C or T
R= A or G
N= A or C or G or T
M= A or C

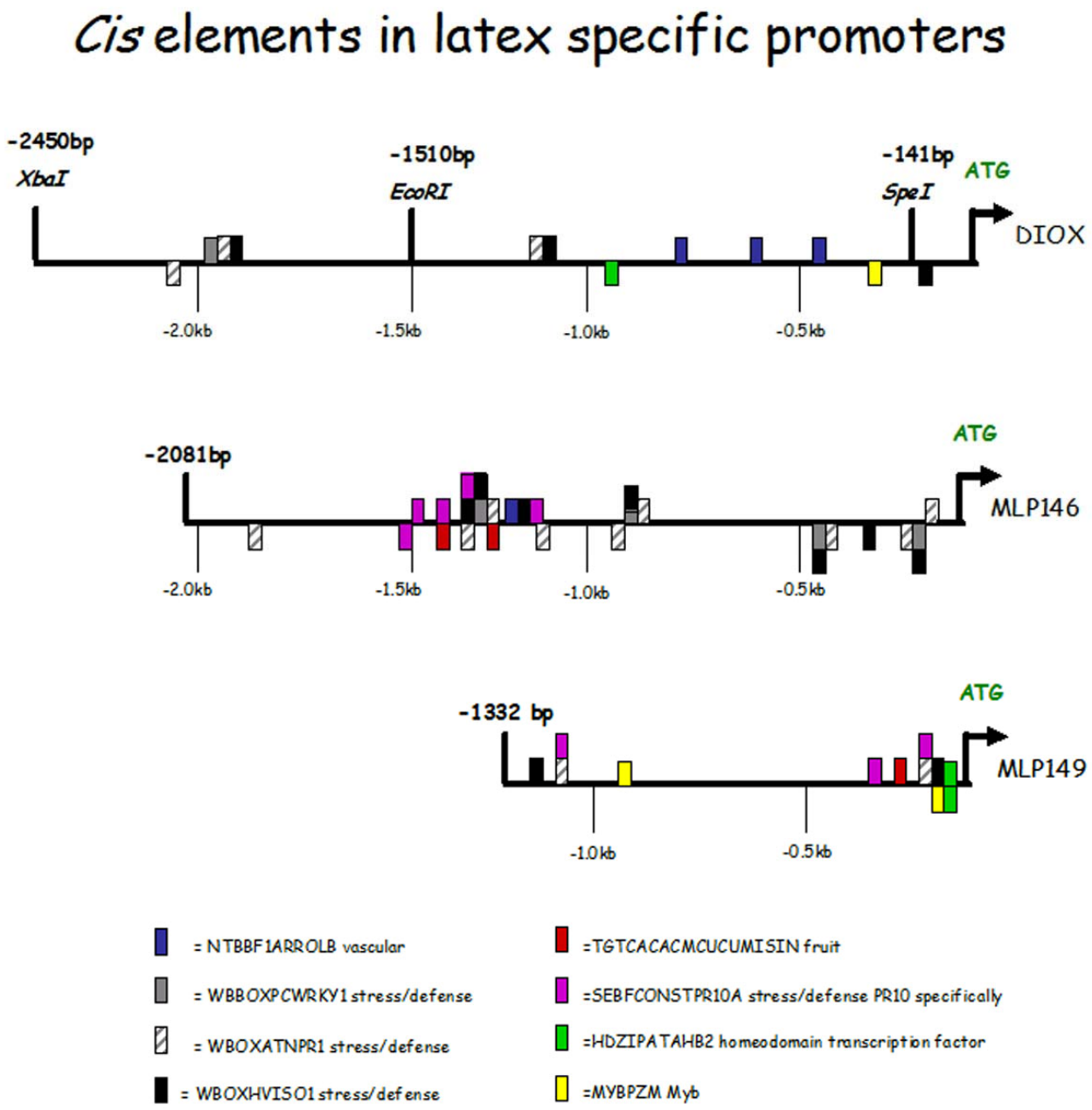

Figure 3.2 Diagram of approximate *cis* element locations. The factors listed are likely to be involved in tissue-specific gene expression of latex.

3.4 DISCUSSION

3.4.1 RT-PCR

The strong RT-PCR signal from the *MLP149* promoter indicates relative promoter strength greater than the *MLP146* gene (Figure 3.1). Strong tissue-specific promoters are ideal candidates for metabolic engineering of secondary metabolites. The strong localized expression would drive high levels of enzyme production leading to the accumulation of desired compounds in a specific cell type or tissue.

3.4.2 *Cis*-element comparative promoter analysis

Cis elements are regulatory elements located proximal to the coding region of a gene (Buchanan *et al.*, 2000). The PLACE (Plant *cis* acting regulatory DNA elements) database was created as a compilation of plant-specific nucleotide sequence motifs and a tool for doing homology searches to estimate the mode of gene regulation (Higo *et al.*, 1999). The PLACE database (http://www.dna.affrc.go.jp/PLACE/) was used as a research tool here rather than another available plant *cis*-acting regulatory element database designated PlantCARE (http://intra.psb.ugent.be:8080/PlantCARE/) (Rombauts *et al.*, 1999). The PLACE database allowed for ease of use in comparative promoter analysis with export of data into a Microsoft Excel (Pullman, WA) file.

Using the PLACE database, sixty-five elements were found to be shared in two out of the three promoters analyzed (Table 3.2). The majority of the predicted *cis*-elements include the functions of basic gene regulation and carbon metabolism. Basic regulation elements include the TATA box and CAAT box which are involved in the positioning of RNA polymerase II (RNA pol II) to initiate transcription (Buchanan *et al.*, 2000).

3.4.2.1 MYB signaling factors

Another large class of *cis*-acting regulatory elements classified in Table 3.2 is the MYB class. The MYB class of proteins was first discovered when the C1 gene in *Zea mays* was sequenced and was recognized to have high similarity with the well known oncogene c-MYB isolated from the cellular avian myeloblasts (Jin and Martin, 1999). MYB proteins have particular importance in plants because of the number of genes involved and their roles that control plant-specific processes (Martin and Paz-Ares, 1997). A distinct function of MYB proteins in plants is controlling secondary

metabolism (Martin and Paz-Ares, 1997). The MYB domain of these proteins is comprised of a region of 52 amino acids that binds DNA in a sequence-specific manner (Jin and Martin, 1999). The typical MYB protein (c-MYB) contains three repeats of the MYB domains (R1,R2,R3) and each imperfect repeat adopts a helix-turn-helix fold (HTH) (Martin and Paz-Ares, 1997; Jin and Martin, 1999).

In plants, the predominant family of MYB proteins has two amino acid repeats, and they are known as the R2R3 MYBs (Petroni *et al.*, 2002). The large R2R3 MYB family is diverse, but its only uniting feature is that most members of the family seem to be involved in plant-specific processes involving control of secondary metabolites, response to secondary metabolites unique to plants, or cellular morphogenesis unique to plants (Jin and Martin, 1999). The first MYB protein characterized in plants was the R2R3 MYB identified in the C1 gene of maize, which is involved in phenylpropanoid biosynthesis (Petroni *et al.*, 2002). Other R2R3 MYB proteins have been implicated in phenylpropanoid biosynthesis, such as the P1 and P protein in maize (Petroni *et al.*, 2002). A putative MYB transcription factor from opium poppy was previously isolated in the Nessler lab from a latex cDNA library (Nessler, unpublished results). The latex cDNAs were sequenced and the expressed sequence tags (ESTs) were analyzed for sequence homology using BLAST searches (Altschul *et al.*, 1997). The isolated MYB like protein, clone 10H7, showed highest sequence homology to a MYB-like transcription factor 6 from cotton (*Gossypium raimondii*)(61% identity 83% similarity, GenBank accession number AAN28287.1). The homologous cotton R2R3 type MYB transcription factor's function has yet to be determined (Cedroni *et al.*, 2003). Other proteins that share sequence homology with the 10H7 clone are the R2R3 type MYB factor GMYB10 from *Gerbera hybrida*, a gene that is associated with anthocyanin pigmentation, and a P-type R2R3 MYB from rice (57% identity 78% similarity, GenBank accession number BAD08950.1) (Elomaa *et al.*, 2003). The P protein was originally isolated in maize and determines pigmentation of floral organs by directly activating a flavonoid gene subset (Grotewold *et al.*, 1994). The P binding consensus is CCWACC (Grotewold *et al.*, 1994). The binding sequence CCWACC (MYBPZM) is found in both the DIOX and MLP149 promoter regions (Figure 3.2). This suggests that this may be an important binding sequence for secondary metabolism in opium poppy. Future

experiments to test if the 10H7 protein binds to the DNA of the MYBPZM *cis* element would determine the importance of the MYBPZM *cis*-regulatory region in secondary metabolism. The MYBPZM is only one of the many MYB *cis*-elements discovered among the latex-specific promoters. Isolation of additional MYB-like proteins from opium poppy would aid in the understanding of the control/regulation of secondary metabolites.

3.4.2.2 Light *cis* elements

The light class of *cis*-elements is also a large class of elements found among the three latex promoters (Table 3.2). It is logical that this group is large because light is essential for all normal plant growth and development. Light also acts as a stimulus that regulates numerous developmental and metabolic processes (Gilmartin *et al.*, 1990). The specific *cis* element GATABOX with the binding sequence of GATA is found in the upstream region of light responsive elements (Gilmartin *et al.*, 1990). The GATA sequence is found in all of the latex-specific upstream regions. Mutation of the GATA sequence in the petunia *CAB22R* promoter substantially decreased expression of the light regulated gene RBCS (ribulose-1,5, bisphosphate carboxylase) (Terzaghi and Cashmore, 1995). Deletion and mutation of the GATA box affects light induction in transient assays in homologous systems (Terzaghi and Cashmore, 1995). Since this element functions similarly in different systems, this light-inducible element may affect the regulation of latex-specific genes in opium poppy.

