A genomic study of morphinan biosynthesis in the "opium poppy"

Papaver somniferum L.

Submitted by

Anthony G. Millgate

B.Sc. (Hons)

A thesis submitted in total fulfilment

of the requirements for the degree of Doctor of Philosophy

School of Biochemistry and Molecular Biology
Faculty of Science
The Australian National University
Acton, ACT 2601

Australia

December 2004

Table of Contents

Statement of Authorship

Except where reference is made in the text of this thesis, this thesis contains no material published elsewhere or extracted in whole or in part from a thesis by which I have qualified for or been awarded another degree or diploma.

No other persons work has been used without due acknowledgment in the main text of the thesis.

This thesis has not been submitted for the award of any degree or diploma in any other tertiary institution.

Technical assistance was received in the course of this research and the nature of this assistance is stated in the following acknowledgements section.

Anthony G. Millgate

December 2004

Acknowledgements

I would like to thank my supervisors, Dr Philip Larkin and Dr Barry Pogson for their valuable advice, enthusiasm and invaluable support over the term of this project. In particular I would like to especially like to thank Dr Philip Larkin for his immense support and faith in me over the years.

Thanks to Dr. Iain Wilson for his support, teaching and guidance over the course of my thesis that without most of this thesis would not be possible.

I am also indebted to Plant Industry, Commonwealth Scientific and Industrial Research Organisation (CSIRO) for permitting me to undertake these studies in their laboratories and would like to thank Dr. Jim Peacock and Dr. Jeremy Burdon. In particular I would like to thank all members of the Tissue Culture Laboratory at Plant Industry CSIRO for their help and support.

Gavin Kennedy, your help was invaluable with the statistical analysis of the microarray data.

I wish to thank Tasmanian Alkaloids Pty. Ltd. and Horticulture Australia Ltd for funding this research. In particular I would like to thank Dr. Tony Fist and Dr. Wayne Gerlach for their continued faith, humour and support.

I would like to thank my parents to whom I owe so much, my Nana, my sister and niece for their continued support and encouragement during the course of this research. Finally I would like to thank Sally for being there for me over this time and for her support and guidance.

Summary

Papaver somniferum (opium poppy) plants have been known to mankind for thousands of years, and used for food, oil, relief of pain, and recreation; it has also been misused with much social harm. The major active ingredient for pain relief is the alkaloid, morphine. The morphinan alkaloids are nitrogenous compounds that are complex in structure and synthesis. The complex reactions that are involved in the plant are difficult to replicate *in vitro*. Advancing our understanding of the process involved in the *in vivo* synthesis of morphine can eventually lead to an insight of where these complex molecules are made and give an insight into the types of enzymes required. Ultimately this work can lead to the future understanding as to how and why plants invest so much energy into these molecules.

This research investigates poppy germplasm (natural genotypes, mutants and transgenics) that have varying profiles of benzylisoquinoline alkaloids. First it was necessary to characterise several new *Papaver somniferum* morphinan alkaloid mutants by end product analysis. Second, in order to enable a global analysis of the transcriptome of these various genotypes, I have developed a cDNA *Papaver somniferum* library and produced a low-redundancy 17 000 cDNA microarray.

By looking at the molecular phenotype of a number of different morphinan alkaloid accumulating germplasm and comparing these to the wildtype or control morphinan alkaloid accumulation any gene changes can be attributed either directly or indirectly to morphine biosynthesis.

Microarray analysis of the transcriptomes of the characterised germplasm, with particular emphasis on the thebaine and oripavine accumulating mutant *top1*, has shown a set of ten genes to be consistently differentially expressed in several of the alkaloid variants compared to wildtype poppy. Further analysis has revealed that this set of ten genes can be reduced to just three different genes, which are

associated with morphinan alkaloid accumulation in the opium poppy. Using database searches these genes have homology closest to: cDNA 116D7, light-harvesting chlorophyll a/b binding protein Lhca2.1; cDNA 221C9, an aminotransferase cDNA which is differentially spliced in the mutant; and cDNA 151H8, which has no homology in the database. This work has opened up new insights and new avenues for further research into the levels of control and regulation on the morphinan pathway in the unique species *P. somniferum*.

List of Figures

List of Tables

Abbreviations

°C-	Celsius temperature
µL-	microLitre
µM-	micromolar
4OMT-	3'-hydroxy-*N*-methylcoclaurine 4'-*O*-methyl transferase
6OMT-	*S*-adenosyl-L-methionine:norcoclaurine 6'-*O*-methyltransferase
AFLP-	amplified fragment length polymorphisms
A.thaliana-	*Arabidopsis thaliana*
ATP-	adenine triphosphate
BBE-	berberine bridge enzyme
cDNA-	complementary DNA
C.japonica-	*Coptis japonica*
CNMT-	Coclaurine N methyl transferase
C.roseus-	*Catharanthus roseus*
cv-	cultivar
DD-PCR-	differential display PCR
DNA-	Deoxy ribonucleic acid
E.californica-	*Eschscholzia californica*
EDTA-	ethylene diamine tetraacetate
EMS-	ethyl methyl sulfate
EtOH-	ethanol
g-	gravity
HPLC-	high performance liquid chromatography
MM-	millimolar
MOPs-	3-(*N*-morphilino) propanesulphonic acid
NADPH-	nicotinamide adenosine di phosphate
NS-	norcoclaurine synthase
OMTs-	*O* methyl transferases
PCR-	Polymerase chain reaction
P.somniferum-	*Papaver somniferum*
qrt-PCR-	quantitative real time PCR
RNA-	ribo nucleic acid
rRNA-	ribosomal RNA
rtPCR-	reverse transcriptase PCR
SAM-	*S*- adenosyl methionine transferase
SAGE-	serial analysis of gene expression
SAT-	salutaridinol acetyltransferase
SDS-	sodium dodecyl sulphate
TLC-	thin layer chromatography
tRMA	tools for R Microarray Analysis (Version 1.7)
T.tuberosum-	*Thalictrum tuberosum*
TYDC-	tyrosine decarboxylase
UV-	Ultra violet

Chapter 1

Literature Review

1.1 Introduction

The study of secondary metabolites in plants has improved in recent times due to the increasing development and availability of sensitive methods of analysis, permitting the discovery of products that are generally produced in small quantities, and were previously difficult to study. Secondary metabolites were initially considered to be debris or plant metabolic waste products until the 1960s when it became clear they had functional characteristics variously contributing to plant survival (Wink, 1998). Many secondary metabolites represent a significant investment of energy and resources by the plant and often include the valuable plant resource nitrogen. Nitrogen is essential for primary metabolism and is often of limited availability to plants, suggesting that the plant must gain some selective advantage in continuing to produce these supposedly 'redundant' molecules.

Secondary metabolites include molecules responsible for flower colour, scent and nectar to ensure pollination and continuance of the species. They also contribute to plant protection from damaging UV light; some of the genes responsible for the synthesis of phenolics and terpenoids are induced by UV light (Wingender *et al.*, 1989; reviewed in Ensminger, 1993). The main purpose of plant secondary metabolites appears to be plant protection, since many are found to accumulate in response to treatment with pathogens. A well-studied example is the induction of phenylpropanoids in soybean (reviewed in Hahlbrock and Grisebach, 1979). It is now becoming evident that secondary metabolites function in a variety of plant functions beyond plant protection. Some secondary metabolites interact with symbionts to enhance plant growth; the flavone luteolin is expressed and exuded from the roots of alfalfa to activate rhizobial nodulation genes thus playing an important role in root colonisation (Peters *et al.*, 1986; Hartwig *et al.*, 1991).

Another secondary metabolite, jasmonic acid, is actually part of a signal transduction chain that ultimately ends in the activation of defence genes and plant senescence (Gundlach *et al.*, 1992; Reinbothe *et al.*, 1994; reviewed in Kutchan 2001).

The most diverse group of secondary metabolites are the alkaloids with as many as 12,000 alkaloids having been documented in plants (Wink, 2004). Some alkaloids are extremely toxic to animals; ouabain is used in arrow poisons (Neuwinger, 1998) and coniine is the active ingredient of hemlock (Wink, 1998). Many alkaloids are the active ingredients of folk medicines and stimulants in the human diet: caffeine in tea from the shrub *Camellia sinensis*, and in coffee from *Coffea arabica*; theobromine, the active ingredient of chocolate from the cacao plant *Theobroma cacao*; the antibacterial sanguinarine from *Sanguinaria canadensis*; atropine from *Atropa belladonna* was used as a medieval anaesthetic and is still used today in ophthalmology and as an antispasmodic; hemlock (*Conium maculatum*) which contains the piperidine alkaloid coniine used for murder; the powerfully addictive cocaine which comes from the coca plant (*Erythroxylum coca*); mescalin which has activities similar to LSD is produced in the peyote cactus, *Lophphora williamsii* (Wink, 1998); nicotine from *Nicotiana tabacum* is useful as an insecticide as well as the main active stimulant in tobacco; and the medicinally important morphinan alkaloids from *Papaver somniferum*, such as morphine and codeine, used as painkillers and to suppress coughing. The medicinal properties of secondary metabolites may be the indirect result of their role as defence compounds; the same effects they have on a predator's fitness through binding to various receptors or interfering with biochemical pathways, allows them to be recruited for human use (Wink, 1998).

The benzylisoquinoline alkaloids contain some of the most pharmacologically active compounds known to man, including morphine. This class of alkaloids can be found primarily in plants that belong to five families: the *Berberidaceae*, *Fumariaceae*, *Menispermaceae*, *Ranunculaceae* and the morphinan containing *Papaveraceae* (Facchini *et al.*, 1997). Many of these pharmacologically active compounds are still harvested from plants due to the complexity of their chemical structures. Approximately 25% of all available modern drugs that play a major

role in the treatment of human disease are derived from or were first discovered in higher plants (Kutchan, 1995: Farnsworth and Bingel, 1997: reviewed De Smet, 1997).

Due to their medicinal importance and the difficulty of *in vivo* synthesis, the elucidation of plant alkaloid formation has been of great interest, however progress has been limited due to the low concentrations of these molecules. More recently progress has improved due to a number of advancements in the fields of radiochemistry, *in vitro* plant cell growth and in molecular biological techniques.

With the introduction of radiolabelled precursors in the 1950s, chemical elucidation of alkaloid biosynthetic pathways became possible. Plant cell cultures provided the next step in providing a renewable and abundant source of secondary metabolic enzymes. The introduction of molecular techniques, in particular the polymerase chain reaction (PCR), have now allowed for many of these genes to be cloned. In the last few years it has been estimated that at least 30 genes (Hashimoto and Yamada, 2003) and up to 80 genes (Kutchan, 1995) have been cloned which are involved in the biosynthesis of various classes of alkaloids. Of the 20,000 to 60,000 genes thought to occur in various plant genomes, 15-25% are estimated to encode enzymes for secondary metabolism (Bevan *et al.*, 1998: Somerville and Somerville, 1999). The latest wave of mass screening technologies includes proteomics, metabolomics (Fiehn and Weckwerth, 2003) and DNA microarrays for transcript analysis, sometimes termed transcriptomics (Schena *et al.*, 1995). With these technologies researchers are well placed to expand the understanding of genes of secondary metabolism and their coordination.

Morphinan alkaloids are one such class of medically important alkaloids that are still widely used today and the synthesis of these alkaloids, while possible (Gates and Tsundi, 1952; Ginsburg and Elad 1954;), is not commercially viable (Rice, 1985). *Papaver somniferum* L. and *P. setigerum* are the only plant species known to accumulate useful quantities of morphine (Meijerink *et al.*, 1999). The aim of this project is to further the characterisation of several *P. somniferum* mutants

that are altered in their ability to synthesize alkaloids. It is hoped this study will further enhance our understanding of alkaloid biosynthesis and in particular that of the morphinan alkaloids. Specifically one mutant, *top1*, will be characterised using classical techniques such as alkaloid profiles and feeding studies together, with molecular techniques using transcriptomics.

1.2 Alkaloids

The term alkaloid is derived from the Arabic word "*al-qali*" meaning *from the ashes* and referring to the ashes of the saltwort plant from which soda was first obtained (Kutchan, 1995). Alkaloids often exhibit physiological activities (and have been exploited by man for their medicinal and toxic properties (reviewed in Roberts and Wink 1998). These molecules have been predominantly found in about 20% of Angiospermae plant species (De Luca and Laflamme, 2001; Facchini, 2001). Over 12,000 alkaloids have been isolated from plants for which many structures have been ascribed function in plant defence against herbivores, pathogens and insects (Carporale, 1995; Wink 1999; De Luca and St Pierre, 2000; Samanani *et al.*, 2002).

The pharmacological activities of these metabolites that make them useful as pharmaceuticals often provides an insight into the biological role they may perform in the plant (Carporale, 1995). However many alkaloids still have no real assigned function in the plant; and the fact that they generally give a bitter taste to the plant is often simplistically used to attribute to them the plant function of a feeding deterrent. However bitterness to human taste is not really very informative and studies have shown that some animals will still consume the plants regardless of the alkaloids (Wink *et al.*, 1993). The most widely used plant-derived alkaloids include: stimulants such as caffeine and nicotine; anticancer chemotherapeutics including vincristine, vinblastine, camptothecin derivatives and paclitaxel; and of course the strong analgesics, morphine and codeine (See Figure 1.1 for some examples of important alkaloids).

The diverse range of structures and activities of alkaloids has made it difficult to classify them easily. Alkaloids are generally defined as low molecular weight, cyclic and structurally diverse secondary metabolites that contain nitrogen (De Luca and St Pierre, 2000). Most alkaloids are derived from amino acids such as lysine, histidine, phenylalanine, tyrosine, tryptophan, ornithine, and arginine, although some are derived from other substrates such as anthranilic acid, nicotinic acid (a non-protein amino acid), and purines (Southton and Buckingham, 1989; Dewick, 1998). Most classification systems of alkaloids are based on their biochemical origins (Southton and Buckingham, 1989; Dewick, 1998; Roberts and Wink, 1998) as the nitrogen atom and carbon skeleton originating from the amino acid are often retained in the final alkaloid. A large group of alkaloids, often termed *pseudoalkaloids*, acquire their nitrogen atom via transamination reactions and the rest of the molecule may be from acetate, shikimate, terpenoid or steroid molecules (Dewick, 1998).

Some of the more well known alkaloids are the indole alkaloids derived from tryptophan and these include the antimalarial quinine from *Cinchona ledgeriana* and the anti-cancer compound camptothecin, from *Camptotheca accuminata*, This class of alkaloid also includes strychnine, ajmalicine, vindoline, tabersonine, catharanthine, and the antileukaemia compounds vincristine and vinblastine. The indole alkaloids number in excess of 1800 different compounds (Kutchan, 1995). The tropane alkaloids, another large group of alkaloids, include the anticholinergic drugs atropine, hyoscyamine and scopolamine. The strong narcotic cocaine is also classified in this group. Nicotine, a pyridine alkaloid, is not classified in the tropane class of alkaloids but has a common intermediate found in both the tropane and nicotinic alkaloid pathways. Tropane alkaloids are derived from ornithine or arginine. Lysine is the precursor of the quinolizidine alkaloids said to constitute about 2 % of all the known plant alkaloids (Wink, 1987). This group contains the antiarrhythmic sparteine and plant quinolizidine alkaloid biosynthesis occurs in the stroma of chloroplasts, the site of lysine synthesis (Wink *et al.*, 1980; Wink and Hartman, 1982). The purine alkaloids

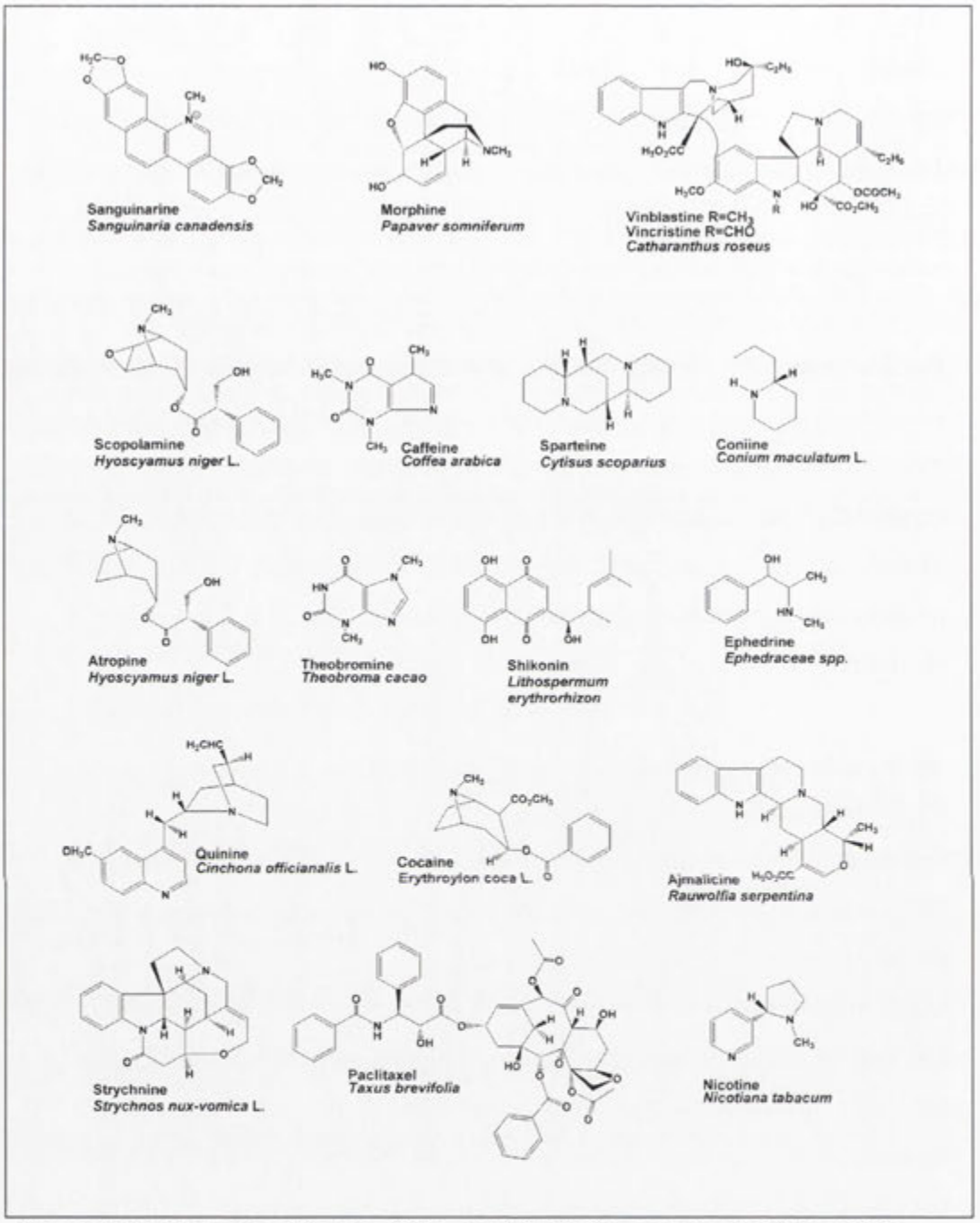


Figure 1.1: Examples of some of the medicinally important alkaloids and the plant species from which they are predominantly found. (Adapted from Kutchan, 1995; Schmeller and Wink, 1998; Verpoorte *et al.*, 2000; and Facchini *et al.*, 2004).

include: caffeine; the active stimulant of chocolate, theobromine; and the bronchodilator, theophylline.