Another *cis* element that functions similarly between homologous systems is the GT1CONSENSUS DNA sequence GRWAAW, found in the 5' upstream region of tryptophan decarboxylase (*TDC*) which interacts with the GT-1 nuclear factor (Ouwerkerk *et al.*, 1999). The DNA consensus sequence is found in all three promoters that were analyzed. The upstream regions of *TDC* in tobacco and Madagascar periwinkle (*Catharanthus roseus*) were analyzed for the GT-1 binding site, identified sites were then mutagenized (Ouwerkerk *et al.*, 1999). Induction of the mutated *TDC* promoter under UV light was significantly lower, demonstrating a functional role of GT-1 in the induction of *TDC* expression by UV light (Ouwerkerk *et al.*, 1999). The GT-1 binding studies in other upstream regions suggest that GT-1 is a regulator of both light and defense related genes (Ouwerkerk *et al.*, 1999).

3.4.2.3 Stress *cis* elements

The largest group of shared *cis* elements among the *DIOX*, *MLP146* and *MLP149* promoters are the stress and stress/defense classes. This reinforces the function of poppy latex as a protection mechanism following wounding. A group of W-box WRKY *cis*-elements were identified in all latex promoters (Table 3.2, Figure 3.2). The W-box cis elements (WBBOXPCWRKY1, WBOXATNPR1, and WBOXHVISO1) are associated with a class transcription factors known as WRKY proteins. WRKY proteins have been found in a variety of plant species and appear to be restricted to the plant kingdom (Eulgem *et al.*, 2000). Many WRKY proteins have been shown to bind *in vitro* to W-box DNA sequences (Eulgem *et al.*, 1999). The WRKY1, -2 and-3 proteins have been implicated in the regulation of the parsley (*Petroselinum crispum*) *PR1* genes encoding PR-10 class proteins (Eulgem *et al.*, 1999). The common feature of WRKY proteins the occurrence of either two or one copies of the highly conserved WRKY domain consisting of approximately 60 amino acids (Eulgem *et al.*, 2000). This conserved domain contains a unique zinc finger motif Cys_2His_2 or $Cys_2HisCys$ in the C-terminus and the N-terminal sequence WRKYGQK which is strictly conserved (Eulgem *et al.*, 2000). The WRKY proteins also demonstrate a requirement of metal ions for binding (Eulgem *et al.*, 1999).

The biological roles for WRKY proteins show a regulatory function in pathogen infection and other stresses (Eulgem *et al.*, 2000). The position and number of W-boxes are important factors related to gene expression (Eulgem *et al.*, 1999; Yu *et al.*, 2001). *WRKY1* expression is mediated by 3 palindromically positioned W-boxes that function as rapid acting elicitor response elements, an indication of autoregulation (Eulgem *et al.*, 1999). The *NPR1* upstream region, a requirement for systemic disease resistance, also contains 3 W-boxes that when mutagenized abolish recognition of the WRKY DNA binding protein and demonstrate compromised disease resistance (Yu *et al.*, 2001). The 5' upstream regions of the three latex promoters contain W-boxes and they seem to occur in clusters (Figure 3.2). A set of pathogenesis related 1 (*PR1*) genes were found to have an average of 4.3 W-boxes in their promoters and were arranged in clusters, whereas a set of randomly selected genes contained fewer than two boxes per promoter (Dong *et al.*, 2003). This strongly suggests that the clustered W-boxes in the latex promoters are involved in signaling stress responses. In the upstream regions of the WRKY genes there

is a significant enrichment of MYB1 binding sites that may play a role in regulating the basal and constitutive level of *WRKY* genes (Dong *et al.*, 2003). The association of WRKY binding sites and the enrichment of MYB binding sites can also be observed in the latex promoters. The *WRKY* genes are thought to be a critical factor in differentially regulated expression of plant defense genes and are proposed to be controlled by autoregulation or by other members of the WRKY superfamily (Sun *et al.*, 2003). It is interesting to note that the protein associated with the *DIOX* promoter has high homology with the SRG1 protein associated with senescence (Callard *et al.*, 1996). Senescence promoters have also been found to have a W-box mediated response to oxidative stress and pathogens (Laloi *et al.*, 2004).

3.4.2.4 Vascular *cis* elements

The *DIOX* promoter has shown vascular-specific expression *in planta* (Figure 2.7). The gene expression pattern combined with the observation of a vascular-specific *cis* element identified in the 5'upstream region of *DIOX* reinforces that there is tissue-specific regulation of this gene. The element NTBFF1ARROLB with the ACTTTA target sequence is found three times in the *DIOX* promoter and once in the *MLP146* promoter. Since laticifers are limited to the vascular tissue associated with the phloem (Nessler and Mahlberg, 1976) this expression pattern and observed *cis*-elements appear to be correlated. The DOF protein transcription factor NtBBF binds the ACTTTA target sequence and shows expression specific to the vascular system (Baumann *et al.*, 1999). DOF proteins are only found within the plant kingdom and in a variety of species such as pumpkin, corn and *Arabidopsis* (Baumann *et al.*, 1999). All DOF proteins contain a 52 amino acid segment that contains a zinc finger like motif, and received its name from this unusual variant of zinc fingers, DNA-binding with one finger domain (Yanagisawa, 2004). Biological functions of the DOF transcription factors are quite varied, but they appear to play critical roles in plant growth and development (Yanagisawa, 2004). Many other DOF *cis* element sites were predicted in the analyzed latex promoters, Appendix C.

3.4.2.5 Homeodomain transcription factor *cis* element

One of the longer *cis* elements observed in *DIOX* and *MLP149* promoters is the HDZIPATATHB2 element that has the TAATMATTA regulatory sequence. The ATHB2 protein that binds the regulatory sequence is a member of the *Arabidopsis*

homeodomain leucine zipper II (HD-Zip II) protein family (Ohgishi *et al.*, 2001). ATHB-2 directly represses its own expression by autoregulation suggesting that regulation is very rapid (Ohgishi *et al.*, 2001). A putative HD-Zip transcription factor from opium poppy was identified previously by EST analysis (Nessler, unpublished results). The clone, 16C4, has high homology with the homeobox protein GLABRA2, also known as ATHB-10, which belongs to the fourth HD-Zip family (DiCristina *et al.*, 1996). Although less conserved, the leucine zipper of ATHB-10 can functionally replace ATHB-2 in an *in vitro* DNA binding assay (59% identity 74 % similarity, GenBank accession number AAG52245.1) (DiCristina *et al.*, 1996). This suggests that the 16C4 protein may interact with the identified DNA consensus sequence for a homeodomain protein. Homeobox genes are a molecular switch that establishes and maintains the cell state directly by gene autoregulation (Schena and Davis, 1992).

3.4.2.6 *Cis* elements shared between the MLP promoters

The MLP promoters *146* and *149* have the same expression pattern and both belong to the *MLP15* subfamily (Nessler, 1994). Therefore, shared regulatory elements are likely to be identified in their 5'upstream regions. Two *cis* regulatory elements of particular interest are the fruit-specific element TGTCACACMCUCUMISIN, and the defense related element SEBFCONSSTPR10A. The TGTCACA motif is involved in fruit-specific expression of the *cucumisin* gene in melon (Yamagata *et al.*, 2002). As previously mentioned, proteins homologous to MLPs have been identified in various types of fruit including melon, wild strawberry and peach (Aggelis *et al.*, 1997; Nam *et al.*, 1999; Ruperti *et al.*, 2002).

A stress/defense element that is shared between the MLP upstream regions is the SEBFCONSSTPR10A sequence YTGTCWC. The silencing element binding (SEBF) nuclear binding factor was identified in the *PR-10a* promoter (Boyle and Brisson, 2001). Database searches show that the YTGTCWC sequences are found in the promoters of numerous PR genes (Boyle and Brisson, 2001). The SEBF protein acts as a transcriptional repressor, and shares homology with RNA binding domains similar to mammalian hnRNPs (heterogeneous nuclear ribonuclear proteins) (Boyle and Brisson, 2001). MLPs have been shown to share significant structural homology with the IPR/PR10 family of proteins (Osmark *et al.*, 1998). Yet it is interesting to note that the

exact biological function of both the MLP and IPR/PR10 group has not been determined (Osmark *et al.*, 1998). The presence of the MLPs and PR10s indicates an important role in plant function. Over 4,500 latex-specific cDNAs were sequenced and the ESTs were classified into various groups (Nessler, unpublished results). The largest groups of ESTs identified were the MLPs (Figure 3.3).

Another connection can be made about fruit ripening, MLPs, and the IPR/PR10 families. People who are latex-sensitive also experience allergic reactions to various fruits and vegetables, and the phenomenon is called latex-fruit syndrome (Yagami, 2002). MLPs were first isolated from latex, MLP homologs have been found in fruits and MLPs show structural homology to the IPR/PR10 proteins. The BET V I family of IPR proteins, a major birch pollen allergen, was used to elucidate similarities between the IPR/PR10 and MLP families (Hoffmann-Sommergruber *et al.*, 1997; Osmark *et al.*, 1998). It is interesting to note that the human immune system identifies a similarity among these proteins.

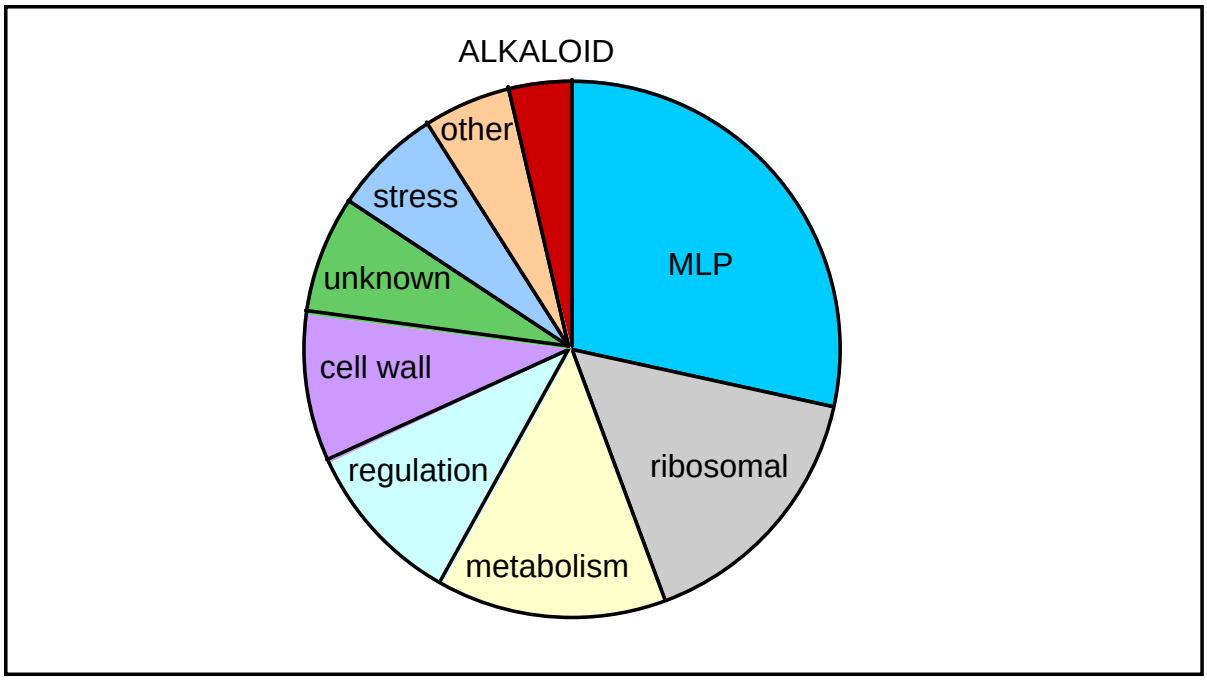

Figure 3.3 Distribution of groups of ESTs identified from a latex cDNA library (Nessler, unpublished results).

3.5 Conclusions

The higher expression of *MLP149* indicates the promoter will be potentially useful for driving expression of metabolic genes for secondary metabolite synthesis. Many *cis* elements were identified among the latex-specific promoters by bioinformatics analysis. Potential regulatory interactions between *cis* elements have been identified and should be tested. The known opium poppy transcription factors; the MYB-like R2R3 protein and the HD-Zip homeodomain transcription factor are of unique interest since potential experiments to test binding and function could be performed. Identification of more potential transcription factors from EST analysis could improve the understanding of latex/fruit ripening genes. Examples of these potential plant-specific transcription factors have been identified recently, such as the WRKY and DOF, and are currently being explored (Dong *et al.*, 2003; Yanagisawa, 2004). The function of MLPs has not been determined, but testing mutants for phenotypic variations in fruit development or pathogen response may be a place to begin to identify the elusive function of these abundant proteins.

In order to engineer secondary metabolites one must either understand the enzymatic steps in a complex pathway or understand the complex regulatory network of the genes of interest, and both of these steps may be equally time consuming (Gantet and Memelink, 2002). It is likely that a combination of these approaches may lead to productive metabolic engineering of plant alkaloids.

3.6 REFERENCES

Aggelis A, John I, Karvouni Z, Grierson D (1997) Characterization of two cDNA clones for mRNAs expressed during ripening of melon (*Cucumis melo* L.) fruits. Plant Molecular Biology **33:** 313-322

Altschul SF, Madden TL, Schaffer A, Zhang J, Zhang Z, Miller W, Lipman D (1997) Gapped BLAST and PSI-BLAST; a new generation of protein database search programs. Nucleic Acids Research **25:** 3389-3402

Baumann K, De Paolis A, Costantino P, Gualberti G (1999) The DNA binding site of the Dof protein NtBBF1 is essential for tissue-specific and auxin-regulated expression of the *rolB* oncogene in plants. Plant Cell **11:** 323-333

Boyle B, Brisson N (2001) Repression of the defense gene PR-10a by the single-stranded DNA binding protein SEBF. Plant Cell **13:** 2525-2537

Buchanan B, Gruissem W, Jones RL (2000) Biochemistry and Molecular Biology of Plants. American Society of Plant Physiologists, Rockville

Callard D, Axelos M, Mazzolini L (1996) Novel molecular markers for late phases of the growth cycle of *Arabidopsis thaliana* cell-suspension cultures are expressed during organ senescence. Plant Physiology**:** 705-715

Cannon SB, Young ND (2003) OrthoParaMap: Distinguishing orthologs from paralogs by integrating comparative genome data and gene phylogenies. BMC Bioinformatics **4**

Cedroni ML, Cronn RC, Adams KL, Wilkins TA, Wendel JF (2003) Evolution and expression of MYB genes in diploid and polyploid cotton. Plant Molecular Biology **51:** 313-325

DiCristina M, Sessa G, Dolan L, Linstead P, Baima S, Ruberti I, Morelli G (1996) The Arabidopsis Athb-10 (GLABRA2) is an HD-Zip protein required for regulation of root hair development. Plant Journal **10:** 393-402

Dong JX, Chen CH, Chen ZX (2003) Expression profiles of the Arabidopsis WRKY gene superfamily during plant defense response. Plant Molecular Biology **51:** 21-37

Elomaa P, Uimari A, Mehto M, Albert VA, Laitinen RAE, Teeri TH (2003) Activation of anthocyanin biosynthesis in *Gerbera hybrida* (*Asteraceae*) suggests conserved protein-protein and protein-promoter interactions between the anciently diverged monocots and eudicots. Plant Physiology **133:** 1831-1842

Eulgem T, Rushton PJ, Robatzek S, Somssich IE (2000) The WRKY superfamily of plant transcription factors. Trends in Plant Science **5:** 199-206

Eulgem T, Rushton PJ, Schmelzer E, Hahlbrock K, Somssich IE (1999) Early nuclear events in plant defence signalling: rapid gene activation by WRKY transcription factors. The EMBO Journal **18:** 4689-4699

Gantet P, Memelink J (2002) Transcription factors: tools to engineer the production of pharmacologically active plant metabolites. Trends in Pharmacological Sciences **23:** 563-569

Gilmartin PM, Sarokin L, Memelink J, Chua NH (1990) Molecular light switches for plant genes. Plant Cell **2:** 369-378

Grotewold E, Drummond BJ, Bowen B, Peterson T (1994) The myb-homologous P-gene controls phlobaphene pigmentation in maize floral organs by directly activating a flavonoid biosynthetic gene subset. Cell **76:** 543-553

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

Hoffmann-Sommergruber K, Vanek-Krebitz M, Radauer C, Wen J, Ferreira F, Scheiner O, Breiteneder H (1997) Genomic characterization of members of the Bet v 1 family: genes coding for allergens and pathogenesis-related proteins share intron positions. Gene **197:** 91-100

Jin HL, Martin C (1999) Multifunctionality and diversity within the plant MYB-gene family. Plant Molecular Biology **41:** 577-585

Laloi C, Mestres-Ortega D, Marco Y, Meyer Y, Reichheld JP (2004) The Arabidopsis cytosolic thioredoxin *h5* gene induction by oxidative stress and its W-box-mediated response to pathogen elicitor. Plant Physiology **134:** 1006-1016

Martin C, Paz-Ares J (1997) MYB transcription factors in plants. Trends in Genetics **13:** 67-73

Nam YW, Tichit L, Leperlier M, Cuerq B, Marty I, Lelievre JM (1999) Isolation and characterization of mRNAs differentially expressed during ripening of wild strawberry (*Fragaria vesca L.*) fruits. Plant Molecular Biology **39:** 629-636

Nessler CL (1994) Sequence analysis of two new members of the major latex protein gene family supports the triploid-hybrid origin of the opium poppy. Gene **139:** 207-209

Nessler CL, Allen RD, Galewsky S (1985) Identification and characterization of latex-specific proteins in opium poppy. Plant Physiology **79:** 499-504

Nessler CL, Burnett RJ (1992) Organization of the major latex protein gene family in opium poppy. Plant Molecular Biology **20:** 749-752

Nessler CL, Griffing LR (1989) Immunolocalization of the major latex proteins in developing laticifers of opium poppy (*Papaver somniferum).* Journal of Plant Physiology **134:** 357-363

Nessler CL, Kurz WGW, Pelcher LE (1990) Isolation and analysis of the major latex protein genes of opium poppy. Plant Molecular Biology **15:** 951-953

Nessler CL, Mahlberg PG (1976) Laticifers in stamens of *Papaver somniferum* L. Planta **129:** 83-85

Nessler CL, Vonder Harr RA (1990) Cloning and expression analysis of DNA sequences for the major latex protein of opium poppy. Planta **180:** 487-491

Ohgishi M, Oka A, Morelli G, Ruberti I, Aoyama T (2001) Negative autoregulation of the Arabidopsis homeobox gene ATHB-2. Plant Journal **25:** 389-398

Osmark P, Boyle B, Brisson N (1998) Sequential and structural homology between intracellular pathogenesis-related proteins and a group of latex proteins. Plant Molecular Biology **38:** 1243-1246

Ouwerkerk PBF, Trimborn TO, Hilliou F, Memelink J (1999) Nuclear factors GT-1 and 3AF1 interact with multiple sequences within the promoter of the *Tdc* gene from Madagascar periwinkle: GT-1 is involved in UV light-induced expression. Molecular and General Genetics **261:** 610-622

Petroni K, Tonelli C, Paz-Ares J (2002) The MYB transcription factor family: From maize to Arabidopsis. Maydica **47:** 213-232

Rombauts S, Dehais P, Van Montagu M, Rouze P (1999) PlantCARE, a plant cis-acting regulatory element database. Nucleic Acids Research **27:** 295-296