Tyrosine is the precursor for a very large family of alkaloids consisting of the phenylethylamines and the benzylisoquinolines. The phenylethylamines include: noradrenaline; adrenaline; and mescaline an alkaloid with psychoactive and hallucinogenic properties. The benzylisoquinolines include such compounds as: papaverine, a vasodilator; noscapine, a cough suppressant; codeine, a cough suppressant and analgesic; berberine, a potent antimicrobial; sanguinarine, an antiplaque agent; and morphine, used in the treatment of severe pain and diarrhoea. The benzylisoquinoline morphinan alkaloids (Figure 1.2) constitute some of the most medically important of all the alkaloids. Man has used morphinan alkaloids in the form of opium, which contains over 40 alkaloids, for thousands of years both as a source of medicinal relief of many ailments and for recreational use. Unfortunately the recreational use of opium and derivatives also results in addiction and associated social problems.

1.3 The History of Opium

Mankind has cultivated opium poppy, *Papaver somniferum* L., for thousands of years for use as a source of food, medicinal and recreational use. Remains of cave-dwellers living in Neolithic times (4-5 thousand years BC) in the territories of what is now Spain, France, Germany and Hungary, contained evidence for the use and knowledge of poppies (Tétényi, 1997). Clay tablets of the Sumerians refer to the poppy as 'hul gil' or plant of joy and its collection as far back as 3 thousand years BC (Brownstein, 1993). There is evidence in the Egyptian Ebers papyrus (1500 BC) of its use; flowers and pieces of poppies were found in tombs from the XVIIIth Dynasty. There are also many instances through classic literature of the mention of poppy and its use even in Homer's *Odyssey* and *Iliad*. The word *opium* is probably of Greek origin, *opos* juice, *opion* poppy juice (Schiff, 2002). Hippocrates (460- 377BC) knew of the use of poppy as a hypnotic, narcotic, styptic and cathartic agent as well as the nutritive value of the seeds. It is thought that the poppy was introduced to India when the Persians under Alexander the Great (4^{th} Century BC) invaded and used the plant for the needs of their army. Herakleides (340 BC) reported that women on the Greek Islands used poppy extracts for euthanasia. The Arabs formerly called the opium

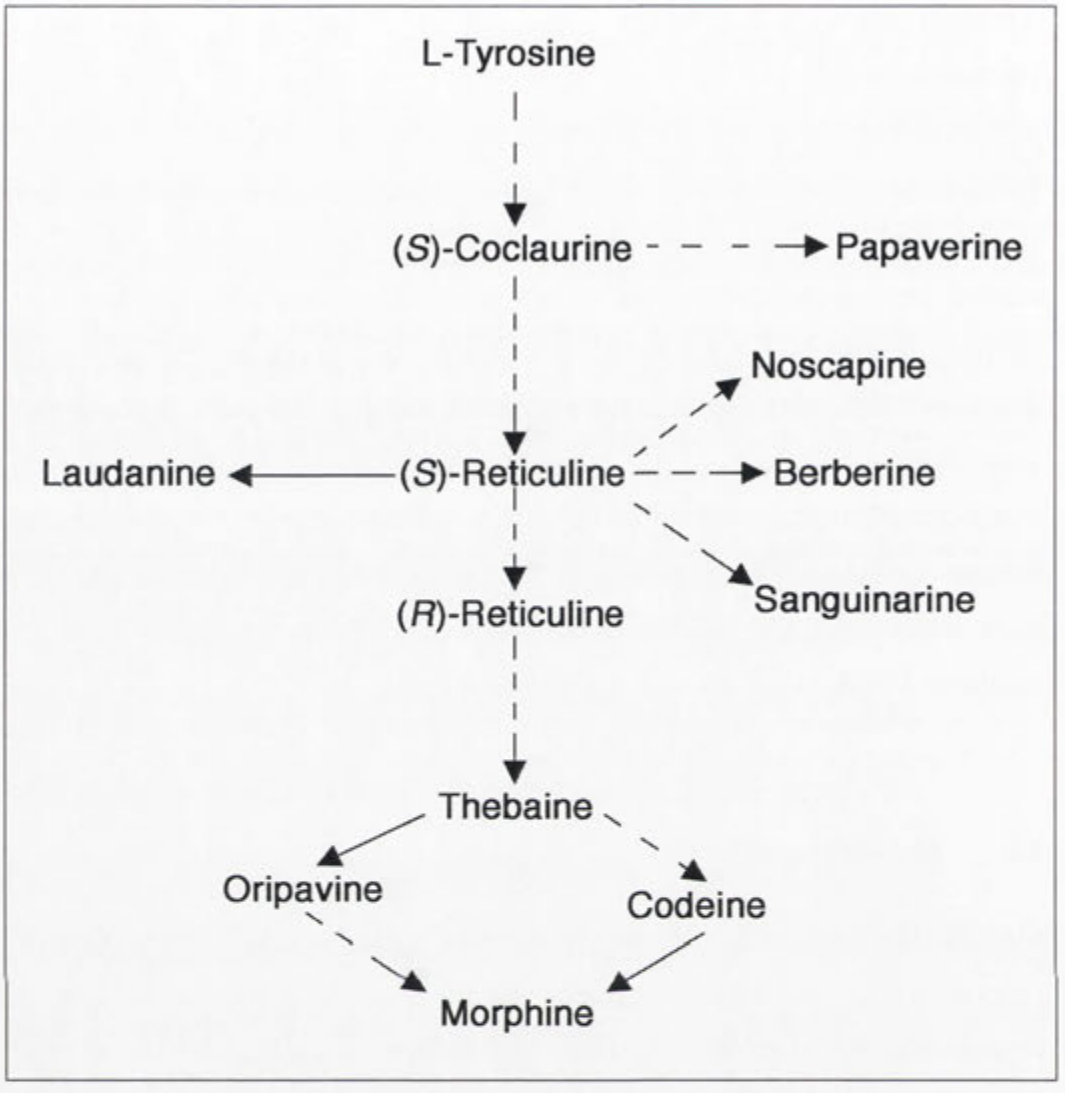


Figure 1.2: A simplified pathway illustrating the major branchpoints and compounds of the benzylisoquinoline morphinan pathway. (Adapted from Dewick, 1998; Facchini *et al*., 2000; and Ounaroon *et al*., 2003).

poppy "Abou-el-noum", father of sleep and it was widespread throughout the Arab Empire (Schiff, 2002).

The spread of opium poppy as a source of medicine and food is thought to have coincided with the spread of the Roman Empire with evidence of its use throughout Europe and Asia. However the uses of the opium poppy plant became differentiated between the East and West. The East concentrated more on poppies for opium production and the West for the food value of the seed and oil. On their return to Western Europe, the Crusaders (between 1095 and 1270)

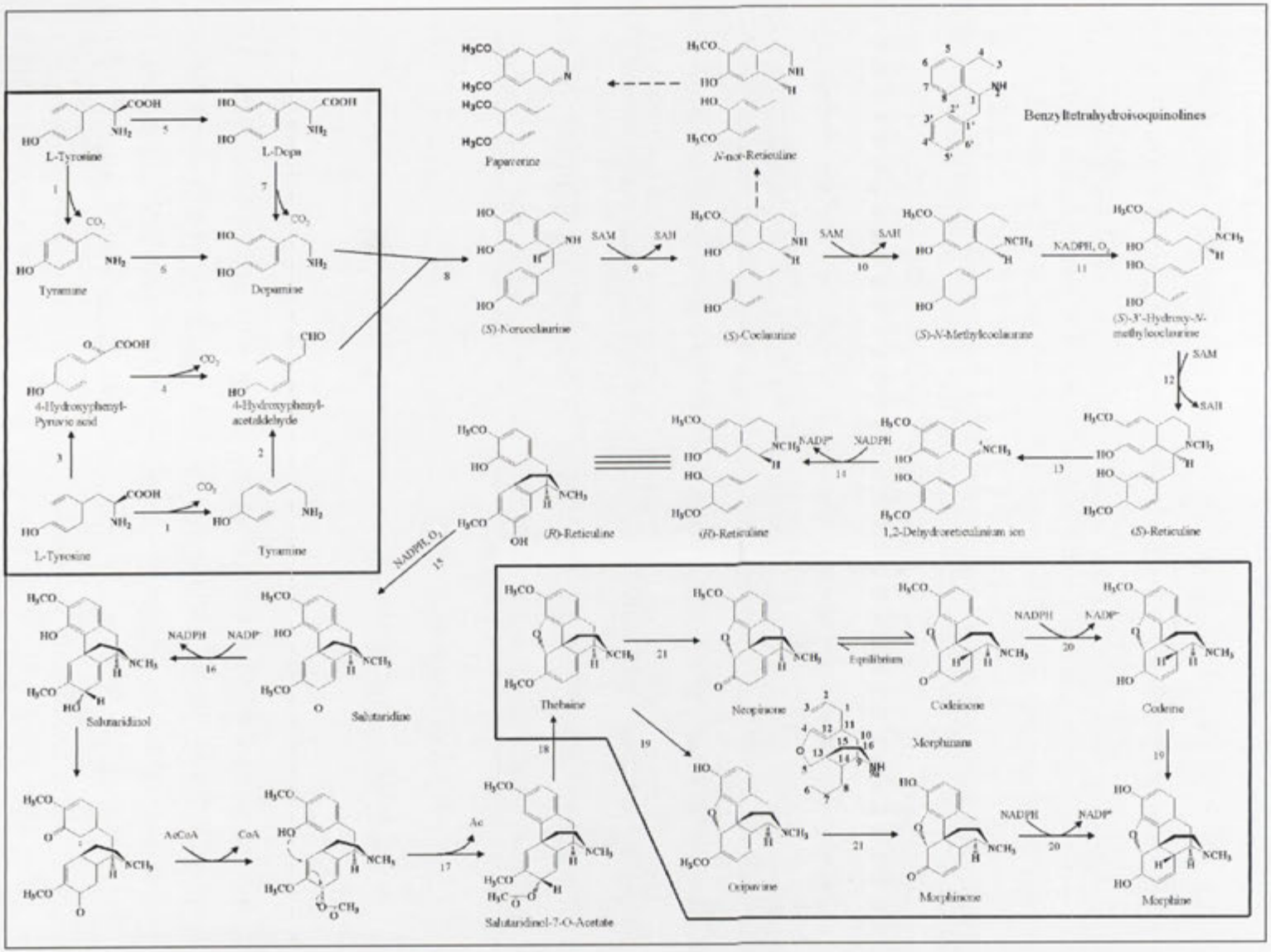

L-Tyrosine
L-Dopa
Tyramine
Dopamine
4-Hydroxyphenyl-Pyruvic acid
4-Hydroxyphenyl-acetaldehyde
Papaverine
N-nor-Reticuline
Benzyltetrahydroisoquinolines
(S)-Norcoclaurine
(S)-Coclaurine
(S)-N-Methylcoclaurine
(S)-3'-Hydroxy-N-methylcoclaurine
(S)-Reticuline
1,2-Dehydroreticulinium ion
(R)-Reticuline
Salutaridinol
Salutaridine
Salutaridinol-7-O-Acetate
Thebaine
Neopinone
Codeinone
Codeine
Morphinone
Morphine
Oripavine
Morphinans
SAM
SAH
NADPH
NADP+
NADPH, O2
AcCoA
CoA
Ac
Equilibrium

Figure 1.3 (Previous Page): The benzylisoquinoline pathway from tyrosine to morphine. The morphinan alkaloids are in the irregular shaped box. 1: Tyrosine decarboxylase, 2: Phenol oxidase, 3: Transaminase, 4: Decarboxylase, 5: Tyrosine aminotransferase, 6: Phenol oxidase 7: Tyrosine decarboxylase, 8: (*S*)-Norcoclaurine synthase 9: Norcoclaurine 6'-*O*-Methyl-transferase 10: (*S*)-Coclaurine *N*-Methyltransferase 11: (*S*)-*N*-Methylcoclaurine 3'-hydroxylase (Cyp80B1) 12: 3'-Hydroxy-*N*-Methylcoclaurine 4'-*O*-methyl transferase 13: (*S*)-Reticuline oxidase 14: 1,2 Dehydroreticulinium ion reductase 15: Salutaridine synthase 16: Salutaridine reductase 17: Salutaridinol acetyl-transferase 18: Thebaine synthase 19: Thebaine demethylase 20: Codeinone reductase 21: Codeine demethylase. Figure adapted from Kutchan 1998; Facchini *et al.*, 1997 and Facchini *et al.*, 2000).

---

reintroduced the use of poppy as a drug. Later it is the alchemist, Paracelsus (1493-1541), trained in Basle but widely travelled through Europe, Russia, China and Turkey, who is generally credited with the introduction of laudanum, an alcoholic extract of opium, which remained popular for social and medicinal use well into the 20th century. After making his first fortune as a buccaneer on the Spanish Main, Thomas Dover (1660–1742) returned to Bristol and made his second fortune from his concoctions called *Dovers Powders*, which had opium and ipecacuanha as active ingredients. The latter is a mixture of isoquinoline alkaloids from *Psychotria ipecacuanha*, a small shrub found in Brazil and Columbia and its root powders were used as an emetic, digestive aid and expectorant. Dover included such large amounts of opium that many apothecaries were said to suggest that their patients settle their affairs and write their wills before using it; a high incidence of mortality obviously resulted from its use.

Addiction from the prolonged misuse of opium was noted as early as the 10th century in Arab literature and manuscripts have been found describing opium abuse in Turkey, Egypt, Germany and England from the 16th century. With the outlawing of tobacco in China, the smoking of opium became prolific, resulting in much addiction. The Chinese passed imperial laws, including capital punishments for the sellers and even confiscated and destroyed opium from

American and British ships importing opium into Canton as an exchange for tea. This angered the British who started the first of the Opium Wars with China (1839-1842), and led to the concession of Hong Kong Island to the British as a reimbursement for the lost opium. The second Opium War in the next decade included American and British fleets taking even more concessions from the Chinese. Each time the concessions included the free trade of opium into China from the British-controlled Bengal. The Boxer rebellion at the turn of the 20th century was another uprising in China to evict the foreigners and with it the opium they brought with them. Again the Chinese lost further economic and territorial concessions. Not until the communists took control of China after World War II could the large rate of addiction be curtailed.

The principal active ingredient of opium poppy, a weak base, was first isolated by Sertürner in 1806 and was named morphine after Morpheus the Greek god of dreams. The purification of morphine gave a 10-fold increase in analgesic effect over opium and provided 19th century physicians with a unique tool for treating both postoperative and chronic pain, and as an adjunct to general anaesthetics. In 1853 this was further enhanced with the invention of the hypodermic syringe allowing a more effective method of delivery. This revolutionised the treatment of the wounded in the Crimean and American Civil wars with soldiers able to self-administer the drug during conflict.

By the time of the 1870s a new phenomenon had arisen with American upper class women hosting morphine parties where the hostess would inject morphine into the arms of the guests. It was widely believed that injection of the drug avoided the addictive qualities of opium smoking and injection of morphine was often prescribed as a cure for this addiction. This resulted in a surge of addiction which coincided with the return home of thousands of Civil War soldiers who were also dependent on the drug from its liberal use. A new search was undertaken to develop an alternative non-addicting opiate.

Heroin (diacetylmorphine) (Figure 1.4) was synthesised in 1874 by Alder Wright and thought to be more effective than morphine and less addictive. In fact heroin was even more addictive than morphine due to its increased lipophilicity and

consequent ability to cross the blood-brain barrier. Heroin was originally marketed by Bayer in 1898 as a cough suppressant in syrup form and quite often given to children as an aid to quieten them. Shortly after this, various governments realised its highly addictive potential and passed legislation to ban its use.

Heroin has been shown to be no more effective as an analgesic and antitussive than hydromorphine (introduced in 1926 by catalytic reduction of morphine to dihydromorphine followed by oxidation), and the drug is now rarely used for this purpose (O'Brien, 2001) (Figure 1.4). Unfortunately it is still available on the illicit market and is a major problem in world illicit drug use and addiction today.

1.4 Importance, Uses and Analogues of Morphinan Alkaloids

Research into the mechanism of opium addiction led to the discovery of highly specific opiate receptors found in brain membranes (Pert *et al.*, 1973; Simon *et al.*, 1973; Terenius 1973). Further studies have revealed three classes of receptors: mu (μ), kappa (κ), and delta (δ), present in nerve cells of the brain, medulla, gastrointestinal regions, and the urogenital regions. These receptors are conserved across species, however their concentration and distribution is not uniform across species (Meijerink *et al*, 1999). The finding of these specific receptors led to the search for endogenous compounds similar to the opiates as it was thought at the time that these receptors were highly unlikely to have co-evolved as a result of using plant opiates or poppy cultivation; these receptors also occur widely in vertebrates. The search for endogenous opioid compounds that also bind the opioid receptors, led to the discovery of three distinct peptide families coded by three separate genes: enkephalins (Hughes *et al.*, 1975), endorphins (Cox *et al.*, 1976), and dynorphins (Goldstein, 1981) (reviewed in Akil *et al.*, 1984). Each of the three families are commonly termed endorphins (from *endo*genous and m*orphine*) and like the receptors are variously distributed across vertebrate species.

The endorphins and opiates work by altering K^+ and Ca^{2+} flux in and out of nerve cells, inhibiting the release of excitatory neurotransmitters and in synaptic activity (Schiff, 2002). The search for endogenous opioids eventually led to morphine itself being isolated from toad skin (Oka *et al.*, 1985). Subsequently several opiates, other than morphine, have been isolated from various mammalian tissues (Goldstein *et al.*, 1985; Donnerer *et al.*, 1986; Weitz, *et al.*, 1986; Donnerer *et al.*, 1986; Kodaira *et al.*, 1989), and recently reticuline has been found in rat brain (Zhu *et al.*, 2003).

Liver homogenates were used to show that transformation of racemic reticuline to salutaridine was possible (Weitz *et al.*, 1987). A cytochrome P450 has been purified and characterised from pig liver and shown to catalyse the phenol oxidative coupling of (*R*)-reticuline to salutaridine (Amann *et al.*, 1995). Cells of rat liver, kidney and brain microsomes were able to transform thebaine to morphine via the same intermediates that occur in plant biosynthesis, namely oripavine and codeine (Kodaira and Spector 1988). These studies illustrate that several of the required enzymes are present in mammalian cells to produce morphine, in particular the critical step from reticuline to salutaridine. There is also some evidence that endogenous morphine and/or codeine levels are elevated in rats after surgery or in humans after fasting (references in Meijerink *et al.*, 1999). While not conclusive these data support current views that morphine is endogenous in mammals.

An aromatic ring and a piperadine ring maintain the stereochemistry at the chiral centre of morphine necessary for its centrally active analgesic activity. By altering the composition of various groups and side groups around the basic molecule, it is possible to alter the effects of morphine to make new drugs with different qualities to morphine such as having fewer side effects, longer lasting analgesia, or having activity as an antagonist of opiates to help against drug addiction.