Ruperti B, Bonghi C, Ziliotto F, Pagni S, Rasori A, Varotto S, Tonutti P, Giovannoni JJ, Ramina A (2002) Characterization of a major latex protein (MLP) gene down-regulated by ethylene during peach fruitlet abscission. Plant Science **163:** 265-272

Sambrook J, Russell DW (2000) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Schena M, Davis RW (1992) HD-Zip Proteins - Members of an Arabidopsis homeodomain protein superfamily. Proceedings of the National Academy of Sciences of the United States of America **89:** 3894-3898

Stromvik MV, Sundararaman VP, Vodkin LO (1999) A novel promoter from soybean that is active in a complex developmental pattern with and without its proximal 650 base pairs. Plant Molecular Biology **41:** 217-231

Sun CX, Palmqvist S, Olsson H, Boren M, Ahlandsberg S, Jansson C (2003) A novel WRKY transcription factor, SUSIBA2, participates in sugar signaling in barley by binding to the sugar-responsive elements of the *iso1* promoter. Plant Cell **15:** 2076-2092

Terzaghi WB, Cashmore AR (1995) Light-regulated transcription. Annual Review of Plant Physiology and Plant Molecular Biology **46:** 445-474

Thompson JD, Higgins DG, Gibson TJ (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

Yagami T (2002) Features and mode of action of cross-reactive plant allergens relevant to latex-fruit syndrome. Food and Agricultural Immunology **14:** 241-253

Yamagata H, Yonesu K, Hirata A, Aizono Y (2002) TGTCACA motif is a novel cis-regulatory enhancer element involved in fruit-specific expression of the cucumisin gene. Journal of Biological Chemistry **277:** 11582-11590

Yanagisawa S (2004) Dof domain proteins: Plant-specific transcription factors associated with diverse phenomena unique to plants. Plant and Cell Physiology **45:** 386-391

Yu DQ, Chen CH, Chen ZX (2001) Evidence for an important role of WRKY DNA binding proteins in the regulation of *NPR1* gene expression. Plant Cell **13:** 1527-1539

Chapter 4

SUMMARY AND FUTURE CONCLUSIONS

4.1 SUMMARY AND FUTURE CONCLUSIONS

The laticifer is an unusual cell type found in various species of plants. Latex, whether from the opium poppy or the rubber tree (*Hevea brasiliensis),* has been an essential part in human culture for many centuries. Not all plant species contain the laticifer cell type, yet the isolated latex-specific promoter in this study demonstrated vascular-specific expression in *Arabidopsis* and tobacco, species that do not contain laticifers. This suggests that laticifers may share an evolutionary link with the plant vascular system. Important alkaloids such as morphine and codeine accumulate in laticifers. Other research has shown that alkaloid biosynthesis enzymes such as codeinone reductase (COR) are localized to sieve elements in the phloem whereas the major latex protein was localized in the cytoplasm of laticifers. Understanding the gene expression of laticifers, such as the DIOX gene, and how accumulation of the alkaloid enzyme products occurs in laticifers may provide another genetic engineering approach to acquire, isolate, and enhance alkaloid production.

Cis element analysis of laticifer-specific promoters identified potential targets for experimental analysis to test vascular-specific expression, binding of transcription factors, and environmental stress regulation. The putative MYB and HD-ZipII transcription factors could be tested with the identified cis elements in the *DIOX* and *MLP* promoters using electrophoretic mobility shift assays. Additional elements and/or conserved sequences could be further identified using motif elucidation tools such as MEME (Multiple Em for Motif Elucidation) (http://meme.sdsc.edu/meme/website.intro.html).

Designing promoter deletion constructs of the *MLP149* would further the understanding of latex-specific gene expression, and may confirm results observed with the *DIOX* promoter. The function of the MLP proteins has still not been determined, but homologs are found in multiple species, and these MLP-LPs are associated with various functions, mainly plant defense. Experiments to identify MLP function could be

designed. For example, MLP homologs that have been identified in *Arabidopsis* could be screened using available T-DNA insertion libraries such as the SALK lines.

Ultimately, the knowledge of laticifer-specific expression could aid in metabolic engineering from various points of view. Understanding how the opium poppy accumulates the desired alkaloids in laticifers, and what regulatory elements (*cis* elements and transcription factors) are necessary to effect latex-specific expression are key concepts in elucidating the metabolic network. The goal of producing higher alkaloid yielding opium poppies through use of cell-specific expression may reduce the pharmaceutical reliance on organic chemistry as a production method.

APPENDIX A

Primer Sequences

Name of primer	Sequence of primer
2L1.1	CGAAGGGTCTGAGCAACC
2L1.2	CTTTAACAATTTGAATACC
2L1.3	CTGGTTCTTTTTCTCC
2L1.4	CTCAAAGATTGATATATG
2L1.5	TTGTTGAATTATCGCCG
2L1.6	TCTCGTTCCTCAGTCGC
2L1.7	ACCCGTTCTTGAACAC
2L1.8	TGTTTATATTGCCACAG
2L1.9	CCAACGAATCTTTATGC

Table A.1 Primers used to sequence the *DIOX* 5' upstream region.

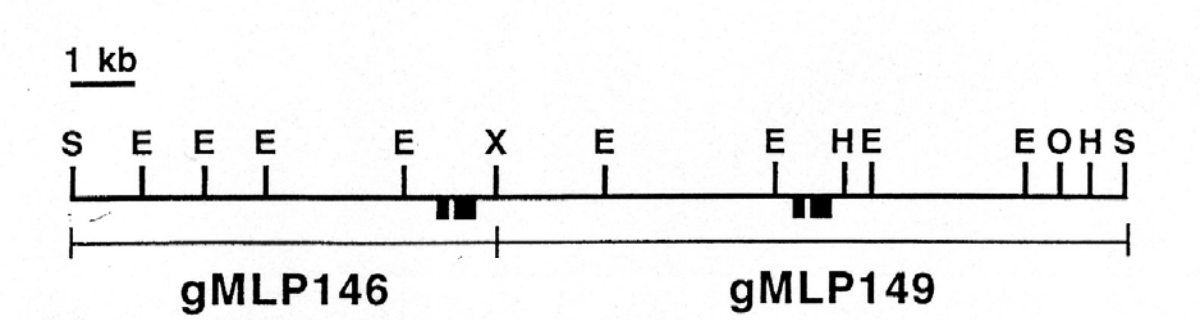

Fig. 1. Restriction map of *MLP14* which includes a 6-kb subclone (*MLP146*) and a 9-kb subclone (*MLP149*). Coding regions are shown as small solid boxes. E, *Eco*RI; H, *Hin*dIII; O, *Xho*I; S, *Sal*I.

Figure A.1 Arrangement of the MLP146 and MLP19 subclones.
(Nessler, 1994)

from
Nessler CL (1994) Sequence analysis of two new members of the major latex protein gene family supports the triploid-hybrid origin of the opium poppy. Gene 139**:** 207-209

Primer name	DNA sequence	Tm
14six966.rev	CCATGGCTAGTAGTTGATACTTATTG	58
14six653.rev	CTGAACATTTGTGACGGATA	58
Pro6NcoZ	ATGAATCGTTGACCCAATAC	58

Table A.2 Primers used to sequence the MLP146 5' upstream region from the inserted *Nco I* restriction site.

Primer name	DNA sequence	Tm
T3	AATTAACCCTCACTAAAGGG	56
six488.for	GTGTTTACTGTTGGATCAGAG	55
Pro6T3A	GCTTCCATGACATTTACACA	55
Pro6T3B	AATTCTTCTCAATCCCCTTC	55

Table A.3 Primers used to sequence the MLP146 5' upstream region from the T3 side in pBluescript genomic clone.

Primer name	DNA sequence	Tm
9T7.2R	TCACAATCACAATCGTCCTA	55
14nine1486.rev	CCATGGCTAAAAGTTGATACTTACG	60
nine532.rev	CATGTTGTGACGTGTGGTT	60

Table A.4 Primers used to sequence the MLP149 5' upstream region from the inserted *Nco I* restriction site.

Primer name	DNA sequence	Tm
T7	GTAATACGACTCACTATAGGGC	56
Nine620.for	GAAGTGTGATGACGATGATG	60
Pro9T7A	GCTTCCATGACATTTACA	55

Table A.5 Primers used to sequence the MLP149 5' upstream region from the T7 side in pBluescript genomic clone.

APPENDIX B

Nucleotide and Amino Acid Sequences

>2L1/DIOX cDNA
GGATCCGAATTCGGCACGAGAAGAAAGTTCATCAAATATAGAGTTCATGGAGACACCAA
TACTTATCAAGCTAGGCAATGGTTTGTCAATACCAAGTGTTCAGGAATTGGCTAAACTC
ACGCTTGCAGAAATTCCATCTCGATACACATGCACCGGTGAAAGCCCGTTGAATAATAT
TGGTGCGTCTGTAACAGATGATGAAACAGTTCCTGTCATCGATTTGCAAAATTTACTAT
CTCCAGAACCCGTAGTTGGAAAGTTAGAATTGGATAAGCTTCATTCTGCTTGCAAAGAA
TGGGGTTTCTTTCAGCTGGTTAACCATGGAGTCGACGCTTTACTGATGGACAATATAAA
ATCAGAAATTAAAGGTTTCTTTAACCTTCCAATGAATGAGAAAACTAAATACGGACAGC
AAGATGGAGATTTTGAAGGATTTGGACAACCCTATATTGAATCGGAGGACCAAAGACTT
GATTGGACTGAAGTGTTTAGCATGTTAAGTCTTCCTCTCCATTTAAGGAAGCCTCATTT
GTTTCCAGAACTCCCTCTGCCTTTCAGGGAGACACTGGAATCCTACCTATCAAAAATGA
AAAAACTATCAACGGTTGTCTTTGAGATGTTGGAAAAATCTCTACAATTAGTTGAGATT
AAAGGTATGACAGACTTATTTGAAGATGGGTTGCAAACAATGAGGATGAACTATTATCC
TCCTTGTCCTCGACCAGAGCTTGTACTTGGTCTTACGTCACACTCGGATTTTAGCGGTT
TGACAATTCTCCTTCAACTTAATGAAGTTGAAGGATTACAAATAAGAAAAGAAGAGAGG
TGGATTTCAATCAAACCTCTACCTGATGCGTTCATAGTGAATGTTGGAGACATTTTGGA
GATAATGACTAATGGGATTTACCGTAGCGTCGAGCACCGGGCAGTAGTAAACTCAACAA
AGGAGAGGCTCTCAATCGCGACATTTCATGACTCTAAACTAGAGTCAGAAATAGGCCCA
ATTTCGAGCTTGGTCACACCAGAGACACCTGCTTTGTTCAAAAGAGGTAGGTATGAGGA
TATTTTGAAGGAAAATCTTTCAAGGAAGCTTGATGGAAAATCATTTCTCGACTACATGA
GGATGTCGAGCTC