1.5 Opium and Poppies

Lancing the unripe capsule of *P. somniferum* L. and air-drying the resulting milky exudate, latex, results in the production of opium. The latex is found in a system of specialised cells called laticifers. The laticifers are a series of non-articulated fused cells (anastomising) that form by the gradual thinning of the walls between adjacent laticifer cells (Nessler and Mahlberg 1979). It is within the vacuoles of the laticifers that the opiates accumulate (Homeyer and Roberts 1984). The opium, now an oxidised and hardened product, becomes more dense and tough on storage. This traditional method of harvest of the opium poppy is still practiced today in the major opium producing regions in India. However most of the world's licit poppies are now harvested by taking the dried upper stalk of the plant, including the capsule, which is called the poppy *straw,* and crushing it for solvent extraction of the alkaloids. The seed does not contain any significant quantities of alkaloids and is generally removed before extraction of the alkaloids (Bernath, 1998)

Poppy seed is still used today in Austria, Czech Republic, Estonia, Germany, Netherlands, Poland, Slovakia, Ukraine and the United Kingdom (International Narcotics Control Board, 1993) for culinary purposes as a source of: oils for human consumption (having similar properties to sunflower and olive oil (Beare *et al*., 1979)); for baking in breads and cakes; and as a replacement for soybean meal in animal feed stocks. The seed oil is also used for industrial products such as paints and varnishes (Balbi, 1960). The world licit and illicit drug trade far exceeds the value and importance of the poppy seed trade.

The illicit production of opium poppies has traditionally been found in remote border areas of the Golden Triangle (including Laos, Thailand, and Myanmar), the Golden Crescent (Afghanistan, Iran and Pakistan), India, Lebanon and Mexico. Estimates of world illicit opium production have remained relatively unchanged since the 1990s of between 4000 to 5000 metric tonnes per annum (United Nations World Drug Report 2004). Recently opium production in the Golden Triangle is on the decline and production has shifted more to

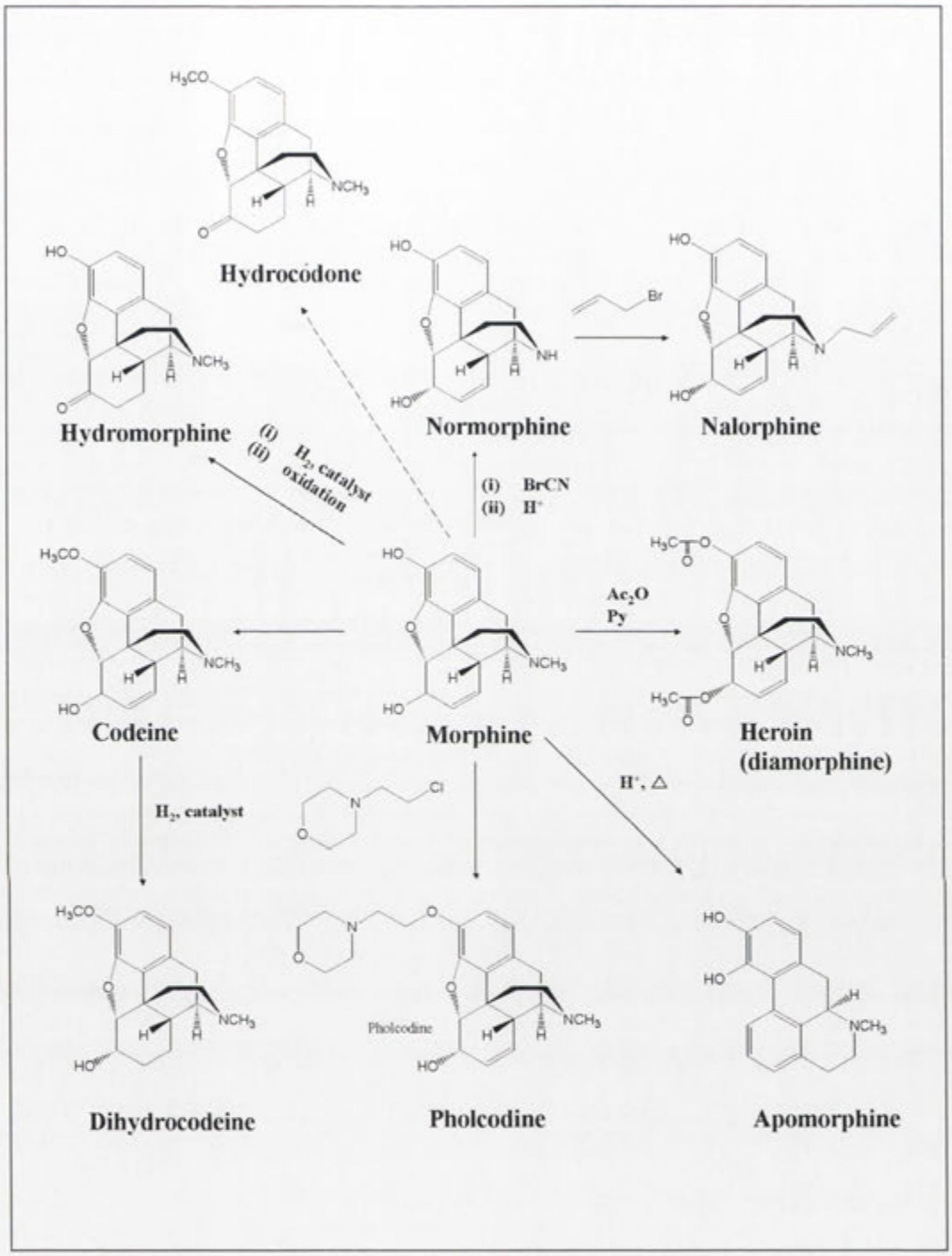


Figure 1.4: The various alkaloids chemically synthesised from morphine. (Adapted from Dewick 1998 and Gagliardi *et al.*, 2003)

Afghanistan, which now accounts for about three quarters of the world's illicit opiate supplies (United Nations World Drug Report 2004).

The licit production of opium poppies for the production of medicinal alkaloids such as the analgesics codeine and morphine, rather than for seed, is the more

important aspect of the cultivation of opium poppy. The cultivation of poppies for medicinal use is increasing with the demand for raw materials to make new semi-synthetic pharmaceuticals including: etorphine, oxycodone, buprenorphine, (Figure 1.5), normorphine, nalorphine, hydromorphine, hydrocodone and dihydrocodone (Figure 1.4). Some countries still produce opium for the home market and for export, including China, India, and the Democratic People's Republic of Korea. Australia, China, Czech Republic, France, Hungary, Slovakia, Spain, Macedonia, Turkey, Yugoslavia all grow *P.somniferum* for the production and sale of alkaloids.

The Australian island state of Tasmania produces approximately half of the world's licit crop with approximately 17,500 hectares planted out in the 2002/3 season (Fist, pers. commun.). Figures collated by the International Narcotics Control Board (2003) reveal 18,477 hectares were planted in 2002 for the production of licit opium, with an expected opium yield of 720 tonnes. In the same year 117,000 hectares were planted worldwide for the production of refined opiates, including purified thebaine, and morphine. The combined production of these two pharmaceuticals is in excess of 62,000 tonnes (International Narcotics Control Board 2003). Worldwide sales of alkaloid-containing drugs were believed to exceed US$4 billion for 2002 (Raskin *et al.*, 2002).

Opium poppy is known to produce over 40 different alkaloids with the morphinan alkaloids, the benzylisoquinoline papaverine (Figure 1.2 and 1.3) and the phthalideisoquinoline noscapine (Figure 1.6) accounting for most of the alkaloid content of opium.

1.6 The Benzylisoquinoline Alkaloid Pathway

The alkaloid pathway (Figure 1.2 and 1.3) in *P. somniferum* can be broken down into two separate parts. The first part of the pathway, central to the production of all benzylisoquinoline alkaloids, entails the differentiation of primary metabolism to secondary metabolism with the concomitant production of 4-hydroxyphenylacetyl-aldehyde and dopamine via decarboxylation, ortho-

hydroxylation, and deamination of L-tyrosine (Rueffer and Zenk, 1987; Stadler *et al.*, 1988)).

The enzyme responsible for the initial decarboxylation step of L-tyrosine, tyrosine decarboxylase, (TYDC) is the only enzyme purified of this part of the pathway, from *P. somniferum* (Facchini and De Luca 1994), *Eschscholzia californica* and *Thalictrum rugosum* (Marques and Brodelius, 1988). TYDC has been cloned from *P. somniferum* (Facchini and DeLuca, 1994; Maldonado-Mendoza *et al.*, 1996) and this aromatic L-amino acid decarboxylase accepts both and exclusively, L-tyrosine and L-dopa as substrates *in vitro*. Expression of two TYDC genes as fusion proteins in bacteria identified the enzyme as accepting DOPA and tyrosine as substrates (Facchini and De Luca 1995a; 1995b).

The TYDCs found in opium poppy are encoded by a minimum 14 genes divided into three different classes (Facchini and De Luca, 1994; Maldonado-Mendoza *et al.*, 1996; Facchini *et al.*, 1998) with different spatial and temporal expression patterns (Facchini and De Luca 1995; Park *et al.*, 1999; El-Ahmady and Nessler 2001)). Rueffer and Zenk (1987) used precursor-feeding studies, isolation of intermediates and *in vivo* NMR techniques to find evidence suggesting the existence of more than one biosynthetic route from L-tyrosine to dopamine, via conversion of tyrsoine to either tyramine or L-Dopa (3,4-dihydroxyphenylalanine).

The other intermediate involved in the norcoclaurine synthase reaction, 4-hydroxyphenylacetylaldehyde could also have two alternate precursors, both direct products of tyrosine, tyramine or 4-hydroxyphenyl-pyruvic acid.

The stereospecific enzyme norcoclaurine synthase (NS) uses a Mannich-like reaction to condense dopamine and 4-hydroxyphenylacetylaldehyde to yield (*S*)-norcoclaurine, the precursor to all benzylisoquinoline alkaloids in plants (Stadler *et al.*, 1987; 1989, Samanani and Facchini 2001; 2002). Norcoclaurine synthase fails to discriminate between 4-hydroxyphenylacetylaldehyde and 3,4-dihydroxyphenylacetalaldehyde (Loeffler *et al.*, 1987) and it was originally

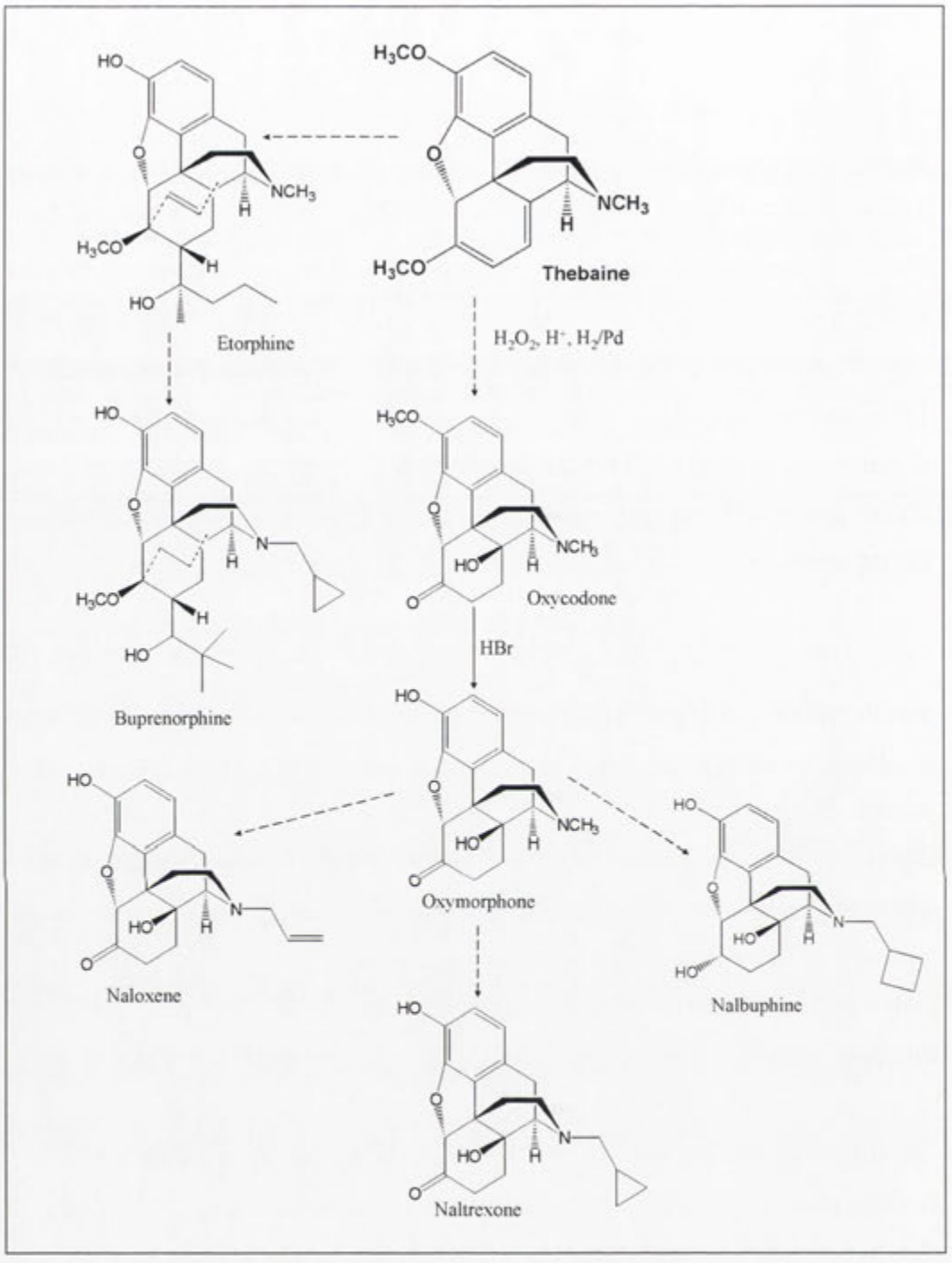


Figure 1.5: Molecules chemically synthesised from thebaine. (Adapted from Dewick 1998)

believed that the tetrahydroxybenzylisoquinoline alkaloid (*S*)-norlaudanosoline was the precursor to benzylisoquinoline alkaloids (Rueffer *et al.*, 1981; 1983), however, only (*S*)-norcoclaurine is found in plants. (*S*)-norcoclaurine undergoes three methylation steps and one hydroxylation step to form (*S*)-reticuline, a key

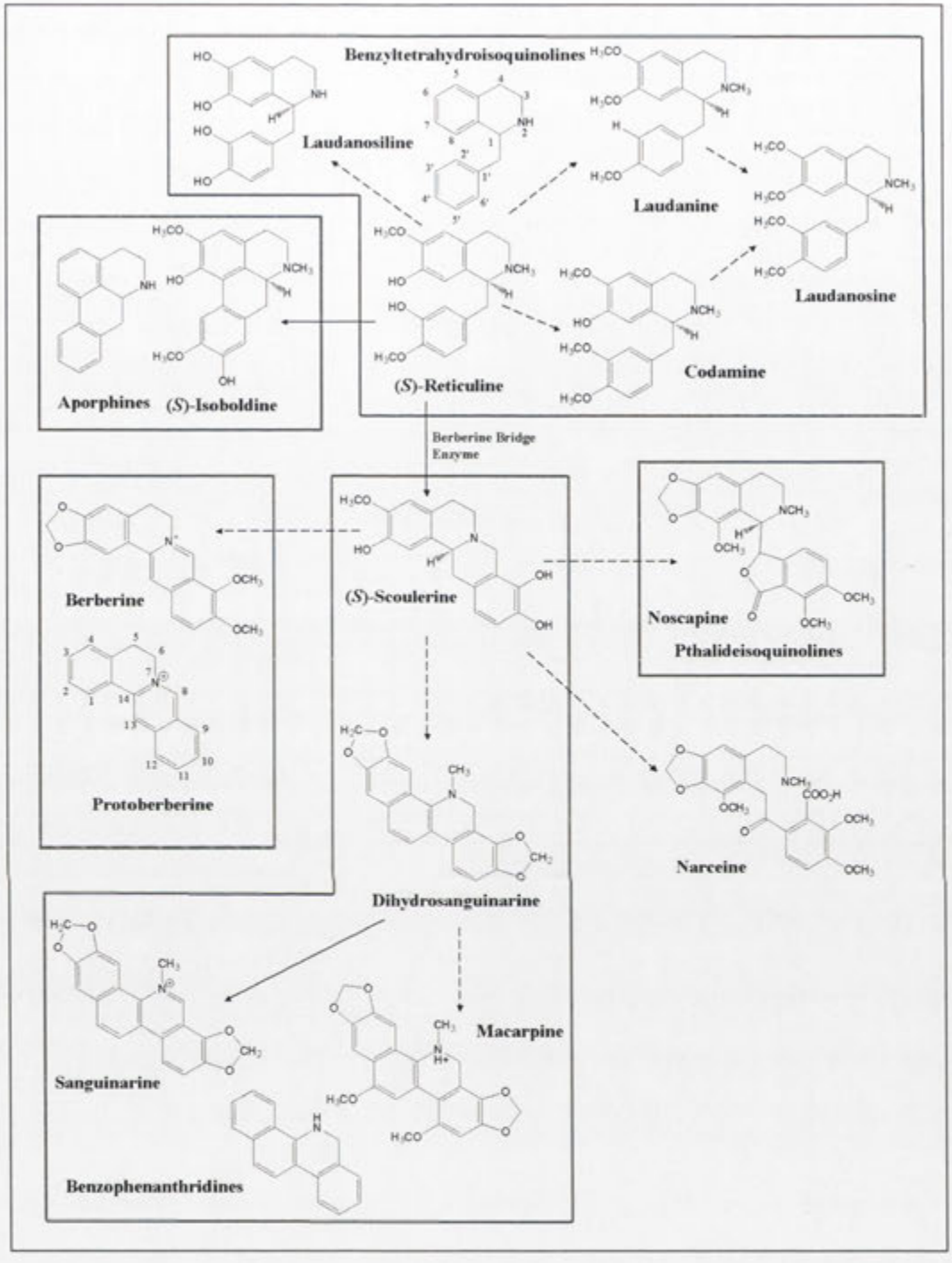


Figure 1.6: The different alkaloids derived from the central benzylisoquinoline intermediate (*S*)-reticuline. (Adapted from Fachini *et al.*, 1997 and Facchini *et al.*, 2000).

branchpoint in the manufacture of tetrahydroberberine, protopine, benzophenanthridine and morphinan alkaloids.

(*S*)-coclaurine is produced from the first of the methylations of (*S*)-norcoclaurine to its 6-carbon moiety and occurs by the transfer of the *S*-methyl group of *S*-adenosyl-L-methionine (SAM). The cytosolic enzyme catalysing this reaction is norcoclaurine 6-*O*-methyltransferase (6-OMT) (Reuffer *et al*., 1983; Frenzel and Zenk, 1990; Hara, et. al., 1994; Sato *et al*., 1994). The gene for this enzyme has been cloned from *Thalictrum tuberosum* (Frick and Kutchan, 1999), *Coptis japonica* (Morishige, *et al*., 2000) and more recently from *P. somniferum* (Ounaroon, *et al*., 2003).

The next catalytic step in the pathway to (*S*)-reticuline is the methylation of the *N* moiety of (*S*)-coclaurine by coclaurine *N*-methyltransferase (CNMT) (Frenzel and Zenk, 1990; Wat *et al*., 1986; Choi *et al*., 2001). CNMT catalyses the transfer of a methyl group from SAM to (*S*)-coclaurine to form (*S*)-*N*-methylcoclaurine (Table 1.1) and the gene has been cloned from cultured cells of *Coptis japonica* (Choi *et al*., 2002) and more recently from *P.somniferum* (Ounaroon, *et al*., 2003). (*S*)-coclaurine can also be converted to another major alkaloid, papaverine, via *N*-nor-reticuline (Uprety *et al*., 1975).

(*S*)-*N*-methylcoclaurine was previously thought to be hydroxylated by a phenolase (Kametani *et al*., 1970,1973), but studies since have revealed that this reaction is catalysed by CYP80b1 (Table 1.1), a methyl jasmonate inducible cytochrome P450-dependent mono-oxygenase also called (*S*)-*N*-methylcoclaurine 3'-hydroxylase (Pauli and Kutchan, 1998).

Subsequent to this hydroxylation reaction is the final methylation to produce (*S*)-reticuline undertaken by *S*-adenosyl-L-methionine: 3'-hydroxy-*N*-methylcoclaurine 4'-*O*-methyltransferase (4'-OMT). This enzyme was co-purified with 6-OMT from *Coptis japonica* (Morishige *et al*., 2000); the two enzymes showed high substrate specificity for their respective substrates.

Cloning and expression of the genes encoding different OMTs from *Thalictrum tuberosum* (Frick and Kutchan 1999) showed that these OMTs acted enzymatically as heterodimers. Expressing different monomer combinations of these heterodimers in baculoviruses, led to different substrates being

preferentially *O*-methylated (Frick *et al.*, 2001). This suggested that different plant species have evolved different methods of *O*-methylation. 7-OMT is responsible for the methylation of reticuline to form laudanine (Figures 1.2 and 1.7). Recent cloning of the 6-OMT and 7-OMT (Table 1.1) genes from *P. somniferum* has shown that these genes most closely resemble the OMT genes isolated from *C. japonica* rather than the *T. tuberosum* genes (Ounaroon, *et al.*, 2003).

(*S*)-reticuline itself is a key intermediate in the biosynthesis of a number of structurally diverse isoquinoline alkaloids (Figure 1.6). These include the morphinan alkaloids and benzophenanthridine alkaloids, the latter including berbamunine, (+)-tubocurarine, sanguinarine and approximately 270 other dimeric bisbenzylisoquinoline alkaloids (Facchini, 2001). The phthalideisoquinoline group of alkaloids, including narcotine, narcotiline and narceine are formed from (*S*)-reticuline via scoulerine and noscapine (Nyman and Hanson, 1978). Not all of these pathways appear in *P. somniferum*, however the antitussive agent noscapine (Ye et. al., 1998) and the antimicrobial sanguinarine do accumulate in opium poppy.

The *P. somniferum* benzylisoquinoline alkaloid papaverine (Figure 1.1 and 1.3) is derived from the hydroxylation and methylation of (*S*)-coclaurine to *N*-nor-reticuline and subsequent methylations and dehydrogenations (Dewick, 1998). To date no radioactive precursor studies of possible intermediates, or enzyme activity studies have been instigated into the synthesis of the muscle relaxant papaverine, leaving the author (Dewick, 1998) to postulate this route of synthesis.

A number of other alkaloids have been found in *P. somniferum* whose exact biochemical origin is yet to be determined. These alkaloids include (*S*)-norreticuline, (*S*)-norlaudanine, (*S*)-norcodamine, (*S*)-codamine, (*S*)-norlaudanosine, (*S*)-laudanosine and (*S*)-laudanosiline (Uprety *et al.*, 1975; Ayyangar and Bhide, 1988). These products have been difficult to study due to their infrequent occurrence and low levels of accumulation.

1.7 Morphinan Alkaloids

Morphine and codeine are present in considerable quantities in the opium poppy with morphine being the dominant alkaloid present in commercial crops. Thebaine, the methyl enol ether of codeinone, is also found in opium poppy and while not used medicinally itself, is important in the manufacture of codeine and several other important semi-synthetic compounds. Thebaine is not readily converted into heroin making it more attractive as an alternative morphinan alkaloid source. Other synthetic morphinan alkaloids include the 14-hydroxymorphinans, such as oxycodone, naloxene, naltrexone, nalbuphine and nalmefene (Figures 1.4 and 1.5), which are important due to their activities as potent analgesics and/or narcotic antagonists and are manufactured from thebaine as a starting material.

Other important opiate derivatives such as the ring-C bridged compounds buprenorphine and etorphine, are also most practically prepared from thebaine. Due to the relatively small amounts of thebaine found in *P. somniferum* most of the material prescribed was previously produced by synthesis from morphine. To better serve this growing market many attempts have been made to find *P. somniferum* or alternative plant species, which accumulate thebaine.

Conversion of (*S*)-reticuline to its (*R*)-epimer is the first committed step in the production of morphinans (Figure 1.3), although it is only with thebaine that the pentacyclic morphinan structure is formed. The epimerisation to (*R*)-reticuline is achieved in two steps, first by an oxidation-reduction process forming the intermediate 1,2-dehydroreticulinium ion (De-Eknamkul and Zenk, 1990; Hirata *et al.*, 2004). The NADPH-dependent enzyme, 1,2-dehydroreticuline reductase, which has been purified from opium poppy seedlings, stereospecifically reduces 1,2-dehydroreticuline to (*R*)-reticuline (De-Eknamkul and Zenk, 1992). This cytosolic enzyme is only found in plants that synthesise morphinan alkaloids.

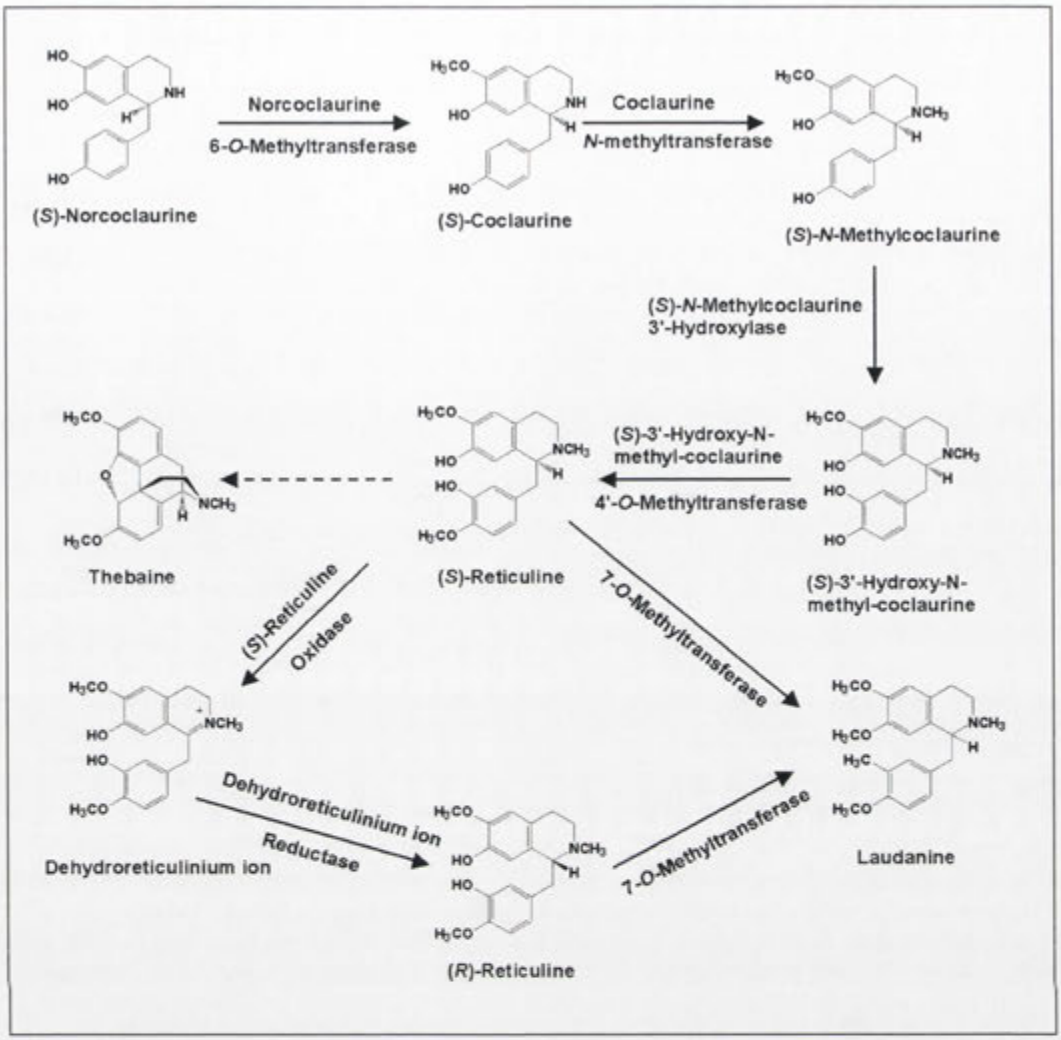


Figure 1.7: Schematic biosynthetic pathway leading from (*S*)-norcoclaurine to (*S*)-reticuline and thebaine. (*R*, *S*)-reticuline 7-*O*-methyltransferase can act on the *R* or *S* enantomer of reticuline to form laudanine (Adapted from Ounaroon *et al.*, 2003).

The *R* enantiomer of reticuline undergoes a C-C phenol-coupling reaction of C12 to C13 forming the phenanthrene ring system of salutaridine. Previously this class of reaction was attributed to laccases, phenolases and peroxidases (Kametani *et al.*, 1970, 1973) because of the relative abundance and broad substrate specificities of these plant enzymes. However *in vitro* analysis of purified extracts revealed catalysis to be performed by a highly substrate specific membrane bound cytochrome P450 enzyme known as salutaridine synthase (Gerardy and Zenk, 1993a).

Salutaridine, found as a minor alkaloid constituent in the opium poppy, undergoes a stereospecific reduction of the carbonyl group by the cytosolic enzyme salutaridine: NADPH 7-oxidoreductase (Gerardy and Zenk, 1993b) to produce 7(*S*)-salutaridinol. Treatment of the phenanthrene salutaridinol with acid *in vitro* can demonstrate the next step in the pathway creating the ether linkage ring closure of C-4 and C-5 by nucleophilic attack of the phenol group on to the dienol system and subsequent displacement of the hydroxyl group. However in the plant an esterification reaction is undertaken to provide a better leaving group by the enzyme acetyl coenzyme A: salutardinol-7-*O*-acetyltransferase (SAT) to produce salutaridinol-7-*O*-acetate. The ring closure then occurs readily but dependent on pH 6-7 and possibly without any enzyme participation (Lenz and Zenk, 1994) to form the pentacyclic morphinan skeleton. SAT has been purified from *P. somniferum* cell cultures (Lenz and Zenk, 1995a) and molecular characterisation of the gene has shown the gene is expressed in root, stem, leaf and capsule (Grothe *et al.*, 2001).

Parker *et al.*, (1972) first proposed the pathway to morphine from thebaine via neopinone, codeinone and codeine. Thebaine undergoes an enol ether cleavage to neopinone then subsequent non-enzymatic isomerization to codeinone (Gollwitzer *et al.*, 1993). Reduction of the enol ether codeinone is catalysed by the NADPH-dependent codeinone reductase purified from both *P. somniferum* cell suspensions and plants. Enzyme activity, from highest to lowest, could be detected in capsule, shoot, root, leaf and cell cultures (Lenz and Zenk, 1995b; Lenz and Zenk, 1995c). Cloning of the gene and the analysis of its expression levels showed that the highest level of expression could be found in roots with decreasing levels of activity found in stem, capsule, leaf and cell suspensions (Unterlinner *et al.*, 1999).

Brochmann-Hanssen (1984) used radioactive feeding studies to show an alternate pathway to morphine in Tasmanian *P. somniferum* cultivars via the demethylation of thebaine at carbon atom 3 to yield oripavine. Oripavine is thought to be demethylated via hydroxylation at carbon atom 6 to produce morphinone with a carbonyl group at carbon atom 6. This in turn undergoes stereospecific reduction of the carbonyl group at carbon atom 6 to morphine.

Codeinone reductase has also been shown to reduce morphinone to morphine (Hodges & Rapoport, 1980; Lenz and Zenk 1995c) suggesting that this enzyme may act in both pathways from thebaine to morphine (Figure 1.3).

1.8 Compartmentalisation of Alkaloids

It is now well established that some alkaloids are produced in one part of the plant and transported to other parts of the plant. For example the tropane alkaloids (reviewed in De Luca and St-Pierre, 2000) and nicotine are mainly synthesised in the root and transported to aerial parts where they accumulate. Sanguinarine and morphine share the common intermediate reticuline yet sanguinarine accumulates in the roots and morphine in the aerial part of the plant (Facchini and De Luca, 1995). Due to the general toxicity of alkaloids especially in large concentrations most of them are sequestered into various compartments in the plant and usually into the plant vacuole.

Several sequential enzymatic steps in alkaloid biosynthesis have been shown to occur in distinct cell types suggesting that intercellular transport of intermediates and specific enzyme localisation occurs (Bird *et al.*, 2003). Measurements of the alkaloids morphine, codeine, thebaine, papaverine, and noscapine in latex showed that more than 90% of the alkaloid content was found in two different types of vacuole (Fairbairn *et al.*, 1974; Fairbairn and Steele, 1981; Homeyer and Roberts, 1984; Roberts *et al.*, 1983; Roberts, 1987). One vacuole type was designated 900xg and the other 1100xg due to their sedimentation properties at these respective centrifugal forces. The 900x g variety seemed to contain the higher proportion of alkaloids of the two types of vacuoles and seemed to accumulate all of the aforementioned alkaloids except thebaine (Figure 1.8) in the three weeks after petal opening (Pham and Roberts, 1991). Dopamine (Homeyer and Roberts 1984; Roberts 1987) and (*S*)-reticuline (Deus-Neumann and Zenk 1986) have also been shown to accumulate in the latex vacuoles.

The study of the morphine pathway is limited by several factors. Firstly morphinan alkaloids are only found in appreciable amounts in the Papaver

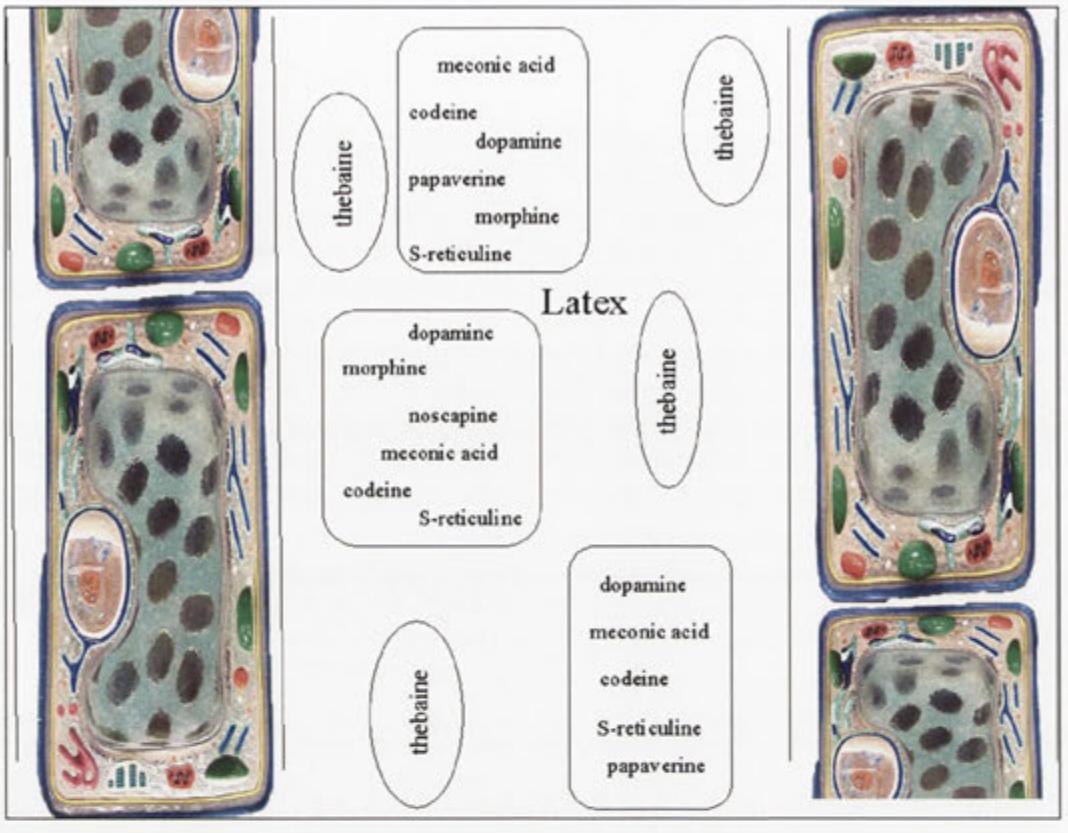


Figure 1.8: Diagram illustrating where the alkaloids have been found to concentrate in the vesicles of the latex. Thebaine accumulates in the smaller (oval shapes) 1100x g vesicle fraction and the other morphinan alkaloids accumulate in the larger (rectangles) 900x g vesicle fraction along with noscapine, papaverine, reticuline and tyrosine (Adapted from Pham and Roberts 1991).

---

species such as *P. orientale*, *P. somniferum*, *P. somniferum subspecies setigerum* and *P. bracteatum*. Secondly very few mutants have been reported in the literature. Thirdly, although a transformation system is now available for opium poppy few genes have been characterised or cloned in order to retransform them into the plant. A further reason is that *P. somniferum* cell suspensions fail to reliably accumulate appreciable quantities of morphinan alkaloids.

Thebaine, a precursor of morphine synthesis, is useful as a starting compound for several important pharamaceuticals including; naltrexone and buprenorphine. Thebaine is the major alkaloid found to accumulate in *P. bracteatum* Lind. and can constitute up to 3.5% of the dry weight material in the mature capsule

(Vincent et. al., 1977). Several studies have tested this species as a potential crop plant (Aynehchi and Jaffarian, 1973; Vincent *et al.*, 1977; Corrigan and Martyn, 1981; Seddigh *et al.*, 1982) and high yields of thebaine are possible from this species, however its biennial growth habit has made it uneconomical to use as a replacement for the annual *P. somniferum,* which is a much more reliable source. *P. bracteatum* yields decrease with overcrowding (Bare, *et al.*, 1987), and the self-incompatibility system of this species makes it impossible to breed pure lines through selfing. *P.bracteatum* has been successfully regenerated in tissue culture (Day *et al.*, 1986) making micropropagation of selected elite genotypes a theoretical possibility, however this would make pure lines for commercial production too expensive. Further to the problems of utilisation of *P. bracteatum* as an alternative source of thebaine, are the socio-political problems associated with countries, such as Turkey and India. These countries have become reliant on the incomes derived from *P. somniferum* crops and the loss of these crops would be problematic if there was increased competition from industrialised western countries growing the non-narcotic *P.bracteatum* (Theuns *et al.*, 1986).