Figure B.1 cDNA sequence of the 2L1/ DIOX clone 1134 bp

>DIOX 5’ Upstream Region
TCTAGAAGATGAAAATTACCTTGCAATGGAAAATATCCCAACAATTGATTAACAATATC
CACAAGAGCTGTAAATCAACAGATTTTATTAAACTAGTTAAAAAACGAAAAATGAAGGG
AAGCAAACAAAGATCAAAACACTATCAATTTAAATCTAAAGAAACACTTAATATATTTT
AATGTAAACTACCAAACCCGTGAACATAAATCAGGTTACAAATAAACAATTCGTAAAAT
CGTTGAATACTCTGTGAATTATTTCGCTAACAGTCATCAAACCTTAAAGAAATTAAATA
ATATTTCCACTTAGTTCTTACCAAATCAAGTTAATCTTCTTGAAGGGGAACTGGTCAAC
ACTATAATTGAATAAATTAAATTCTCAAGCGGCAATGAAACCAAAAATTCGAAAACTTC
TTTAACAATTTGAATACCATTGATTTTCAACAAAATATGAAGTCTTTTTTTGGATCTCA
ACCCGTAAAAGGGAGAAAAAGAACCAGAGTTGTTAAAGAGAGAAAGAGGGAATGAGAAC
CGGAGAAGTTATATGTAGTAAATGGTGTATACGTGGGGTTAGACTTTGACTCCTCGGAA
TATATTAATTCTCTTTCTTTTCATTTTAGGAAGAATAAAAATATATTAATACGTTTAAT
TCTCTTTCTTTCCATTTTATGAAGAATAAAAATATATTAATACATTTTTGCCACACGTA
TAACATGAATCGTAGCAATAAACTGTCTTTTTGCCGCACACCGGACCTTGCTGCGGCAA
AAGCTAAACTTTTTGCGTCACCAGTGGTTTCCATGTGGCTAGAAATCAGATGTTTCCAC
AACATGATATAACTGTGGCAATATAAACATTTTGCAACATACCCTCTAGTCGGTGGTGT
GGCTACAAAGTAATTTTTTTGTAGTGTAAGAAATCTTTATTGATGAATTCTAAGGAAAC
AAAATACCCGTTCTTGAACACAAGGAAACCCTAATTTCACTCTCTACACTAAGCTCTCT
TCTCTCTCTCCCAAAATCCGCCCTCCATGCCTTGTCCAAGGACCTTTATTTATAGGCCA
ATCAACTCTAATGGGGTACAACTCATTTTACTTCCCAAGTTCTTTCGGGGGTGCTTTAT
ACTTCGCCAATCCATTTTCCATAAAATTTTCTCCCTTCTTTCGCCATCTTCACATGGGT
TTGTCGGGACACCTATCTCTTTTGTTGCTCCCCAATATAGTTACTTCGCGATCTATCTT
TTCAACCAAGTTACCGTACTTCTCTCTTTTGGTTTCTTCATTAGAATCTTGTTGGGTAC
TCTAGCTTGGTCTTTCGTAACTTATATTCTTCGTGTAAGTCTCCTTGTTTCTATATGAT
ACTTCGCGGGATTGATCGGTGACGAGGTAATTCTGACATTCGAATTGACAAGTCCTGAC
TGAGATCTTTAAGTGAGATTTCCTAATCATGCTTCTTCCTTTTGTGTTCGACACGTGGC
GAGTCTATTTCGTGTACCACACACATTGTAAAATGAGGGACAATCTCATTGCTTTCTTC
ATGAGCAGGTAATGCTATGATTGTAGAAATCGTTAATAATTACATACCCAATTGCGCAT
CACCATGCAATCGTAAATCATTGCCCAACGAATCTTTATGCATAAGATAATGCAGGACA
TGGGTGTGTGTGAGACCAAGAGTGGTGGCATAATACGTTCTCGAAGGACAAGTGGGCCA
AAGTCGGCGATAATTCAACAAATAGCGCAATGAAGGTAGTGGTAGTGGTGGTAGGATCT
TCAATCTGAATGCAGAGGTAGTAATATTTGAAAATAGTACTAAATTACTAATCGAGATC
TTCACTTTATTTATTTTCTATGATCAACAACCGATTCTTACTAGGTGAATTTTTTCACC
GTTCTTCTATCACTAACTGTCAGTTATCATATTTCCATCTTATCATAAATTCAATCTTA
TGTTTGTTTCTAGGAACTTTAGGTATGGAGATATGGCCTATAGAAAGTACTAAAAAGCA
GTTGCTTGTGTGGATTTGATTGAGATTTGGTGAAGAAAAGGATTAACTTTATTATTAAC
ATGAACTGAAGTAGTGGTTCTCGTTCCTCAGTCGCTACATTAGCTTGGGAAAATACCTT
AGATCGTACACCATTTGCCCACGTCCTCCTGCTCATGTAGGCATGGACCCAATCTAATA
CCTTAGATCGTACCAAGTTCTGCAACAAGAAAGATATAGAAAGTGAGAAGAACAACAAA
AATCAGTCCTAATCACCTAAACAACAACCAAACTTCACCATATTTCGTTGTTGTTAGAT
TATCTTGCAACTTAGACCGGCAGGTTAGATCGTAGACATCATCAGGTTAAAACTGGTAG
GGTAAGCAGTGGTTGCTCAGACCCTTCGTTGAAAAAACACTAGTCATTTATTCATTAGA
CATCGAGTGAAATATAGATATAAGTAAGGCATATCCACTTGAGAGAAACAAAGAGACCT
GAAAGATTGATATATGATCTGAAGATCTGACAAGAAAGTTCATCAAATATAGAGTTCCA
TGG

Figure B.2. DNA sequence of the 2540bp DIOX promoter. The ATG start codon for the initiation site Met is underlined.

>MLP146 Upstream region
TTGGGGCAACGATCATTCAACCTATCTATCCAAACCATATCTTTCAGTGTACATTACAT
ATATTGAGATTGTTACCGTGTTTCGTATGTATGAATAGTTTCATATCATAATCAACTAT
ATGCATTAAAAGTGTGGGGTTTACTGAGTACGGATGAATTACAAACATTACGATTCGAT
CGTCAATGTAACCTGTCAGTCCATCATAGACTACAAATGCGAATTCACAGAATTAGATG
CGACAGCAAAATCACCAGATTTGTTAGCAATACCAGCTATCTACTCATGATTGTCAGTT
GATGCATCTCCATTAGTCTTTTAAGACTGCAGAAGTTGCATCCCACCCAGATTACTTTC
CTGGAAGGTACTGAATAAAATAGAGAACCGCCGTCTGAAATAATCTCTCAAAACAATCA
TCTTAAAATAAGATCTTAATTGCATAGTTTAATGATGAACAAAAACTACATGTGATGCC
CTATTGGGACGGCAGAAAAGTGACAGAATAGCCTGTTTTTCTGTCACAACTGAACATTT
GTGACGGATACAGTACCGTCACCTTAGGTGCTGTCGCAAATACCTTTTGTGACGGCTAC
TGCGACGGATATACTGTCACGAAATCAATTGTGACGGAAGTATTCCTGTCACATTTGCC
TTTGACTAATTTCAGTTGACTTTACCAACAGGGACGGAATAACAACAGGTATTTCTTGG
GACAGCATGCCGTCACCTCTGGCCTGACTTAGCTGTCACCAAAACTAAAAATTTCAAGT
TGTAAAAATTACATCCATCCATCAAAAAACACACTGGAAATTCTCAGTAAGATTCTATT
TCCATTCATCCTCAATGTAGCAAAAAATGTCTTGGTTACCATACAAATTAAAAGACAAA
TGTATGTCAAAATGCCAGAATAATTCTTAGAAGAACTCACCAGTTTCACAACCAATGCC
TCAGCTCAAGAATGAGCTCTCAACCCTTTTTGCACAACCTGCACGACACTTAATGTCCG
ACTTAATGTATACCATTAGGTGATTTTAAATTGCTTATAATACGTGTGCCAACACGACA
CCGCACATCACGTTTATTCGTGTCATGTGCCCGTGGGCTGTGTTGTGCCTAGCTTTTGA
CTAGTAAAGCACAGCACGGCACAGCACGTATGAATCGTTGACCCAATACAGCACGGCGC
ATGGCACGTGAAGCACGGGATTTCGTGTGTCGGACTATGCCGGGCCGGCACGTTTTACA
CCTCTAGTTTTGTGGTGCAGTTATCTTTTATCCCTACTCGATTGAAGTGTCGAAAGTTC
AACTAGATTATTTGTTGTTTGATACCCTGAGTGTTGTCATAATACTGTTCGGCACTTGC
ACGCACCAATCATTCACCCAACGTAAGGGAATCAAACTATTTGTTAAGGTTTCTACTTT
CTATTGTCGAATGTTTCTAACTTTTTAGATTGAACGAGAACAAATTTGGGACCTATATA
GTAACACTGTGGGGTAAAATCCTATTACTACCACCCTAGTAGTCTATATCAAACACGCA
AAGGGAAGAGGATACATGTGATTAAAGGTCATATATCTAACAGTCAAAATGAAGTACAC
AGGCACGTTCATTTTTGTTAGCTCCACTCCATCGTAGACTCAATACTTTTTGTACAGTG
ACTTACGGGGAATAGACATCGGCCTTTGTTAGTATTAGAATTGTACCCCGTACGTAGAG
TTAAGTGATGGTTGTGTCTCTCTTTTTTCTTATAAACTTTTAGCCAAAAGAATGAAGAA
AAAACATCCAAGGTAACCTTTGCAACCTGAGATTTCTGTTGTCAATAGGCAATCTTTTT
CTAACTCAAGGGGCACCATTTAGGCATTTACCAACACAACCCCATAGCCAGTCAAACGT
ACTAGATAATAGATATCCCAACTAGACCTGAAATACCATGATGATGAGCCATTGCTTAT
AAATGATACTGATTGTTGCCTTAACTAAATGTTAGGATAATTTTGAATTTGTTATG

Figure B.3 DNA sequence of 2081bp MLP146 promoter. The ATG start codon for the initiation site Met is underlined.