Attempts to use *in vitro* cell cultures of *P. bracteatum* (Kamo and Mahlberg, 1988) as an alternative source of thebaine have proved unsuccessful, as cultures failed to accumulate appreciable levels (Kutchan et. al. 1983). Kamimura *et al.* (1976) indicated that morphogenetic differentiation from cultured cells of *P. bracteatum* was needed before higher levels of thebaine could be produced. Various researchers have confirmed the need for some vascular differentiation in *P. bracteatum* cultures before any thebaine is detectable (Staba *et al.*, 1982; Kutchan et. al., 1983). Similarly *P. somniferum* cell suspensions were attempted to provide an alternative source of morphinan alkaloids with the same result.

The need to induce morphogenesis in cell cultures dramatically increases the technical difficulties and reduces the economic feasibility of using them as a production system. Reports have been published suggesting that morphinan alkaloids can accumulate in cell suspensions (Kamimura *et al.*, 1976; Tam *et al.*, 1980; Hodges and Rapoport 1982; Kamo et. al., 1982; Staba *et al.*, 1982; Lockwood 1983; Hsu and Pack, 1984; Yoshikawa and Furuya, 1984; Griffing et. al., 1989; Siah and Dorian, 1991; Facchini and Bird, 1998), but these reports

were all of very small quantities and were not reproducible. One cell suspension of *P. somniferum* was shown to reliably produce small quantities of thebaine as the only morphinan alkaloid (Gerardy and Zenk, 1993a). This is despite the fact that enzymes from this part of the pathway could be purified from suspension cultures (Wilhelm and Zenk 1997). This problem is not limited to the morphinan alkaloids; other alkaloids such as vinblastine, hyoscyamine, scopolamine, and vincristine have also proved difficult to accumulate in cell suspensions of their respective source species (Verpoorte et. al., 2000).

Many isoquinoline alkaloids can be produced in de-differentiated cell cultures of the *Berberis* species and this has led to the advancement of the study of many of the alkaloids leading to (*S*)-reticuline (reviewed in Zenk, 1995). The failure to reliably accumulate the morphinan alkaloids in poppy cultures suggests that there is some developmental, compartmental or transport control operating which prevents these compounds being synthesised in the absence of differentiated laticifer cells (Kutchan et. al., 1983; Yoshikawa and Furuya, 1984; Hsu and Pack, 1984; Galewsky and Nessler, 1986; Griffing *et al.*, 1989; Siah and Doran, 1991). As a consequence cell cultures are no longer considered a viable alternative to cultivated poppy crops in the production of morphinan alkaloids. They are also not a viable system to study the final steps in morphine biosynthesis.

1.9 Alkaloid Transport

Some form of alkaloid transport mechanism must exist for morphinan alkaloids as many of the intermediates are known to accumulate in specialised cells in the vacuoles and the enzymes necessary for several of the steps of morphinan biosynthesis are not found in the latex. The cytochrome P-450 enzyme salutaridine synthase has been shown to be active in roots, shoots, capsule and leaf vein but not in the latex (Gerardy and Zenk, 1993b).

Two hypothetical possibilities for alkaloid transport in *P. somniferum* have been suggested: passive diffusion with an "ion trap" mechanism and specific carrier mediated transport.

Gene	Species	Cloning Method	Identity (aa) %	Other	References
O-methyl transferases	*T. tuberosum* 6 alleles	HB	93.2-99,7% to each other	5' and 3' UTRs differ Methylated catechols and phenylpropanoids. Cytosolic. 4 clones isolated. MeJ induction.	Frick and Kutchan (1999) AF064693-7
4' *O*MT	*C. japonica* *P. somniferum* 2 alleles	PP	*P. somniferum* 80-83% homology to *C. japonica.*	Active enzyme was a dimer. No divalent cations req. for activity. Inhibition by Fe^{2+}, Cu^{2+}, Co^{2+}, Zn^{2+}, Ni^{2+}. But none by SH. Cytosolic. 4'*O*MT works better with N-CH3 substrate.	Morishige et. al., (2000) D29811 Facchini and Park (2003) AY217333-4
6'*O*MT	*C. japonica* *P. somniferum* 1 allele	PP	*C. japonica* 6'OMT is 51% to *C. japonica* 4'OMT *P. somniferum* is 63% to the *C. japonica* 6'OMT.	Active enzyme was a dimer. No divalent cations req. for activity. Inhibition by Fe^{2+}, Cu^{2+}, Co^{2+}, Zn^{2+}, Ni^{2+}. But none by SH. Transcript expressed in stem, bud, root, leaf and least in capsule. Cytosolic.	Morishige *et al.*, (2000) D298212 Facchini and Park (2003) AY217335 Ounaroon *et al.*, (2003) AY268894
7'*O*MT	*P. somniferum*	PP	36% to *C.jap* 6'OMT.	Transcript found in bud, stem and less in leaf. Absent in root	Ounaroon *et al.*, (2003) AY268893
cyp80b1	*E. californica* 3 alleles *P. somniferum* 2-3 alleles	HB	*P. somniferum* 77% to *E. californica*	2 isoforms MeJ Induction. Membrane bound. Transcript present in all organs. Lowest in capsule and bud.	Pauli & Kutchan (1998) AF014800-3 Facchini & Yu (1999) Unpublished AF191772
Berberine Bridge Enzyme	*E. californica* 2 alleles *P. somniferum* cv Marianne 2 alleles	PP	76% to *P. somniferum* clone	Putative FAD binding site. MeJ induction. Promoter doesn't express in *P. somniferum* cell suspensions. Membrane bound. 1 pseudogene. Transcript expressed in leaf, stem and root.	Dittrich and Kutchan (1991) M77150 Facchini *et al.*, (1996) U59232-3

				Putative hydrophobic N-terminal signal peptide.	
NADPH cytochrome P450 reductase	*E. californica* 1 allele *P. somniferum* 1 allele	HB	*E. californica* and *P.somniferum* are 69% identical.	Km cytochromeP450, 8.3 uM Km NADPH 4.2 uM pH optimum 8.0. 80 kDa in size 2 isoforms. MeJ induction. Membrane bound.	Rosco *et al.*, (1997) U67185-6
Salutaridine 7-*O*-acetyl transferase	*P. somniferum* 1 allele	PP	37 % identitical to *C. roseus* deacetylvindoline acetyltransferase	Transcript found in *P. orientale* and *P. bracteatum*. Cytosolic. Similar expression in capsule, root, stem and leaf.	Grothe *et al.*, (2001) AF339913
N-methyl transferase	*C. japonica* *P. somniferum*	PP	*P. somniferum* is 80% to *C. japonica*.	Low enzymatic activity at this step in *C. japonicus* (Choi *et al.*, 2001) Cytosolic.	Choi *et al.*, (2002) AB061863 Facchini and Park (2003) AY217336
Tyrosine decarboxylase	*P. somniferum*	HB		TYDC5 not in latex. (El-Ahmady and Nessler 2001)). MeJ induction. Cytosolic. Also in *A. thaliana*.	Maldonado-Mendoza *et al.*, (1996) U16804
Codeinone reductase	*P. somniferum* >6 alleles	PP	Cor 1.1-1.4 share 95-96% sequence identity and only 70% to cor 2 at the nucleotide level. 53% to 6'-deoxy-chalcone synthase aa level.	Each of cor 1.1-1.4 have one intron of 443 bp starts at 561 bp. Constitutive expression. Cor 2 has 2 introns at 321 and 514 bp. Not functional. Cytosolic. No MeJ induction. Cor 1 found in *P. rhoeas*, *P. orientle* and *P. bracteatum* (Huang and Kutchan, 2000).	Unterlinner *et al.*, (1999) AF108432-8

Table 1.1: Genes cloned from the alkaloid pathway. (See Figure 1.3). Berberine bridge enzyme and (*R*, *S*)-reticuline 7-*O*-methyltransferase are also included as side branches. For cloning of genes from another species before *P. somniferum* the cloning method and reference are listed first. The cloning method is described by the abbreviations: PP, protein purification; HB, homology based isolation and functional expression. All transcript levels are listed from highest to lowest expression levels unless otherwise stated. MeJ is methyl jasmonate addition to cell suspensions.

The "ion trap" hypothesis proposes that the lipophilic alkaloids can passively diffuse from the cytoplasm across the vacuolar membranes where they are trapped in the acidic vacuolar compartment by protonation and salt formation (Neumann *et al.*, 1983; Matile 1984; reviewed in Roberts *et al.*, 1991). It has been shown in *Chelidonium majus* that chelidonic acid readily complexes with alkaloids (Hauser and Wink, 1990) and thereby provides the acidic properties required to allow the accumulation against a concentration gradient; this mechanism has gained support from studies of the uptake of sanguinarine and various other lipophilic alkaloids into latex vacuoles (Matile, 1976; Hauser and Wink, 1990). In *P. somniferum* meconic acid is the major complexing agent found to accumulate in the latex in high concentrations (Fairbairn and Williamson, 1978). High levels of sulphate and chloride were also found in latex suggesting that there may be some form of ion trap mechanism for the major alkaloids of opium poppies (Pham and Roberts, 1991).

There is no direct evidence to suggest an "ion trap" is responsible for the accumulation of morphinan alkaloids in latex. Evidence against the ion trap mechanism of transport comes from the observation that morphine accumulates in opium poppy in preference to several more basic alkaloids. An addition of ATP caused morphine uptake into the vacuoles to increase, suggestive of a specific transporter. Despite this, the addition of ATP inhibitors did not reduce morphine uptake into the vacuoles, suggesting that there may be a specific alkaloid channel as well as a specific morphine transporter (Roberts *et al.*, 1991). It is thought that the vacuoles of *P. somniferum* latex may be specialist vacuoles due to their size, resistance to external measures designed to reduce the Δ pH, and capacity to sequester millimolar quantities of alkaloids (Roberts *et al.*, 1991). Further to this the rapid and specific uptake of the major opium alkaloids has no absolute requirement for MgATP, is independent of temperature and does not show saturation kinetics (Homeyer and Roberts, 1984).

The second hypothetical method of alkaloid accumulation in the vacuole involves specific carrier mechanisms. This is evident for (*S*)-reticuline, a charged species

under cytosolic pH, and this mechanism appears to be stereoselective for the *S*-rather than the *R*-configured compound (Deus-Neumann and Zenk, 1986). This discrimination for the naturally occurring (*S*)-enantomiers has also been shown for the alkaloid (*S*)-scoulerine (Deus-Neumann and Zenk, 1986). Transport is thought to occur by a H^+ antiport mechanism for alkaloids such as lupanine (Mende and Wink, 1987), (*S*)-reticuline, (*S*)-scoulerine (Deus-Neumann and Zenk, 1986), and 1-(malonylamino)-cyclopropane-1-carboxylic acid (Bouzayen *et al.*, 1989). H^+-ATPase is a major vacuolar protein (Leigh and Walker, 1980; Sze, 1985; Bremberger *et al.*, 1988) and along with pyrophosphatase (Leigh and Walker, 1980; Rea and Sanders, 1987; Bremberger *et al.*, 1988) is responsible for tonoplast H^+ concentration being greater than in the cytoplasm (Thom and Komor, 1984; Sze, 1985; Rea and Sanders, 1987; Heidrich *et al.*, 1989; Taiz, 1992). The proton gradient is important for substrate-proton antiport transport (Lüttge *et al.*, 1981; Hager and Hermsdorf, 1981; Thom and Komor, 1984, 1985; Briskin *et al.*, 1985; Blumwald and Poole, 1985 a,b; Blumwald, 1987; Blackford *et al.*, 1990; Getz, 1991).

Increasing interest in the transport mechanisms in plants, in particular for secondary metabolites, has led to a number of new classes of transporters being isolated and characterised. These include the multidrug resistance (MDR) class of proteins that have now been found in large numbers in plants (Rea *et al.*, 1998, Theodoulou, 2000), including one recently found in *Coptis japonica* and possibly involved in alkaloid transport (Yazaki, *et al.*, 2001). The glutathione (*S*)-transferase class of transporters may also be involved in alkaloid transport. The maize *Bronze-2* gene, which encodes a protein thought to be responsible for the deposition of anthocyanins in the vacuole, has been shown to be a glutathione (*S*)-transferase (Marrs *et al.*, 1995). There is still some conjecture as to how the glutathione (*S*)-transferase transports the anthocyanins into the tonoplast. A type III glutathione (*S*)-transferase has been cloned from fungal elicited *P. somniferum* cell suspensions (Yu and Facchini, 2000) using degenerate primers to other plant GSTs in the database, however it still remains to be shown if it has any role in alkaloid trafficking.

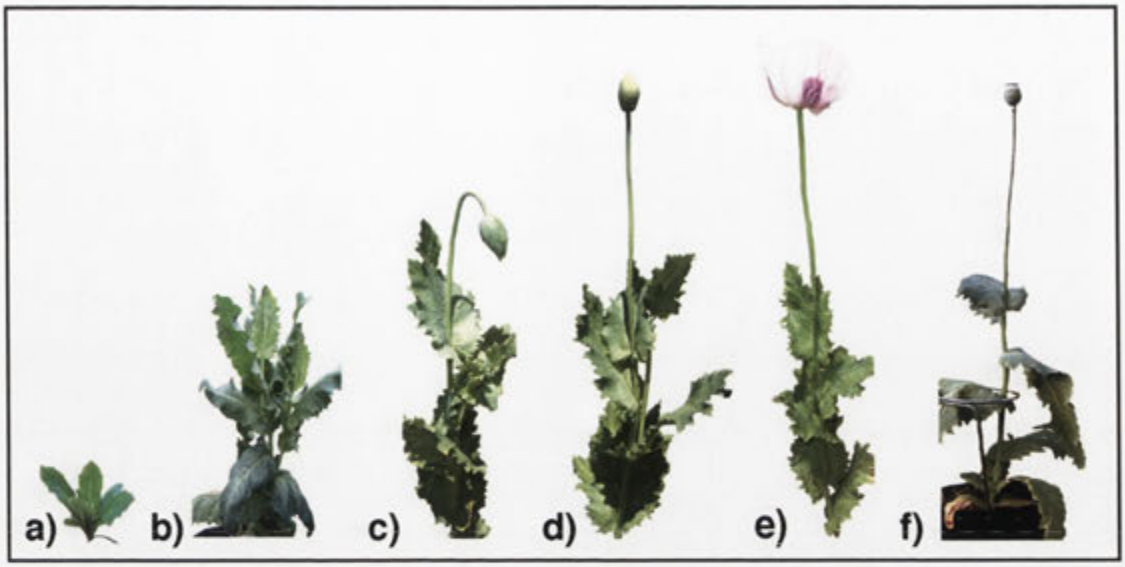


Figure 1.9: The different developmental stages of the opium poppy *P. somniferum*. The six to eight leaf prehook-stage (b) prior to the hook stage (c) is where the maximum production of morphinan alkaloids occurs in the plant. The final accumulation of morphinan alkaloids occurs in the capsule at the final stage (f). (Diagram not to scale).

1.10 Regulation of the Morphinan Alkaloid Pathway

Early studies have shown that alkaloids are regulated in such a way as to allow accumulation of particular alkaloids in certain parts of the *P. somniferum* plant. Morphine is the dominant alkaloid in the aerial part of the plant (Facchini and De Luca 1995) and accumulates in the laticifer system of the mature capsule. The developing poppy plant also accumulates morphinan alkaloids at varying rates as the plant matures (Williams and Ellis, 1989). See Figure 1.9 for the different developmental stages of *P. somniferum*.

Several genes from the alkaloid pathway have been cloned from *P. somniferum* or other alkaloid producing species (summarised in Table 1.1) and have been shown to be inducible and regulated developmentally (Facchini and Park 2003). Gene transcripts of *cyp80b1*, *tydc* and *bbe* are all upregulated with the addition of methyl

jasmonate to cell suspensions. However no corresponding increase in morphinan yields occurred, suggesting that there are still transcripts involved in morphine biosynthesis that are not regulated by methyl jasmonate. *Cyp80b1* and *bbe* were differentially expressed in developing plants, across mature plants and in cell suspensions (Huang and Kutchan 1999), while codeinone reductase isoform 1 was not. *Cor 1* expression remained unchanged with the addition of methyl jasmonate. *Cor 1* gene transcript was also much higher in the stem than root or bud, while *bbe* transcript was not detected in the bud or capsule of the mature plant. Both *tydc* and *bbe* transcript are induced by the addition of methyl jasmonate and both promoters have been isolated, although no common promoter elements have yet been found (Park *et al.*, 1999). More recent work with newly cloned genes *N*-methyl transferase, 4'-*O*MT, 6'-*O*MT and salutaridine acetyl transferase, showed these gene transcripts were also upregulated in cell suspensions with the addition of methyl jasmonate elicitor or wounding (Facchini and Park, 2003).

In cultured cells of *E. californica* the addition of 2mM butyric or pivalic acid led to the acidification of the cytoplasm that subsequently led to the induction of alkaloid biosynthesis (Roos *et al.*, 1998). This mimicked the response to known fungal elicitors such as *Botrytus spp*. The reverse was also true that by depleting the vacuolar proton pool the elicitor-triggered pH shifts and subsequent alkaloid induction were prevented. It is possible that a common regulator is responsible for the coordinated increases in alkaloid biosynthesis.

Common trans- regulators, ORCA2 and ORCA3, have been found to co-ordinately regulate monoterpenoid indole alkaloid biosynthetic genes in *C. roseus* (Van Der Fits and Memelink, 2000). The ORCAs (Octadecanoid-Responsive *Catharanthus* AP2/ERF-domain) are members of the AP2/ERF-domain family of transcriptional regulators unique to plants. They are characterised by the APETALA2 (AP2) and the Ethylene Responsive Factor (ERF) domains. They were discovered (Menke *et al.*, 1999a) using a yeast one-hybrid screening method and using the terpenoid indole alkaloid gene strictosidine synthases' promoter element, involved in jasmonate –and

elicitor-responsive gene expression (JERE), as bait (Menke *et al.*, 1999b). It has been suggested that these plant regulators could be used to alter secondary metabolic pathways that are responsive to the addition of Methyl Jasmonate including benzylisoquinoline biosynthesis (Memelink *et al.*, 2001).

Overexpression of ORCA2 in *C. roseus* increased transcription of an *Str* promoter: reporter gene fusion (Menke *et al.*, 1999b). A further ORCA gene, ORCA3, was characterised by its overexpression in *C. roseus*, which led to an increase in gene transcript of *Str*, *Tdc* and the downstream products of these genes (van der Fits and Memelink, 2000; 2001). ORCA2 and 3 are both transcriptionally induced by jasmonate (Menke *et al.*, 1999b; van der Fits and Memelink, 2001) and it is thought that the ORCA proteins are activated by methyl jasmonate to increase their own transcription and consequently increase the relevant alkaloid genes (vom Endt *et al.*, 2002).

Transcription factors have been isolated involved in strictosidine synthesis, a critical early enzyme in tryptophan derived monoterpenoid indole alkaloids (Hashimoto and Yamada, 1994) including vinblastine and vincristine in *C. roseus* ((Kutchan *et al.*, 1988; De Luca *et al.*, 1989). Such studies confirm that there are regulatory genes, and possibly signal transduction pathways, controlling alkaloid synthesis. To date there have been no reports of transcriptional regulators interacting with the morphine or general benzylisoquinoline pathway of any species.

Further levels of control have been shown to exist on the pathway. The failure of poppy cell suspensions to reliably produce morphinans, even though codeinone reductase and several other cloned genes from the pathway are present, suggests that differentiation of the cells into laticifers is necessary for morphinan alkaloid accumulation (Roberts *et al.*, 1993).