>MLP 149 5′ Upstream region
GTCTTGCTGAATTTAGGATGAGTGTGATTGTGATGAAAATTAGTCTTAGCTGACCNAGT
TTATATAGAAACAACCAGAAACTACAGGAATGATAAAGGGTGGTGACTGGATAATGAGC
TGTATTTAGGGCAACGATCATTCAACTAGAACTATCTTGAAAGACTAAGAGATTTAGAG
ACAATAAAAGATTTTGTGAGATTGTAGTTATAATTTGACCTTTAACCACTAGTATTGTG
ATGAGTTACTAGCACAAACTTTGAATGAATTCTCCGTAAAAACGGCTTAAAGGATCAGC
TTCAAGATAGACAGATGATTGCATATACATCAAATTTATTACTGCATTTTCCCTGATTC
ATTACTTATTATAGATAACCAACCAGTAATCAGAGTGCATGATAGATTTAGTTTCAGTG
TACGTTACATATTTGAGTTTGTTATTGTATTTTTCGTCGATCGATCAACAATTAATCGC
AAATATGAAAAAATTGCACATCATGTTCGATCAACTGTATGCACTAAAGGTGTCCAATG
CATGTTGTGTACGTGTGGTTTGGGGTTTACTGCGTAAGGATGAACAAAAATTCAGTCGT
TGAAACCTATCAGTCCATGATAGACTGCAGATGTGAGGCAACAGAATAAGATCCGAGAA
ACTTTTACTGTTATAAAACATCTCGTTTTCAACAGAGAAAAAACACCAAAAAGATTTTC
AGTTCTACTCAAGAAAAATTCTGCATAGTCTTTTAAAGATAGCTTGTGCTACTAAATTA
ATTGCGCTCAGACGGACCCAAGAATTATTGTTGGTGGGGCTGGACTACTGAAGAACTCA
AGTGAAAAGGAAACTGGACAAATAAAAATAATACTCCCTCCGTCCGATTATAAAAAACG
TGCAACTAGCTCAGCTAGTGATCCCCGTTTCTCAAAGTTTCATGTGCTTCATGATCTCC
CAGGTTCGAGTCCCAATGGCGTAATATTGCAAGAATTAATATTATGTATGGAAGGTAGA
GTTAGTTTGTCTCCTTGTGACACAATCTCAGACCCCCACCCGTTTCGGCTTTTTTGGCT
AAGTTACCCATGAGACCCGCTGGTCGGCTAGCACAGCAACCTGGTTACGTTGTTGGAGC
TTAGCTTGTTAGGCTTCCAACTAGTTAATGTAATACTCTTTATCTAATAATTATTAATT
GACCTATTTATTTTCAGATTTTGTCTCACAATAGATGACTTATTTCACCAAACAAGGGG
ATATTTCTAAAATTACCCTTTTAATTGATTTTTTTTAGTATAAGAAATATATGTAATTT
GATAGGCATGTTTATATTGGTTAAGTAAGTGTTT

Figure B.4. DNA sequence of the 1332bp MLP149 promoter region. The ATG start codon for the initiation site Met is underlined.

MSDTNHSRSSQDSKDATSSTSQNSKPDFMEDEEILIAKMHSLVGDRWSLIAGRIPGRTA
EEIEKYWTSRHSSSHR

Figure B.5 Amino acid sequence of the 10H7 clone from opium poppy similarity to R2R3 myb family.

MTSSFCRGIGASSHNTWTKIVTKAGDDIRIASRKNLNDPGEPLGVIVCAVSSTWVPVAP
NVLFDFLRDDARRMEWDIMSMGGPVQSIANLAKGQDRGNTVTIQASKSKDSSTWILQDC
CTNGFESMIVYASVDIPELQSAMTGCDSSKVAILPSGFSILSDGLETRPLVISSRLEEK
TVEGAGSLFTMVFQILANSTPTAKLTMESLETVNTLVACTLQNIKKGLDCED

Figure B.6 Amino acid sequence of the 16C4 clone from opium poppy similarity to homeodomain leucine zipper transcription factor II family.

APPENDIX C

Shared *cis*-elements and their locations

Appendix C Cis element location in relationship to ATG start codon

			Total 2540bp		Total 2081bp		1332bp	
Class/Type	Cis element	recognition seq	DIOX location	DIOX site	MLP146	MLP146 site	MLP149	MLP149 site
light	-10PEHVPSBD	TATTCT	-1217	(+)	-1101	(-)	-699	(-)
			-1868	(-)	-1506	(-)		
			-1918	(-)				
seed	-300ELEMENT	TGHHAAARK	-1038	(+)	-216	(-)	-503	(+)
			-1685	(-)				
seed	AACACOREOSGLUB1	AACAAAC	-591	(-)			-902	(-)
stress	ABRELATERD1	ACGTG	-396	(-)	-406	(-)	-449	(+)
			-1071	(+)	-774	(-)	-790	(+)
			-1072	(-)	-817	(+)		
			-1837	(-)	-818	(-)		
			-1978	(+)	-857	(-)		
					-932	(-)		
stress	ACGTATERD1	ACGT	-395	(+)	-118	(+)	-223	(+)
			-395	(-)	-118	(-)	-223	(-)
			-853	(+)	-299	(+)	-449	(+)
			-853	(-)	-299	(-)	-449	(-)
			-1071	(+)	-405	(+)	-790	(+)
			-1071	(-)	-405	(-)	-790	(-)
			-1836	(+)	-625	(+)	-917	(+)
			-1836	(-)	-625	(-)	-917	(-)
			-1899	(+)	-773	(+)		
			-1899	(-)	-773	(-)		
			-1978	(+)	-817	(+)		
			-1978	(-)	-817	(-)		
					-856	(+)		
					-856	(-)		
					-931	(+)		
					-931	(-)		
					-958	(+)		
					-958	(-)		
c metabolism	AMYBOX1	TAACARA			-8	(-)	-901	(-)
					-326	(-)		
					-396	(-)		
					-606	(-)		
					-1749	(-)		
stress/defense	ASF1MOTIFCAMV	TGACG	-1163	(+)	-1284	(-)		
			-1757	(-)	-1381	(+)		
					-1422	(+)		
					-1455	(-)		
					-1470	(+)		
					-1827	(-)		
other plastid	BOXIINNTPATPB	ATAGAA	-321	(+)	-588	(-)	-1268	(+)
			-553	(+)	-1183	(-)		
			-647	(-)				
			-695	(-)				
			-1193	(-)				
basic regulation	CAATBOX1	CAAT	-56	(-)	-44	(-)	-18	(-)
			-366	(+)	-64	(-)	-69	(-)
			-515	(-)	-183	(+)	-123	(+)
			-600	(+)	-190	(+)	-154	(-)
			-768	(+)	-311	(-)	-306	(+)
			-801	(+)	-369	(+)	-362	(-)
			-927	(-)	-558	(-)	-374	(+)
			-939	(+)	-584	(-)	-538	(-)
			-955	(-)	-639	(+)	-564	(-)
			-957	(+)	-723	(-)	-805	(+)
			-986	(-)	-838	(+)	-847	(-)
			-1017	(-)	-970	(-)	-870	(+)
			-1024	(+)	-1065	(+)	-895	(-)
			-1040	(-)	-1164	(+)	-1019	(-)
			-1139	(-)	-1385	(-)	-1101	(-)
			-1172	(-)	-1387	(+)	-1134	(-)
			-1327	(+)	-1528	(-)	-1153	(+)
			-1411	(+)	-1571	(-)	-1305	(-)
			-1479	(+)	-1595	(+)		
			-1616	(-)	-1719	(-)		
			-1695	(+)	-1741	(+)		
			-1816	(+)	-1824	(+)		
			-2108	(-)	-1937	(-)		
			-2121	(+)	-1943	(-)		
			-2153	(+)	-2040	(-)		
			-2179	(-)				
			-2316	(+)				
			-2396	(+)				
			-2487	(+)				
			-2496	(-)				
			-2498	(+)				
			-2516	(+)				

mutiple secondary metabolism	CACGTGMOTIF		-1072 (+)	-818 (+)	
			-1072 (-)	-818 (-)	
seed	CANBNNAPA	CNAACAC	-1080 (-)	-143 (+)	
				-478 (+)	
				-951 (+)	
other flowering	CARGCW8GAT	CWWWWWWWWG		-60 (+)	-75 (+)
				-60 (-)	-75 (-)
					-973 (+)
					-973 (-)
other heat shock	CCAATBOX1	CCAAT	-367 (+)	-640 (+)	-18 (-)
			-958 (+)	-839 (+)	-375 (+)
			-1328 (+)	-1066 (+)	-806 (+)
			-1412 (+)	-1528 (-)	
			-1480 (+)		
seed	CICADIANELHC	CAANNNNATC	-241 (-)	-20 (-)	
			-521 (-)	-1600 (+)	
			-2218 (+)		
c metabolism	DOFCOREZM	AAAG	-27 (+)	-179 (-)	-75 (-)
			-60 (+)	-215 (-)	-173 (-)
			-71 (+)	-245 (+)	-280 (-)
			-317 (+)	-255 (-)	-413 (+)
			-327 (+)	-270 (-)	-500 (+)
			-487 (-)	-327 (-)	-589 (+)
			-497 (+)	-364 (-)	-594 (-)
			-539 (+)	-445 (+)	-633 (+)
			-549 (+)	-469 (+)	-681 (-)
			-576 (-)	-566 (-)	-814 (+)
			-706 (-)	-590 (-)	-1047 (+)
			-829 (+)	-711 (+)	-1077 (-)
			-913 (-)	-739 (-)	-1115 (-)
			-1013 (-)	-876 (+)	-1148 (+)
			-1085 (-)	-888 (-)	-1174 (+)
			-1117 (-)	-1033 (-)	-1238 (+)
			-1230 (-)	-1126 (+)	
			-1275 (-)	-1334 (-)	
			-1303 (-)	-1354 (-)	
			-1341 (-)	-1428 (-)	
			-1381 (-)	-1514 (+)	
			-1424 (-)	-1655 (-)	
			-1436 (-)	-1692 (-)	
			-1493 (-)	-1878 (+)	
			-1620 (-)	-1964 (-)	
			-1648 (+)		
			-1764 (-)		
			-1773 (+)		
			-1805 (-)		
			-1883 (-)		
			-1887 (-)		
			-1933 (-)		
			-1937 (-)		
			-1965 (-)		
			-2025 (+)		
			-2033 (+)		
			-2050 (+)		
			-2060 (+)		
			-2083 (-)		
			-2127 (-)		
			-2258 (+)		
			-2384 (+)		
			-2413 (+)		
seed/aba	DPBFCOREDCDC3	ACACNNG	-143 (+)	-413 (+)	
			-529 (-)	-799 (-)	
			-1073 (+)	-1205 (+)	
			-1578 (+)		
			-1794 (+)		
stress	DRECRTCOREAT	RCCGAC	-826 (-)		-247 (-)
			-1665 (-)		

seed	EBOXBNNAPA	CANNTG	-85 (+)	-454 (+)	-406 (+)
			-85 (-)	-454 (-)	-406 (-)
			-404 (+)	-652 (+)	-507 (+)
			-404 (-)	-652 (-)	-507 (-)
			-535 (+)	-818 (+)	-714 (+)
			-535 (-)	-818 (-)	-714 (-)
			-839 (+)	-917 (+)	-828 (+)
			-839 (-)	-917 (-)	-828 (-)
			-957 (+)	-1121 (+)	-1025 (+)
			-957 (-)	-1121 (-)	-1025 (-)
			-1072 (+)	-1341 (+)	
			-1072 (-)	-1341 (-)	
			-1368 (+)	-1361 (+)	
			-1368 (-)	-1361 (-)	
			-1726 (+)	-1387 (+)	
			-1726 (-)	-1387 (-)	
			-1741 (+)	-1476 (+)	
			-1741 (-)	-1476 (-)	
			-2498 (+)	-1484 (+)	
			-2498 (-)	-1484 (-)	
				-1541 (+)	
				-1541 (-)	
				-1714 (+)	
				-1714 (-)	
				-1794 (+)	
				-1794 (-)	
stress/defense	ELRECOREPCRP1	TTGACC	-2192 (-)	-844 (+)	-153 (+)
					-1120 (+)
other gibberellins	GAREAT	TAACAAR		-8 (-)	-206 (-)
				-326 (-)	-901 (-)
				-396 (-)	
				-606 (-)	
				-1749 (-)	
light	GATABOX	GATA	-53 (+)	-20 (+)	-33 (+)
			-89 (-)	-51 (+)	-93 (+)
			-104 (+)	-101 (-)	-170 (-)
			-238 (-)	-103 (+)	-586 (+)
			-324 (+)	-110 (+)	-723 (+)
			-563 (+)	-435 (-)	-734 (-)
			-611 (-)	-458 (+)	-937 (+)
			-627 (-)	-481 (-)	-964 (+)
			-644 (-)	-684 (+)	-1031 (+)
			-820 (+)	-735 (-)	-1181 (-)
			-901 (+)	-742 (-)	-1223 (+)
			-1185 (+)	-1405 (+)	-1241 (+)
			-1306 (-)	-1465 (+)	
			-1346 (-)	-1731 (-)	
			-1708 (+)	-1902 (-)	
			-2399 (-)	-1967 (-)	
			-2484 (-)	-1978 (-)	
			-2506 (-)	-1982 (-)	
				-2019 (+)	
				-2069 (-)	
light	GT1CONSENSUS	GRWAAW	-149 (+)	-20 (+)	-81 (-)
			-240 (-)	-110 (+)	-142 (-)
			-426 (+)	-147 (-)	-172 (-)
			-427 (+)	-178 (-)	-611 (+)
			-621 (-)	-235 (+)	-629 (-)
			-660 (-)	-268 (-)	-646 (+)
			-698 (-)	-515 (+)	-853 (+)
			-740 (+)	-737 (-)	-889 (-)
			-820 (+)	-1179 (-)	-990 (-)
			-901 (+)	-1200 (+)	-991 (-)
			-998 (+)	-1333 (-)	-1223 (+)
			-1106 (-)	-1495 (-)	-1241 (+)
			-1157 (+)		-1297 (+)
			-1393 (-)		
			-1404 (-)		
			-1405 (-)		
			-2053 (+)		
			-2104 (-)		
			-2242 (-)		
			-2434 (+)		
			-2511 (+)		
			-2512 (+)		
			-2525 (-)		
			-2529 (+)		

light	GT1CORE	GGTTAA	-194 (+)		-15 (+)
					-1113 (-)
other homologous to pectate lyase	GTGANTG10	GTTA	-114 (+)	-287 (+)	-107 (-)
			-263 (-)	-353 (+)	-126 (-)
			-285 (-)	-451 (+)	-312 (+)
			-314 (+)	-632 (-)	-429 (+)
			-504 (+)	-815 (+)	-504 (+)
			-642 (-)	-933 (-)	-709 (+)
			-656 (-)	-980 (+)	-1098 (+)
			-666 (+)	-1072 (-)	-1139 (+)
			-709 (-)	-1081 (-)	-1230 (+)
			-878 (+)	-1259 (-)	-1302 (+)
			-947 (-)	-1282 (-)	-1308 (+)
			-1111 (+)	-1364 (-)	
			-1164 (+)	-1382 (+)	
			-1369 (-)	-1396 (-)	
			-1559 (-)	-1423 (+)	
			-1755 (-)	-1453 (-)	
			-2290 (+)	-1471 (+)	
			-2343 (+)	-1487 (-)	
				-1511 (+)	
				-1538 (+)	
				-1757 (-)	
				-1783 (-)	
				-2029 (+)	
				-2044 (+)	
light	HDZIP2ATATHB2	TAATMATTA	-972 (+)		-163 (-)
					-166 (+)
light	IBOXCORE	GATAA	-239 (-)	-20 (+)	-171 (-)
			-612 (-)	-110 (+)	-964 (+)
			-628 (-)	-736 (-)	-1223 (+)
			-820 (+)	-743 (-)	-1241 (+)
			-901 (+)		
light	INRNTPSADB	YTCCANTTY	-317 (-)	-243 (-)	-626 (+)
			-472 (-)	-402 (+)	
			-602 (+)	-423 (-)	
			-710 (+)	-560 (-)	
			-770 (+)	-1050 (-)	
			-1035 (-)	-2042 (-)	
			-1456 (+)	-2047 (-)	
			-1560 (+)		
			-1930 (+)		
			-2019 (-)		
			-2432 (-)		
stress	LTRECOREATCOR15	CCGAC	-826 (-)	-794 (-)	-247 (-)
			-1357 (-)	-1002 (+)	
			-1665 (-)		
myb/aba	MYB1AT	WAACCA	-1271 (-)	-1144 (-)	-16 (-)
			-1748 (-)	-1973 (+)	-228 (-)
			-2148 (+)	-2062 (+)	-785 (-)
					-962 (+)
					-1112 (+)
myb/stress	MYB2AT	TAACTG	-631 (-)	-746 (-)	
			-638 (+)		
			-1704 (+)		
myb/aba	MYB2CONSENSUS	YAACKG	-535 (-)	-746 (-)	-828 (+)
			-631 (-)	-1341 (-)	
			-638 (+)	-1484 (+)	
			-1704 (+)	-1714 (-)	
myb/secondary metabolism	MYBCORE	CNGTTR	-535 (+)	-197 (+)	-653 (-)
			-631 (+)	-431 (-)	-675 (+)
			-638 (-)	-746 (+)	-703 (-)
			-683 (-)	-1311 (-)	-828 (-)
			-1704 (-)	-1328 (-)	
			-2276 (-)	-1341 (+)	
			-2464 (-)	-1484 (-)	
				-1714 (+)	
myb/c metabolism	MYBGAHV	TAACAAA		-8 (-)	-901 (-)
				-326 (-)	
				-396 (-)	
				-606 (-)	
				-1749 (-)	