Expression studies using northern blotting techniques have shown that cyp80b1, coclaurine *N*-methyl transferase, 6' *O*-methyl transferase, 4'-*O* methyl transferase,

berberine bridge enzyme, salutaridine acetyl transferase, and tyrosine decarboxylase are all induced upon wounding or upon addition of elicitor (Facchini and De Luca 1994; Facchini *et al.*, 1996; Park *et al.*, 1999; Facchini and Park, 2003). Several other genes are developmentally controlled such as berberine bridge enzyme and cyp80b1 (Huang and Kutchan, 2000).

Tyrosine decarboxylase message was shown to be most abundant in metaphloem cells of the stem and roots and at low levels in the developing capsules (Facchini and De Luca, 1995). Metaphloem cells are closely associated with laticifers (Nessler *et al.*, 1985). Further to this is the evidence that several key enzymes, salutaridine synthase and salutaridine: NADPH 7-oxidoreductase, show greater activity in stems and roots than in capsules (Gerardy and Zenk 1993a; 1993b). This evidence suggests that the alkaloids are produced in the roots and stem and accumulate in the capsule. Codeinone reductase is relatively unchanged developmentally or with the addition of elicitors to cell suspensions (Huang and Kutchan 2000; Facchini and Park, 2003), however there are spatial distribution differences across plant tissue types (Huang and Kutchan, 2000). Codeinone reductase message levels are highest in the stem leaf and capsule (Huang and Kutchan 2000) suggesting the first part of the morphinan pathway is produced in the stem and roots while the latter synthesis takes place in the stem and capsule.

1.11 The General Study of Alkaloids (A review of methods used)

Many of the modern day drugs were originally isolated from plants producing large enough quantities of the relevant alkaloid to give their pharmacological effect. Several of the less complex alkaloids can be efficiently and cost effectively chemically synthesised for modern use, however many of the more complex alkaloids are difficult or expensive to chemically synthesise. It was hoped plant cell suspensions would enable large-scale factory production of useful alkaloids and other secondary metabolites, however these efforts have been met with limited

success. Very few successful cases have been reported where the yield of the required secondary metabolite was sufficient to offset the costs of maintaining large batch cultures of aseptic plant cell suspensions. Two examples of moderate success were in the production of the antitumour compound paclitaxel, and the antiviral and anti-inflammatory shikonin (Verpoorte *et al.*, 2000). Yields of the alkaloids are often much lower in cell suspensions than in the native plant which also have their limitations subject to seasonal variation and environmental factors. Political pressure can also affect the supply of raw materials for use in the pharmaceuticals industry because of the raw materials being diverted for illicit purposes.

This has led to research targeted at better understanding the pathways involved in alkaloid production. There are several factors that hamper the understanding of many alkaloid pathways and consequently make alkaloid production in cell suspensions difficult to achieve. Most alkaloids have complex pathways with a large number of enzymatic steps. For example the benzylisoquinoline macarpine has 20 enzymatic steps from its precursor tyrosine (Kutchan and Zenk, 1993; Kammerer *et al.*, 1994). This creates a daunting prospect to upregulate all the genes, or find, those genes whose modification will have the greatest effect, for improved alkaloid yield. Even more daunting would be attempting to produce these alkaloids in transgenic microbial cell systems.

One of the first attempts at over expression of a gene in a medicinally important alkaloid pathway was of the gene coding for hyoscyamine-6β-hydroxylase. In one primary transformed plant of *Atropa belladonna* L. increases were found of the end product scopolamine (Yun *et al.*, 1992). This effect was inherited by the progeny to the extent that when mature these plants contained scopolamine as almost the only tropane alkaloid.

Another problem facing efforts to increase alkaloid yields in cell suspensions is that many alkaloids are stored in vacuoles or specialised storage compartments such as the specialised vacuoles in *P. somniferum* latex (Pham and Roberts 1991); the

efficient transportation of these molecules in and out of storage compartments may be a limiting factor.

Many alkaloids are produced in one part of the plant and transported or stored in another part making it difficult to simulate in cell suspensions. Nicotine is possibly the most widely studied of the alkaloids and its role in the plant best defined. It is already well established that nicotine is synthesised in the roots and transported to the leaves (Dawson, 1941; 1942; Wink and Roberts, 1998). Systemic alkaloid induction in *Nicotiana attenuata* has been shown to occur in response to simulated herbivore attack (Baldwin and Ohnmeiss, 1993).

Sanguinarine is produced in the aerial part of the plant and transported to the roots (Facchini and De Luca, 1995) and quinolizidine alkaloids are made in the chloroplast of green cells (Wink *et al.*, 1980; Wink and Hartman, 1982) and yields are increased in the presence of light (Wink *et al.*, 1981; Wink and Hartman, 1980). Further to these problems of alkaloid production in cell suspensions is that many of the desired alkaloids are not always the end product but rather pathway intermediates which do not normally accumulate. For example the indole alkaloid ajmalicine, was shown to be continually synthesised but rapidly degraded in cell cultures of *C. roseus* (Dos Santos *et al.*, 1994).

Gene silencing has been successfully utilised to knock out individual structural genes and their encoded enzymes and eliminate various alkaloid products in the plant. An example is the recent development of caffeine-free coffee plants by targeting the gene for theobromine synthase (Ogita et. al., 2003). There have been several reports of the use of antisense techniques to alter benzophenanthridine alkaloid content. Transformation of *Eschscholzia californica* cell suspensions transformed with the gene encoding either the antisense construct for the berberine bridge enzyme (BBE) or N-methylcoclaurine 3' hydroxylase (CYP80B1) and cell suspensions were found with lowered levels of the respective transcripts. This decreased message led to a

reduced accumulation of benzophenanthridine alkaloids compared to the control cell suspension lines (Park, *et al.*, 2002).

Recently the direct transformation of opium poppy with an antisense construct to the gene encoding the BBE 1 enzyme showed perturbation in the alkaloid content of latex but not in the roots. There was an increase in upstream and side branch alkaloids but no concomitant reduction in the root accumulating sanguinarine, the end product of the BBE branch point enzyme (Frick *et al.*, 2004). These latter results were obtained from only one transgenic event and will need to be repeated and extended. Hairpin RNA induced gene silencing of the penultimate step in morphine synthesis, codeinone reductase, has resulted in the accumulation of (*S*)-reticuline, some seven enzymatic steps back from codeinone reductase (Allen *et al.*, 2004). (*S*)-reticuline is the major benzylisoquinoline branchpoint intermediate. This surprising result illustrates how difficult it is to predict the outcome (alkaloid phenotype) when individual gene activities are enhanced or eliminated.

The inverted repeat knockout systems (or DNA-directed hairpin RNA gene silencing) could also be quite useful in eliminating negative regulators responsible for inhibiting alkaloid biosynthesis. Conversely it may be possible to knock down the expression of genes responsible for controlling the regulators.

Tryptophan decarboxylase (TDC) decarboxylates tryptophan to produce tryptamine and is condensed with secologinin by the enzyme strictosidine synthase (Str) to produce strictosidine. Both genes have been shown to increase biosynthesis (Aerts *et al.*, 1994) via an increase in message levels in response to a fungal infection (Menke *et al.*, 1999a). This process was shown to occur via a signal cascade involving methyl jasmonate and this cascade is widespread in plants (Sembdner and Parthier, 1993; Wasternack and Parthier, 1997) suggesting that an understanding of the regulation of strictosidine accumulation will be of value in other alkaloid pathways (vom Endt *et al.*, 2002).

Master transcription factors responsible for the regulation of secondary metabolic pathways have been found associated with flavonoid synthesis (Weiss, 2000), suggesting that co-ordinated transcription could also be occurring in alkaloid synthesis. This may not be the only level of control, however, as it is likely there are post-transcriptional controls in place to regulate these pathways (vom Endt *et al.*, 2002).

1.12 The Study of Morphinan Alkaloids

One way to look at the morphinan pathway is to use mutants that are deficient in one or more aspects of the pathway. One advantage of using mutants in secondary metabolism is that, unlike mutants in primary metabolism, they are rarely lethal. Mutants deficient in aspects of a biosynthetic pathway have been very powerful tools in dissecting and characterising secondary metabolism. The proanthocyanidin pathway is one of the most well studied pathways in plant secondary metabolism. Initially mutations from petunia, tobacco, maize and barley have helped to isolate several of the genes in the biosynthetic pathway and to reveal several levels of control.

Arabidopsis mutants are now helping to examine the latter stages of the proanthocyanidin pathway, its regulation and transport of the intermediates and end products. By using the transparent testa (*tt*) mutants (Shirley *et al.*, 1995), named for their transparent or proanthocyanidin free seed coats or the tannin-deficient seeds (*tds*) (Abrahams *et al.*, 2003a) it has been possible to find a number of new genes involved in proanthocyanidin pathway. *Arabidopsis* has become the species of choice in which to study proanthocyanidin synthesis, as it appears all but one of the genes involved in the pathway for flavanoid synthesis are encoded by a single gene copy (reviewed in Winkle-Shirley, 2001). Arabidopsis' small genome, fast life cycle, ease of transformation, and a low level of gene duplication has helped with the

discovery of several new genes involved in proanthocyanidin biosynthesis that have been difficult to discover in other species such as maize and barley.

The *Arabidopsis* mutants have helped to isolate several proanthocyanidin structural genes encoding for example: flavanol 3'hydroxylase with *tt*7; flavonoid 3'hydroxyase with *tt*6; dihydroflavanol reductase with *tt*3 (Shirley *et al.*, 1995) and leucoanthocyanidin dioxygenase (LDOX) (Abrahams *et al.*, 2003b). Several genes encoding different classes of transcriptional regulators of the proanthocyanidin pathway have been isolated using *tt* mutants such as: the *tt*2 mutant which encodes an R2R3 MYB regulator, a key determinant in the latter genes of proanthocyanidin accumulation in developing seed (Nesi *et al.*, 2001); a MADS-homeobox containing gene from *tt*16 (Nesi *et al.*, 2002) and WIP from *tt*1, a zinc-finger class of protein different enough to be put into its own subclass. Other examples include: the *ref2* and *ref8* mutants of *Arabidopsis* that have helped to provide a better understanding of the phenylpropanoid pathway and lignin deposition (Hemm *et al.*, 2003; Franke *et al.*, 2002).

A proanthocyanidin specific transporter has also been discovered using *tt*12 that is specific for the vacuolar membrane (Debeaujon *et al.*, 2002). These are several examples of the many genes cloned that are involved in the proanthocyanidin pathway giving new insights into the complexity of this secondary metabolic pathway.

Using the flavanoid mutants of *Arabidopsis tt4*, *tt5*, and *tt3* it was shown that insertion and expression of the maize orthologs could rescue the *Arabidopsis* mutant (Dong *et al.*, 2001) illustrating that secondary metabolic enzymes may not be subjected to such rapid divergence as previously thought.

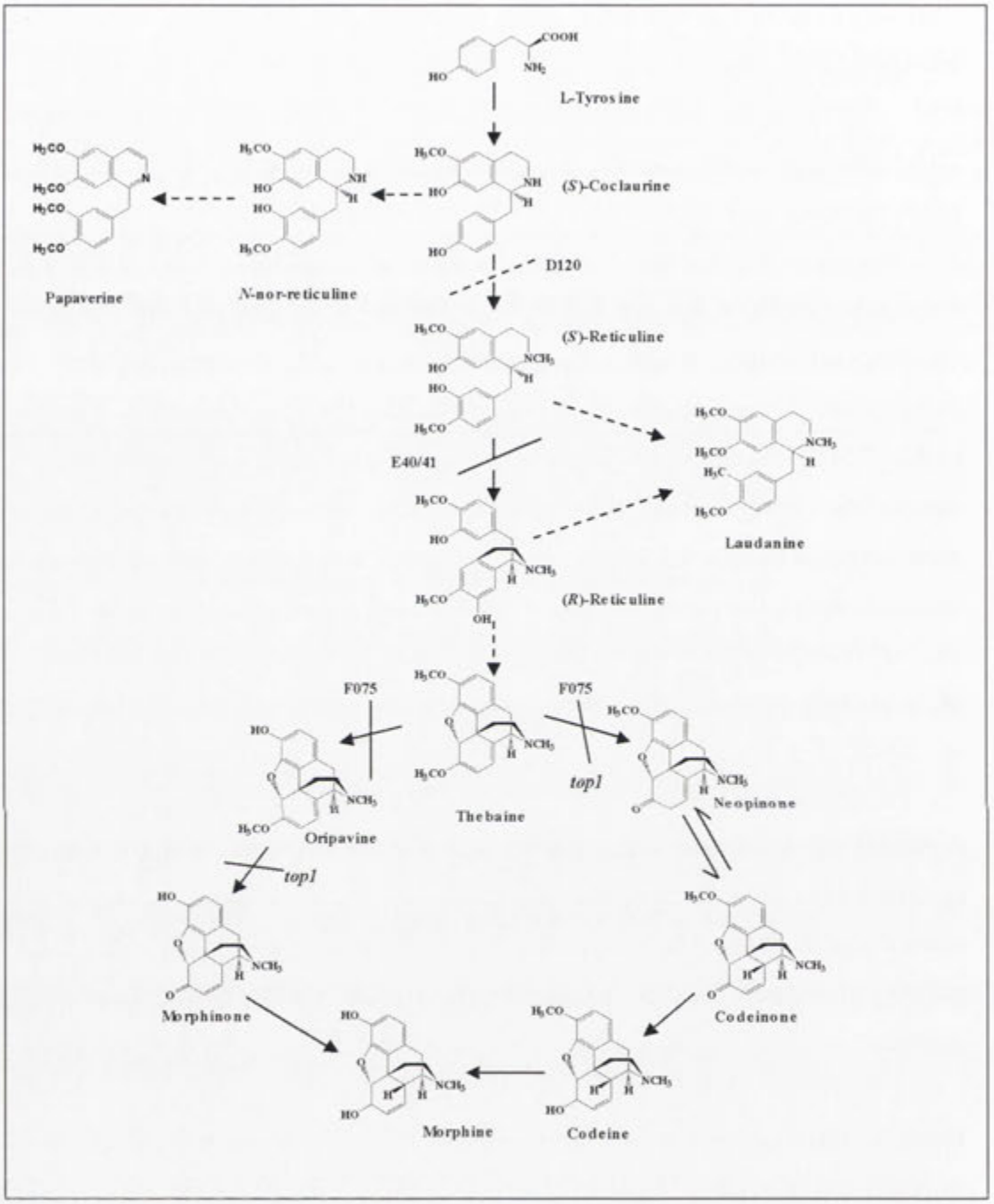


Figure 1.10: An abbreviated alkaloid pathway showing where some of the morphinan mutants, produced by Dr Fist at Tasmanian Alkaloids Pty Ltd., disrupt the pathway. A broken line across the arrow representing the enzymatic step/s represents that the mutant still progresses past this point but the yields are very much reduced. None of these mutants have been characterised and the figure is based on several morphinan alkaloid yields that have been measured on these plants. See

Table 4.1 for a detailed explanation of the mutants' new alkaloid phenotype. (Adapted from Kutchan, 1998).

While there is still much to learn about the proanthocyanidin pathway, the proanthocyanidin mutants in *Arabidopsis* have illustrated how rapidly a complex secondary metabolic pathway can be determined. The morphinan alkaloid pathway has several disadvantages to the proanthocyanidin biosynthetic pathway. Firstly, benzylisoquinoline alkaloids appear not to be synthesised in *Arabidopsis* or any other prospective model species. Secondly, few of the genes isolated from opium poppy have been found in *Arabidopsis* except for possible analogues to the berberine bridge enzyme (The *Arabidopsis* Genome Initiative, 2000). Thirdly, although a robust transformation system is now available for the opium poppy (Park and Facchini, 2000; Chitty *et al.*, 2003) its life cycle is much longer than that of Arabidopsis. Fourthly, there have been few mutants characterised in the benzylisoquinoline alkaloid pathway, in particular the morphinan branch.

Papaver somniferum is a predominately self-pollinating species and has 2n=22 chromosomes. Levy and Milo (1998) studied 8 of the 11 chromosomes of *P. somniferum* and found mainly sub-terminal chromosomes, 2 pair very long, 4 pair of medium length, and 3 pair of short chromosomes. Hybrids between *P. somniferum* (2n=22) and *P. setigerum* (2n=44) suggested that *P. setigerum* had three genomes: A with 11 pair; B with 3 pair; and C with 8 pair. There is still much conjecture as to the relationship of *P. setigerum* and *P. somniferum* however importantly all 11 *somniferum* chromosomes in the *P. somniferum* x *P. setigerum* hybrid paired with 11 *setigerum* chromosomes (presumably the "A" genome). By contrast no *somniferum* chromosomes paired with those of *P. orientale* (2n=28) or *P. bracteatum* (2n=14) when crossed. This is consistent with *somniferum* being a true diploid (x=11). Kadereit and Sytsma (1992) claim evidence from chloroplast DNA supporting the view that *P. setigerum* is just a tetraploid version of *P. somniferum*. *P.setigerum* behaves as a diploid and crosses readily with *P.somniferum* in breeding studies (Fist pers. communication.). Several recessive spontaneous mutants with altered flower

qualities have been identified (Sharma and Singh, 1983). Khanna and Singh (1975) have used gamma rays on opium poppy seeds and found several M1 mutants with different characteristics such as: male sterility, opium less plants, high morphine and increased capsule number. Nigam *et al.*, (1990) also used gamma rays to find dwarf and early flowering varieties. Using a combination of EMS and gamma ray treatment seemed to enhance the mutation frequency (Patra and Chauhan, 1990; Chauhan and Patra, 1993).

Very few of these mutations have been utilised to further the understanding of the alkaloid pathway, and at least some of them appear to have been unstable. Polyploid *P. somniferum* plants have been reported to have higher morphine content and capsule number than the parent from which they were obtained (Levy and Milo, 1998), however late seed setting and poor seed yield were major shortcomings for potential commercial production.

Nyman and Hall (1976) described a spontaneous mutant in poppy accumulating thebaine as the major alkaloid instead of morphine, however this phenotype was unstable and true-breeding lines of the mutant could not be obtained (Nyman, 1978; 1980). A non-narcotic variety of opium poppy named Sujata has been obtained using EMS (Sharma *et al.*, 1999a,b). This mutant is intended for cultivation for poppy seed oil and culinary seed, not for alkaloid production. Little genetic or biochemical characterisation of this mutant has been reported. An Indian landrace mutant, *aco*, has been reported which form androcarpels in place of inner whorl stamens in their flowers (Prajapati *et al.*, 2001). Limited analysis suggested that this mutant accumulated more morphinan alkaloids in the capsule due to the increase in androcarpel wall mass, which accumulates alkaloids.

A series of stable mutations in the morphine pathway have now been identified by Dr. Fist at Tasmanian Alkaloids, which hold great promise as tools to explore the regulation of the pathway (Figure 1.10). The present study employs traditional

characterisation techniques of metabolite analysis to analyse these morphinan alkaloid mutants.