myb/secondary metabolism	MYBPLANT	MACCWAMC	-272 (+)		-106 (+)
					-961 (+)
myb/secondary metabolism	MYBPZM	CCWACC	-184 (-)		-959 (+)
			-779 (-)		
myb	MYBST1	GGATA	-89 (-)	-21 (+)	-94 (+)
			-2484 (-)	-101 (-)	-1224 (+)
			-2506 (-)	-459 (+)	
				-735 (-)	
				-1406 (+)	
				-1466 (+)	
				-1978 (-)	
stress	MYCATERED1	CAGTGT	-1368 (-)	-454 (+)	-406 (+)
			-1741 (+)	-917 (+)	
stress/aba	MYCATRD22	CACTG	-1368 (+)	-454 (-)	-406 (-)
			-1741 (-)	-917 (-)	
				-1541 (-)	
stress/aba	MYCCONSENSUS	CANNTG	-85 (+)	-454 (+)	-406 (+)
			-85 (-)	-454 (-)	-406 (-)
			-404 (+)	-652 (+)	-507 (+)
			-404 (-)	-652 (-)	-507 (-)
			-535 (+)	-818 (+)	-714 (+)
			-535 (-)	-818 (-)	-714 (-)
			-839 (+)	-917 (+)	-828 (+)
			-839 (-)	-917 (-)	-828 (-)
			-957 (+)	-1121 (+)	-1025 (+)
			-957 (-)	-1121 (-)	-1025 (-)
			-1072 (+)	-1341 (+)	
			-1072 (-)	-1341 (-)	
			-1368 (+)	-1361 (+)	
			-1368 (-)	-1361 (-)	
			-1726 (+)	-1387 (+)	
			-1726 (-)	-1387 (-)	
			-1741 (+)	-1476 (+)	
			-1741 (-)	-1476 (-)	
			-2498 (+)	-1484 (+)	
			-2498 (-)	-1484 (-)	
				-1541 (+)	
				-1541 (-)	
				-1714 (+)	
				-1714 (-)	
				-1794 (+)	
				-1794 (-)	
vascular	NTBBF1ARROLB	ACTTTA	-488 (+)	-1335 (+)	
			-577 (+)		
			-707 (+)		
secondary metabolism	PALBOXAPC	CCGTCC		-1322 (-)	-466 (+)
				-1524 (-)	
basic regulation	POLASIG1	AATAA	-133 (-)	-928 (-)	-145 (-)
			-486 (-)	-1637 (+)	-485 (+)
			-701 (-)		-1002 (-)
			-705 (-)		-1152 (+)
			-1492 (-)		
			-1619 (-)		
			-1815 (+)		
			-1866 (+)		
			-2175 (+)		
			-2322 (+)		
			-2456 (-)		
basic regulation	POLASIG2	AATTAAA	-1896 (-)	-1131 (+)	-73 (-)
			-2171 (+)		
			-2253 (+)		
basic regulation	POLASIG3	AATAAT	-483 (-)	-698 (-)	-161 (-)
			-971 (+)	-1099 (+)	-165 (+)
			-2248 (+)	-1610 (+)	-479 (+)
			-2286 (-)		-541 (-)

other pollen	POLLENLELAT52	AGAAA	-29 (+)	-176 (-)	-50 (+)
			-77 (+)	-200 (-)	-89 (-)
			-319 (+)	-236 (+)	-419 (-)
			-329 (+)	-266 (-)	-612 (+)
			-500 (+)	-573 (-)	-647 (+)
			-551 (+)	-589 (-)	-686 (+)
			-586 (-)	-596 (-)	-1256 (+)
			-696 (-)	-1302 (-)	-1266 (+)
			-981 (+)	-1493 (-)	
			-1012 (-)	-1517 (+)	
			-1194 (-)		
			-1268 (-)		
			-1391 (-)		
			-1626 (+)		
			-1732 (+)		
			-1886 (-)		
			-1936 (-)		
			-2027 (+)		
			-2054 (+)		
			-2256 (+)		
			-2382 (+)		
c metabolism	PYRIMIDINEBOXOSAMY1A	CCTTTT	-498 (-)	-1034 (+)	-76 (+)
			-1086 (+)	-1429 (+)	-501 (-)
			-2061 (-)		
other pollen	QELEMENTZMZM13	AGGTCA		-443 (+)	-152 (-)
					-1119 (-)
AP2 like	RAV1AAT	CAACA	-249 (-)	-42 (-)	-219 (-)
			-276 (+)	-142 (+)	-536 (-)
			-304 (+)	-196 (-)	-653 (+)
			-334 (+)	-673 (-)	-703 (+)
			-686 (+)	-692 (-)	-798 (-)
			-813 (+)	-900 (-)	-873 (+)
			-1251 (-)	-950 (+)	
			-1337 (-)	-1311 (+)	
			-1679 (+)	-1328 (+)	
			-1714 (+)		
			-2099 (+)		
			-2189 (+)		
			-2464 (+)		
			-2501 (+)		
light	REALPHALGLHCB21	AACCAA	-272 (+)	-1068 (+)	-17 (-)
			-1272 (-)	-1145 (-)	-961 (+)
			-1297 (+)		
			-2147 (+)		
other root	ROOTMOTIFATPOX1	ATATT	-16 (-)	-1945 (+)	-20 (+)
			-110 (-)		-47 (-)
			-258 (+)		-92 (+)
			-623 (+)		-349 (+)
			-746 (+)		-350 (-)
			-747 (-)		-364 (+)
			-1218 (+)		-365 (-)
			-1326 (-)		-858 (-)
			-1694 (-)		-910 (+)
			-1857 (+)		
			-1860 (-)		
			-1907 (+)		
			-1910 (-)		
			-1948 (+)		
			-1951 (-)		
			-2093 (-)		
			-2244 (+)		
			-2245 (-)		
			-2369 (+)		
			-2372 (-)		
			-2486 (-)		
			-2508 (-)		
seed/aba	RYREPEATBNNAPA	CATGCA	-943 (+)		-802 (-)
					-942 (-)
other plastid	S1FBOXSORPS1L21	ATGGTA	-2112 (-)	-81 (-)	
				-989 (-)	
				-1140 (-)	
stress/defense	SEBFCONSSTPR10A	YTGTCWC		-1262 (+)	-131 (+)
				-1367 (+)	-322 (+)
				-1399 (+)	-1157 (-)
				-1490 (+)	
				-1511 (-)	

seed	SEF4MOTIFGM7S	RTTTTTR	-301 (-)	-399 (+)	-483 (-)
			-1847 (+)	-1233 (-)	-756 (-)
			-1864 (-)	-1249 (-)	-1059 (-)
			-2144 (-)	-1550 (-)	
c metabolism	TAAAGSTKST1	TAAAG	-487 (-)	-446 (+)	-173 (-)
			-576 (-)	-877 (+)	-590 (+)
			-706 (-)	-1334 (-)	-815 (+)
			-913 (-)		-1048 (+)
			-1117 (-)		-1115 (-)
			-1424 (-)		-1239 (+)
			-1493 (-)		
			-1620 (-)		
			-2034 (+)		
			-2127 (-)		
			-2259 (+)		
			-2385 (+)		
basic regulation	TATABOX2	TATAAAT	-1489 (-)	-58 (+)	
basic regulation	TATABOX3	TATTAAT	-1855 (-)		-159 (+)
			-1856 (+)		-353 (-)
			-1905 (-)		
			-1906 (+)		
			-1947 (+)		
basic regulation	TATABOX5	TTATTT	-700 (+)	-697 (+)	-112 (+)
			-704 (+)	-1584 (-)	-144 (+)
			-1491 (+)	-1611 (-)	-480 (-)
			-2249 (-)		-486 (-)
			-2285 (+)		
			-2323 (-)		
c metabolism	TATCCAOSAMY	TATCCA	-89 (+)	-1978 (+)	-1225 (-)
			-2484 (+)		
light	TBOXATGAPB	ACTTTG	-830 (-)		-414 (-)
			-1649 (-)		-1078 (+)
			-1966 (+)		
other- fruit	TGTCACACMCUCUMISIN	TGTCACA		-1366 (+)	-313 (-)
				-1489 (+)	
other	TRANSITINMONOCOTS	RMNAUGGC	-1376 (-)	-67 (-)	
stress/defense	WBBOXPCWRKY1 stress	TTTGACT	-1964 (+)	-124 (-)	
				-427 (-)	
				-886 (+)	
				-1353 (+)	
stress/defense	WBOXATNPR1	TTGAC	-1138 (+)	-123 (-)	-153 (+)
			-1963 (+)	-192 (-)	-1120 (+)
			-2191 (-)	-426 (-)	
				-844 (+)	
				-885 (+)	
				-1112 (-)	
				-1338 (+)	
				-1352 (+)	
				-1826 (-)	
stress/defense	WBOXHVISO1	TGACT	-138 (-)	-124 (-)	-116 (+)
			-1127 (+)	-352 (+)	-1229 (+)
			-1962 (+)	-427 (-)	
			-2272 (-)	-884 (+)	
				-1270 (+)	
				-1337 (+)	
				-1351 (+)	
light	ZDNAFORMINGATCAB1	ATACGTGT	-1838 (-)	-960 (+)	

Vita

Michelle Raymond was born in Dallas, TX on June 27, 1979 and was raised in Wilmington, DE by George and Jean Raymond. She is the older of two children. She attended Concord High School in Wilmington, DE and graduated in 1997. After high school Michelle was accepted to the University of Delaware in the College of Agriculture and Life Sciences. After a summer internship at the AgrEvo research station in Fresno, CA she changed her major from horticulture to plant biology. Michelle continued to intern and started research on *Bradyrhizobium* in the Plant and Soil Sciences Department under advisement of Dr. Jeffry Fuhrmann. For a senior research project Michelle began research with Dr. Janine Sherrier in the area of proteomic analysis of plant-microbe interactions. Michelle earned her Bachelors degree with a Degree of Distinction from University of Delaware in spring 2001. In the fall of 2001, Michelle entered the Masters program at Virginia Tech in the Department of Plant Pathology, Physiology, and Weed Science. She began her Master's research under the advisement of Dr. Craig Nessler in January of 2002. She completed her Masters degree in the fall 2004. Michelle is planning to move to Pittsburgh, PA and pursue a career in the biomedical sciences.