In this thesis, particular attention is given to *top1*, a mutation in which morphine and codeine are entirely replaced by thebaine and oripavine (Fist *et al.*, 2000; Millgate *et al.*, 2004). Oripavine and thebaine have no real value directly as pharmaceuticals since their adverse affects outweigh any analgesic properties. Indeed thebaine induces vomiting and oripavine causes convulsions and death when fed to mice (Gomez-Serranillos *et al.*, 1998). The *top1* mutant was found to have the same high yielding alkaloid characteristics, as the progenitor parent genotype and alkaloid yields have subsequently been further enhanced through breeding. The *top1* mutation in homozygous condition has been utilised in commercial production in Tasmania in a number of released cultivars with the harvested thebaine and oripavine being used as feedstocks for many of the newer pharmaceuticals. This plant constituted almost 50 % of the total *P. somniferum* crop grown in Tasmania in 2002 (International Narcotics Control Board 2003).

It is anticipated that the transcription of various genes associated with morphine accumulation will be affected in these mutants. It is further anticipated that the identity of the genes differentially expressed in one or more mutants, will provide a better understanding of the control of morphine biosynthesis and accumulation.

This study will take advantage of these previously unstudied *P. somniferum* morphinan mutants. All the mutants and *top1* in particular will have their transcriptional changes logged and compared. The genes responsible will be assessed for whether they play a role in the regulatory or transport processes controlling morphine biosynthesis. Using homology to known sequences from other species in the database roles can be inferred. This is dependent upon there being some functional characterisation of these genes, which have been annotated in other species. However it still doesn't rule out the possibility that these genes may have more than one function which is yet to be described or discovered or that these genes

may be different enough from the genes described in the database as to have a entirely unrelated function.

Many questions remain regarding the synthesis and accumulation of morphinan alkaloids. It is still not known where the morphinan alkaloids are synthesised. The morphinan alkaloids accumulate in specialised latex vacuoles but how do they get there and what form of transport mechanisms moves them to the appropriate compartment for the next catalytic step to take place. For example thebaine is known to accumulate in the specialised vacuoles but the next catalytic step is a demethylation, which most likely involves a cytochrome P450 enzyme, which is most likely closely associated with the endoplasmic reticulum; the products of the demethylation again are known to accumulate in the vacuole as codeinone or oripavine.

1.13 Methods of Transcriptional Profiling

There are several traditional ways to map the transcriptional profile of a pathway, such as Northern blotting, differential display-PCR (DD-PCR), serial analysis of gene expression (SAGE) and cDNA amplified fragment length polymorphisms (cDNA-AFLPs). There are positives and negatives for each type of the techniques. Northerns are still considered the benchmark of transcription profiling as they are still used to authenticate real time PCR and microarray results, however they are laborious and time consuming making it impractical to screen large numbers of unknown genes for differences in transcriptional profiles.

cDNA-AFLPs use the power of PCR-based markers for the rapid screening of genetic diversity (Vos *et al.*, 1995) and have been used successfully to isolate a number of genes in non-model organisms. The use of PCR allows this technique to be highly sensitive. This method has been used with some success to analyse the nicotine biosynthesis pathway in tobacco (Goossens *et al.*, 2003). Their limitation

however is that a large number of PCR reactions need to be followed by restriction enzyme analyses which need to be run on agarose gels. This is time consuming. Retrieval and identification of the correct DNA band can also be difficult and time consuming. If the mutation is one that is in the promoter region of a gene or affects the subsequent processing of the expressed gene, then cDNA-AFLPs will be uninformative of the possible gene lesion. cDNA-AFLPs also have the added difficulty in identifying homologous markers (alleles) hindering heterozygosity analyses (Mueller and Wolfenbarger 1999) and fail to give quantitative data.

DD-PCR enables the researcher to identify differentially expressed genes without prior knowledge of sequence and availability of cDNA clones by arbitrary amplification, labelling with radioactive probes and comparison of different mRNA sources. Modifications to this system have enabled researchers to amplify and clone products over 2 kb in size (Diachenko *et al.*, 1996). Again being a PCR based approach it can also be highly sensitive. The inability of the system to quantify the expression differences is DD-PCRs major shortcoming as well as similar issues to cDNA-AFLP analysis of running gels of many reactions and identifying the correct band.

SAGE involves the use of very short cDNA sequence tags, generally 9-14 bps in length. Using these sequence tags as a unique identifier for transcripts can give a quantitative estimate of the expression level of many transcripts. The identity of the gene corresponding to the gene tag can only be inferred if the gene is already annotated revealing a major limitation of this method of analysis especially for a non-model species such as poppy. Not many of the genes responsible for alkaloid synthesis are available in the database at the present time and regulatory sequences of morphinan alkaloid synthesis are yet to be cloned limiting the transcriptional study of the control of the pathway using these methods.

The use of cDNA microarrays has become an important tool in the understanding of global gene expression. The advantages of this system are its ability to spot several

thousand cDNAs from a particular tissue type and to quantitatively compare transcript expression levels of a test and reference sample over them. This technique uses fluorescently labelled reverse transcribed cDNAs rather than radioactive labelled probes. The DNA sequence is not needed to produce primer combinations or DNA probes. The cDNAs that are differentially expressed between the test and reference samples can be sequenced and further annotated.

There are several problems associated with using cDNA microarrays as a tool for gene discovery. Firstly, using cDNA microarray analysis to analyse mutants is reliant upon the mutation causing a change in transcript levels. The mutation could be in a structural enzyme, which affects the enzyme catalytic activity or translational process, but not the gene's expression, making it unlikely that the mutation will be seen using cDNA microarray analysis. Secondly the mutation may affect the post-transcriptional modification, for example causing alternative splicing, which unless specific parts of an expressed clone are placed on the array would not be detected, as similar amounts of transcript would still be present to bind to the array. Thirdly, it is possible that the RNA transcript has a different half-life in a particular tissue and unless material is used from the correct stage and tissue the mutant will not be detected using cDNA microarrays. Microarray technology is also less sensitive than the more specific PCR based approaches and can have cross-hybridisation issues with competition between analogues clones or sequences. The other major drawback with microarrays, especially with ones that do not contain the whole genome, is that not all of the genes/clones of interest are represented on the array.

Microarrays have been shown to be powerful tools in the mass analysis of genome wide transcript profiling and have been widely utilized in model systems such as *Arabidopsis* and rice to help understand plant stress responses such as: cold (Seki *et al.*, 2001; Breton *et al.*, 2003); wounding (Strassner *et al.*, 2002); and pathogen attack (Schenk *et al.*, 2000). The sequencing of the *Arabidopsis* genome has brought with it rapid developments in the use of microarrays as gene analysis on sequenced EST clones is much more informative than if undertaken on unknown clones. Each

subsequent microarray analysis adds more information about the 'unknown', but sequenced gene, allowing possible mode/s of function to be ascertained. Interesting genes can be further analysed through the use of reverse genetics and either partially or totally knocking out the candidate gene to create a phenotype lacking the genes' function/s.

Study of a mutant's phenotype can often be limited using classical characterisation techniques due to the complex nature of the mutation, pleiotropic effects or no obvious phenotype. Study of the transcriptome of mutants enables the researcher to analyse what genes are being affected by the mutation. Already several mutants have been studied using microarrays in *Arabidopsis* allowing the investigators to postulate the gene's role in the plant and relation to several metabolic pathways. In particular microarrays give the sensitivity and a broad net to contribute to the understanding of the mutated function. By using mutants in *Arabidopsis* it was possible to illustrate the *knotted1*-like homeobox encoded protein, BREVIPEDICELLUS (BP), was responsible for the regulation of lignin biosynthesis (Mele *et al.*, 2003) giving a new insight into this genes' function that was previously unknown, as the phenotypic effects were too varied to classify a direct function. This enabled the authors to concentrate their efforts to show that the BP protein binds promoters of several genes involved in the lignin pathway.

It has also been shown that many genes regulated in parallel can share common *cis*-acting regulatory elements for transcription (Harmer *et al.*, 2000; Gross *et al.*, 2000). Even if the gene or genes directly responsible for the mutant phenotypes in *P. somniferum* used in this study are not revealed through microarray analysis it should reveal a class of genes associated with the mutation and possibly co-regulated with the morphinan pathway.

1.14 Aims of the Present Study

Understanding the morphinan alkaloid pathway is important in a commercial aspect, it's biochemical and scientific aspects are equally if not more so important. Several questions arise when considering the alkaloid pathway. Why does the plant invest resources from primary metabolism to these secondary metabolic pathways? Why does the plant invest even more resources and have many levels of control on the production of these secondary metabolites? For example the opium poppy and many other plants have a laticifer system to store these compounds requiring even greater investment of the plants resources. By studying the morphinan pathway this will help further our understanding of the complexities of secondary metabolic pathways and possibly further our understanding the benefits that these pathways confer upon the plant. While this study does not hope to answer such large and involved questions in full it is hoped that some small steps can be made in determining some of the genes that are associated with the morphinan pathway, leading to further avenues of study of the morphinan pathway, its control and regulation.

This thesis is a study of the morphinan alkaloid pathway in opium poppy and involves the production of a 17,000 cDNA clone micrarray and global analyses of transcriptional differences between the normal poppy and various mutations and other genotypes in which the morphine pathway is disturbed. Germplasm to be utilised includes: the mutants deficient in alkaloid synthesis produced at Tasmanian Alkaloids Pty Ltd. (See Figure 1.10 for where the mutations occur); a genetically transformed *P. somniferum* containing an RNAi knockout construct to the codeinone reductase gene; a low alkaloid cultivar *P. somniferum* cv. Marianne; and a different *Papaveraceae* species *E. californica*. The different germplasms can be separated into three groups: 1) those that accumulate at the major branchpoint of (*S*)- reticuline including one mutant (E40/41), the hairpin RNA codeinone reductase transgenic (hpCOR) and *E. californica*; 2) those that accumulate at the next major branchpoint, thebaine, in the morphinan pathway, (including the mutants F075 and *top1*); and 3)

those that have reduced morphinan alkaloid yields (including D120 and the *P. somniferum* cv. Marianne).

Not all of the genes for morphinan alkaloid biosynthesis have been cloned, making it difficult to determine the position of mutations in the pathway. Little is known of many of the morphinan genes' level of expression and where the morphinan alkaloids are produced or transported in the plant. It is often difficult to predict whether the mutation is in: a regulatory pathway; transport process; structural changes in cells or particular cell types; or in the gene coding for the enzyme immediately after the blockage itself. By looking at the phenotype at the molecular level rather than just the classical mutant characterisation it is hoped to create a greater understanding of the morphinan alkaloid pathway.

It is hoped that a global analysis of differential gene expression in *P. somniferum* mutants, with altered morphinan phenotypes, will deliver a short list of candidate genes for having roles in regulation, metabolism and transport of alkaloid intermediates. The shortlisted genes will then qualify for further characterisation both within this present study and beyond. It is anticipated that the mutated genes themselves may not be transcriptionally affected; this would be the case if the mutation encoded a non-functional protein but the transcript level was unchanged. However the lack of a functional protein will have other transcriptional consequences in the pathway, including through feedback controls. It is also possible that the gene corresponding to the mutation is not even printed on the slide, but again the secondary transcriptional affects of the mutation should be evident. It is these secondary transcriptional changes, which are anticipated to provide some new insights into the alkaloid pathway.

The steps involved in attaining these goals can be stated as a series of aims:

1. To establish a *P. somniferum* cDNA library and use this to develop microarray slides.

2. To characterise several different genotypes with disturbed morphine alkaloid biosynthesis characteristics.
3. To undertake a global transcriptional analysis of this germplasm using the cDNA microarrays.
4. To compare differentially expressed clones between the different genotypes, sequence them and use BLAST analysis to determine potential functions.
5. To reduce the number of genes to a short-list of candidates most likely to be involved in or interact with the morphine pathway; a short-list warranting further study.

Chapter 2 contains the methods used in this thesis. Chapter 3 concerns the classical characterisation (where it is original work) of each of the different genotypes being presented. Chapter 4 examines the development of the microarrays and the results of the microarray analyses. A set of differential genes is characterised further in Chapter 5 and a short-list of genes established warranting further investigation beyond the scope of this thesis.

If successful, this work will be the first cDNA microarray study of *P. somniferum* and the first molecular study of mutations affecting morphine biosynthesis. It is hoped that new genes will be uncovered which will expand our understanding of the pathway, especially into areas of regulation and transport, which currently are very poorly understood.

Chapter 2

Methods

2.1 Cultivation of Plants

Tasmanian Alkaloids Pty Ltd isolated the morphinan alkaloid mutant lines used in this thesis. The M2 of EMS treated seed of the Tasmanian cultivar C052 were grown to between the 10-leaf stage and the "running up" stage (Figure 1.7) prior to flowering.

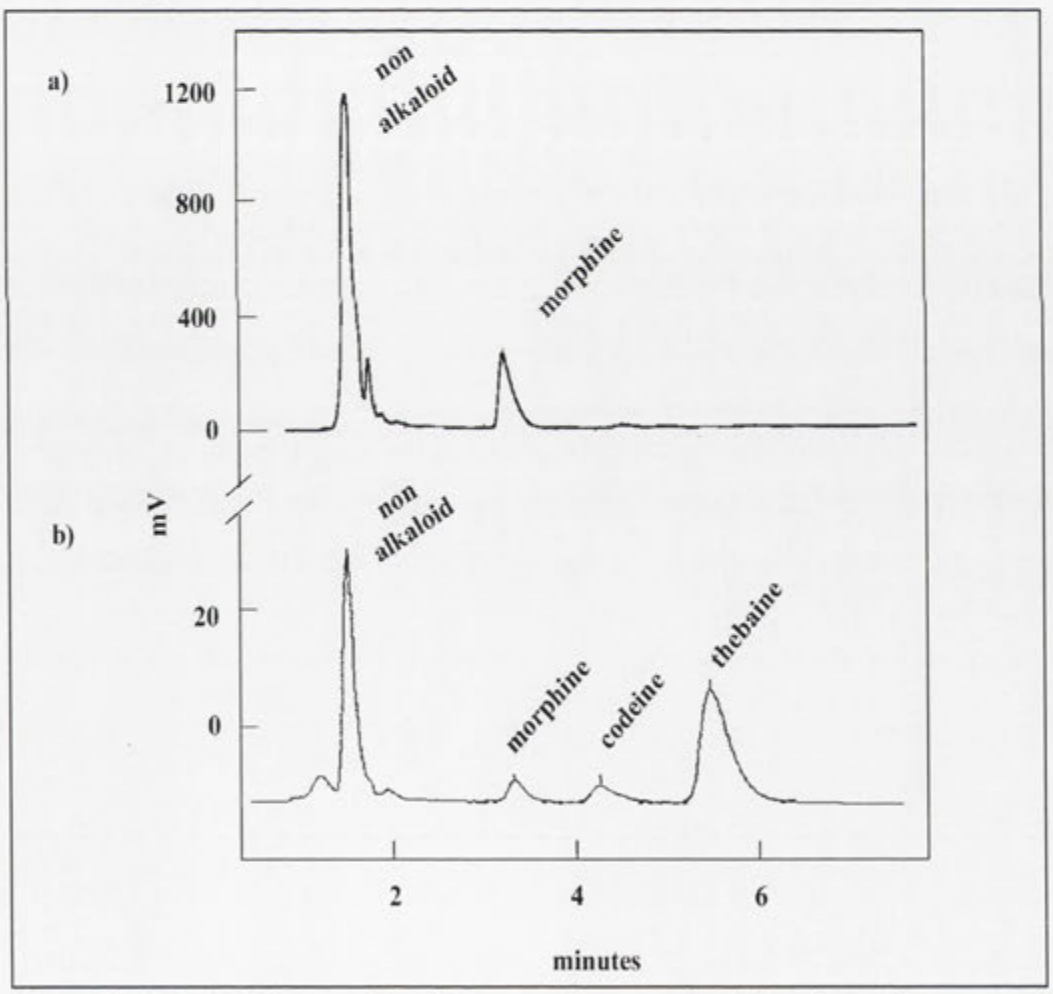


Figure 2.1: a) HPLC trace file of *P. somniferum* at the 6-8 leaf running up stage of development compared to the HPLC trace file **b)** of the immature (2-4 leaf) *P. somniferum*.

This stage is prior to morphinan accumulation in the capsule as can be seen by Figure 2.1, which shows the HPLC trace files of the difference in the levels of accumulation of the morphinan alkaloids of young plants (from 2-4 leaf stage) compared to the older 10 leaf stage. Screening of the EMS generated M2 mutant population was undertaken by solvent extraction of latex collected from leaf, and HPLC analysis.

The elite *P. somniferum* line used for the production of the cDNA library is the parent line C052 provided by Tasmanian Alkaloids Pty Ltd. The control plant was used rather than the mutant to produce the cDNA library for slide printing due to the likelihood that the mutation has caused a down regulation in the pathway. It was preferable that the down-regulated genes be represented on the microarray.

Plant material was germinated and grown under glasshouse conditions of 16 hr days at 25 °C and 8 hr nights at 18 °C. The potting mix contained Scott's Osmocote® Plus controlled release fertilizer at the company's recommended dose.

Plant material for the microarrays was grown in the glasshouse under the same conditions as the C052 elite cultivar used for the production of the *Papaver somniferum* cDNA library. Leaf material was harvested at the running up or prehook stage (See Figure 1.7) just before flowering and RNA prepared as Chapter 2.3.

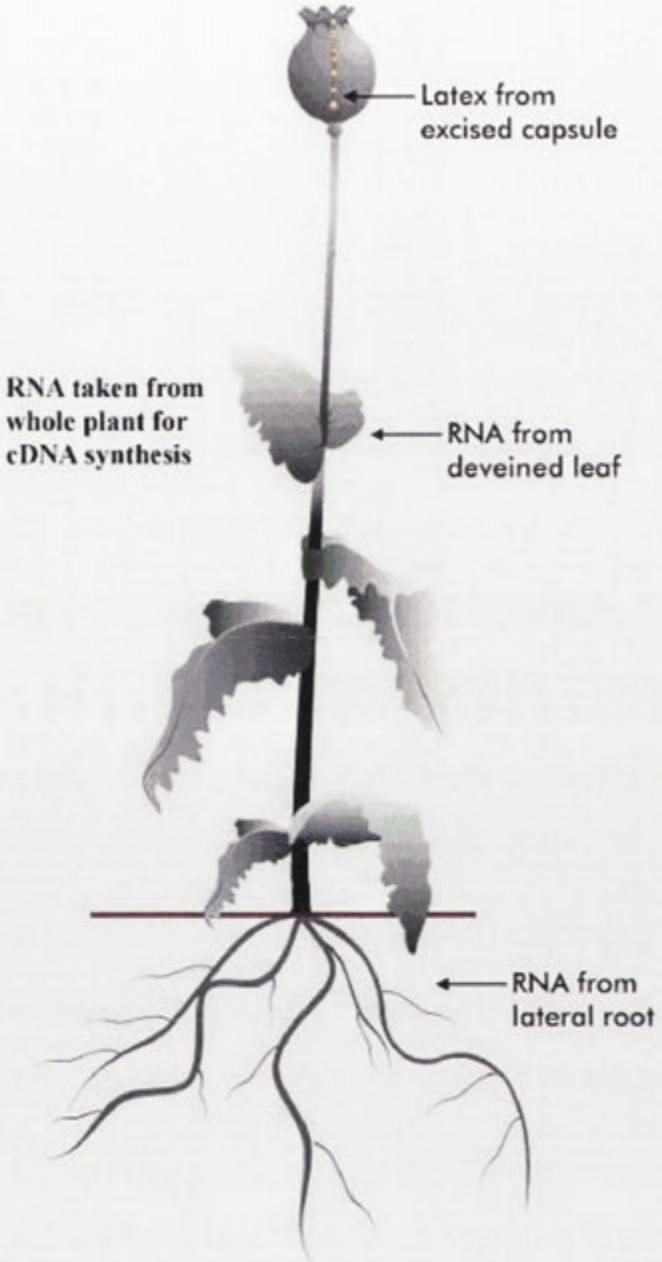


Figure 2.2: Diagramatic representation of a *P.somniferum* and the regions RNA was taken for RNA analysis.

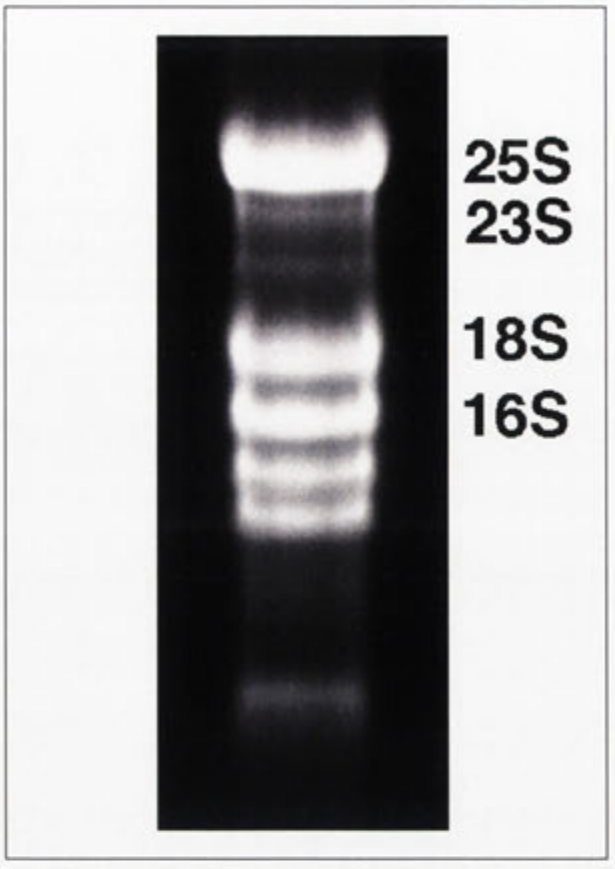


Figure 2.3: Picture showing the quality of *P. somniferum* total RNA run on a formaldehyde gel and used for Poly A^+ mRNA extraction.

2.2 RNA Extraction

RNA was isolated from plant material grown to the running up or prehook stage, just prior to flowering and capsule formation of the poppy (Figure 2.2), using the scaled up method of NTES buffer, phenol/chloroform extraction and LiCl precipitation (Higgins et. al., 1976).

The prehook stage was chosen because this is when the alkaloid production and accumulation is maximal in the poppy plant (Williams and Ellis, 1989). 8 gm of whole plant (excluding roots) from six separate plants was harvested and immediately placed in liquid Nitrogen and ground with a mortar and pestle yielding good quality RNA. The *P. somniferum* Poly A+ RNA (100 μg at a concentration of 20 μg/mL with an A260/280 of 1.9) was isolated using the

Promega PolyATtract mRNA isolation System (Catalogue #Z5210). To check the quality of the RNA and Poly A+ mRNA they were electrophoresed on gels consisting of 0.66 M formaldehyde mixed with 1 x MOPS buffer (0.1 M 3-(*N*-morphilino) propane sulphonic acid (pH 7.0), 40 mM sodium acetate and 5mM EDTA (pH 8.0)) and 1.32 % agarose as described in Sambrook et. al., (1989). 10 µg of total RNA was incubated for 5 minutes at 65 °C in a total volume of 20 µL containing 40 % formamide (deionized), 15 % formaldehyde, 1 X MOPS buffer (Sambrook et. al., 1989) and1 mg/mL Ethidium Bromide. Gels were visualised using a Gel Documentation 2000 system from Bio-Rad (Figure 2.3).

2.3 cDNA Synthesis

The *Papaver somniferum* plasmid cDNA library was provided in the pCMV•Sport6 vector (Figure 2.4) transformed into EMDH10B cells by Life Technologies Catalogue No. 11138-013. The library yielded 2.63 x 10^8 cfu with an average insert size of 1.82 kb and 96% of the colonies having an insert ≥200 bp. To minimize representational biases from occurring during the expansion of plasmid cDNA libraries (Kriegler 1990) semi-solid amplification of the primary cDNA transformants was undertaken. The Life Technologies modified protocol was used to amplify the cDNA transformants at 30°C to help stabilise any unstable clones (Hanahan et. al., 1991). The cDNA library redundancy was calculated by sequencing 1056 clones at AGRF (see section 2.8) and aligning each sequenced cDNA clone to the other 1055 sequenced clones. Thus the percentage of singletons was very high at 98.5%. This gave 16 redundant clones of which there were eleven groups (Appendix 1). Using these figures an approximate percentage of the redundancy of the library could be ascertained (Huang and Weir 2001).

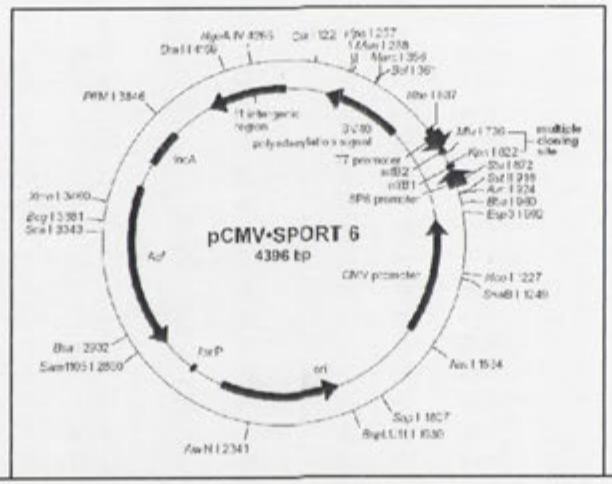
$$\% \text{ Redundancy} = \frac{\text{Total number of clones - (Singletons + redundant groups)}}{\text{Total number of clones}}$$

$$\% \text{ Redundancy} = \frac{1054-(1038 + 11)}{1054} \times 100 = 0.005\%$$

2.4 Microarray Printing

The PCR amplification of a total of 17,000 *P. somniferum* cDNAs (AB Gene Catalogue #AB-0600) was undertaken in 96 well PCR plates by adding 47 µL of a PCR cocktail to each well and adding the 3 µL of the overnight LB culture (see above). The PCR cocktail contained PCR Buffer (500 mM Tris/HCL pH 9.2, 160 mM $(NH_4)_2SO_4$, 22.5 mM $MgCl_2$), 25 ng of each primer (SP6 and T7 (Figure 2.5) each at 100 ng/µL made by Sigma), 0.2 mmoles of dNTPs (Promega Catalogue #U1240) and 1 unit of Taq DNA Polymerase in a final volume of 50 µL. The PCR plates were sealed (AB Gene Clear PCR Plate seals Catalogue #AB-0580) then placed into a Corbett PCR thermocycler (PC-960C) under the following conditions: 5 minutes at 94 ºC, 35 cycles of 30 seconds at 94 ºC, 30 seconds at 48 ºC and 1 minute 30 seconds at 68 ºC, then 1 cycle of 94 ºC for 30 seconds, 48 ºC for 30 seconds, 72 ºC for 10 minutes before a final step at 20 ºC. 1µL of loading dye (from a 6 X stock: 15% Ficoll® 400, 0.03% bromophenol blue, 0.03% xylene cyanol FF, 10 mM Tris-HCl (pH 7.5) and 50 mM EDTA) was added to 5 µL of the PCR reaction and run on a 1% agarose and 0.5 X TBE gel in 0.5 X TBE buffer at 120 mV and 0.8 Amps to check for insert size and amplification. This quality control was undertaken on all of the 96 samples of each 96 well plate. Any plates where there were greater than 5 PCR reactions

a)



b)

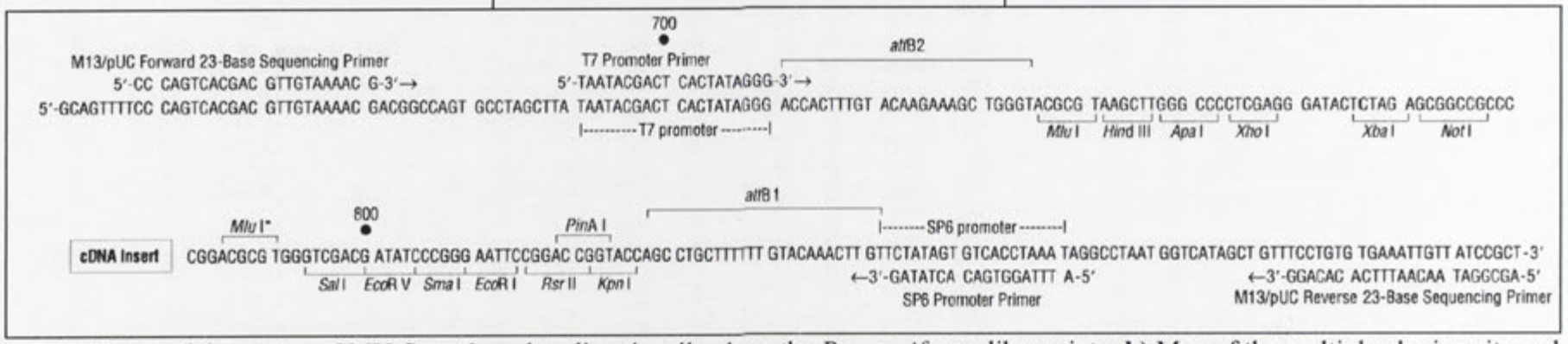


Figure 2.4: a) Map of the vector pCMV•Sport6 used to directionally clone the *P. somniferum* library into. **b)** Map of the multiple cloning site and primer binding regions (641-917) of pCMV•Sport6. The sequence listed here is the negative strand. *This Mlu I restriction site is contained within the Sal I adaptor introduced into the vector upon ligation of the cDNA insert. Adapted from Life Technologies.

SP6-	5'TATTTAGGTGACACTATAG 3'
T7-	5'TAATACGACTCACTATAGGG 3'

Figure 2.5: PCR primers used for the amplification of the *P. somniferum* cDNA inserts contained within the pCMV•Sport 6 vector.

combined, which either did not amplify or gave more than one band were either repeated or replaced with clones that did give one strong band from a new PCR reaction. This was undertaken so that the degree of multiple bands (contamination by more than one cDNA clone) on the array was minimised. This gave confidence that the spots defined as being statistically differentially expressed in the experiment would be single clones and easier to analyse later.

Once plates were confirmed as having amplified products (Figure 2.6) the contents of each well was precipitated with 2.2 volumes of 95% drum EtOH (not 100 % as it contains fluorescent impurities) and 0.1 volumes of 3M sodium acetate, pH 5.2 mixed well by pippeting up and down. The resealed plates were placed at –20 °C for at least an hour before spinning in a Sorvall RC-5B refrigerated bench top centrifuge with a SH-3000 rotor at 4000g for 1 hour then decanting the ethanol acetate solution. The precipitates were washed with 80 % ethanol (also made from 95 % drum ethanol) and centrifuged at 4000g for 45 minutes before carefully decanting the ethanol wash. The precipitates were dried of residual ethanol by spinning the plates at 100 rpm for 1 minute upside down in the Sorvall centrifuge with paper towel cut to size underneath. The plates were allowed to air dry before adding 8 µL of 50% DMSO (giving a DNA concentration of 0.15 to 0.20 µg/µL), resealing and storing at 4 °C overnight to allow resuspension of the DNA into the DMSO solution. Again, quality control was undertaken by checking the first row of each precipitated 96 well plate, this was done by running 0.5 µL of each of the resuspended DNA samples on a 1% TBE agarose gel containing Ethidium bromide as above (Figure 2.6).

To facilitate slide printing the samples were then transferred to 384 well plates (Whatman

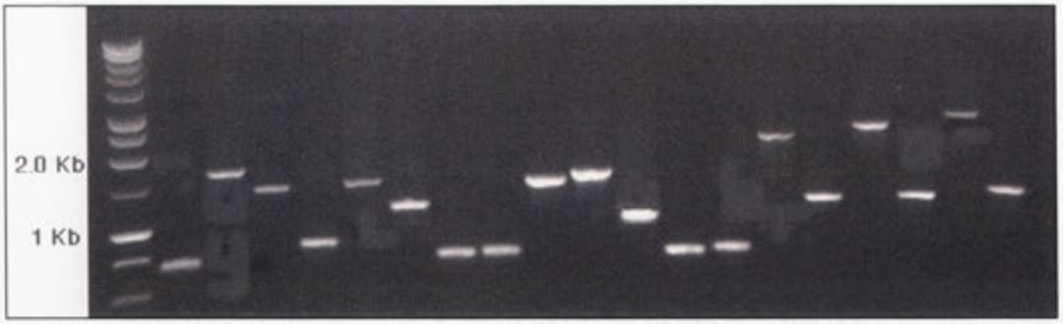


Figure 2.6: 1% Agarose gel run on 0.5X TBE showing amplification of cDNA clones and the range of sizes of the clones

384 well 80 µL polypropylene v-bottom Uniplate Catalogue # 7701-5101) using a multi-channel pipette and the order of the 96 well plates recorded.

The 384 well plates were then sealed using the adhesive plate seal and stored at 4 °C until printing onto glass microarray slides (Corning CMT-GAP2 coated slides Catalogue #40004, non-bar coded) with the microarrayer (Bio-Rad VersArray Chip Writer Pro.) at CSIRO Plant Industry with reservoir pins 100-120 µm in size. After the slides had been printed they were baked at 80 °C for 4 hours to crosslink them and stored dry and in the dark at room temperature.

2.5 RNA Labelling and Hybridisation of Arrays

RNA was extracted from plant material at the same developmental stage (unless otherwise stated in the appropriate chapter) using the same RNA extraction procedure as per Chapter 2.3. RNA quality was checked by running on a MOPs formaldehyde gel as per the methods described in Chapter 2.3. (For a diagrammatic representation of the RNA labelling, hybridisation and analysis see Figure 2.7). 100 µg of test (different morphinan accumulating germplasm) and reference *P. somniferum* (CO52) RNA was labelled, ethanol precipitated and taken up in 10 µL. 4 µL of this was then used for each slide hybridisation.

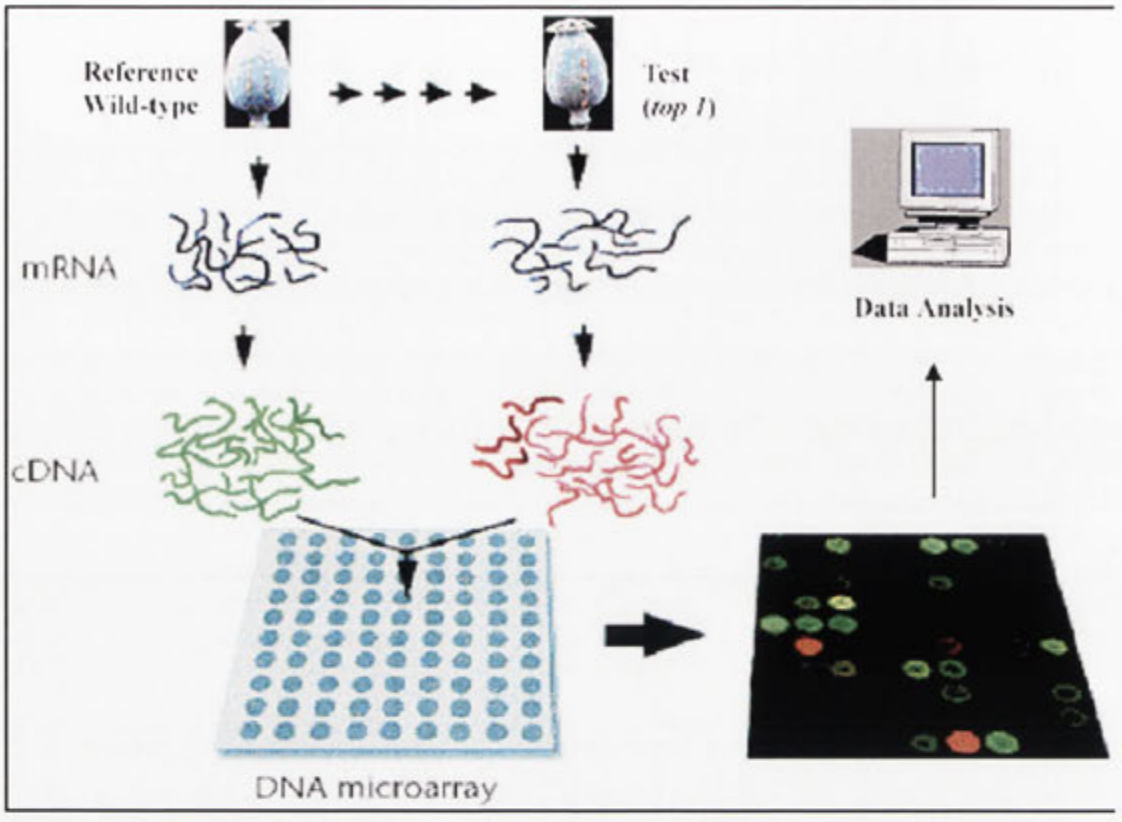


Figure 2.7: Diagrammatic representation of the RNA labelling, hybridisation and analysis used for the *P. somniferum* experiments. (Adapted from Brown and Botstein 1999).

Probes for the microarray slides were labelled using a two-step method in the incorporation of Cy3/Cy5 (Amersham Pharmacia FluoroLink™Cy3/Cy5-dUTP Catalogue #PA 53022 and 55022) dyes by random priming (Schenk et. al., 2000). The first strand synthesis reaction was initially heated at 70 °C to anneal 100 µg of either the test or reference RNA and oligo dT (23 mer dT with C/G/A at 3' end made by Sigma) at a final concentration of 2.5 µg/µL before transferring to ice. The reaction was then continued at 42 °C for 1 hour after adding 2 x 10^5 Units of Superscript II Reverse transcriptase (200 U/mL) (Gibco BRL catalogue # 18064-014), 1st strand synthesis buffer (1x supplied with Superscript), 0.01M DTT, and 0.5mM of dNTP (final concentration at 0.5 mM of each dNTP) mix. RNase H (1 unit Promega Catalogue # M4281) was added and the cDNA-RNA mix incubated at 37 °C for 30 minutes. The reaction was purified using Amicon Microcon YM30 filters (Catalogue # 42410) and eluted to a final volume of 8 µL with TE (pH 8.2) buffer. 1 µL of the elution of unlabelled cDNA was taken and diluted

1/20 with TE (pH 8.2). 1 µL was used in a PCR reaction to check the quality of the cDNA synthesis reaction. Primers were used to the 5' and middle region of the codeinone reductase gene for amplification of the expected band size. This would enable an mRNA and cDNA synthesis quality control point that the 3' to 5' directional synthesis was working efficiently.

The Cy3 or Cy5 labelling reaction of the cDNA involved the annealing of random primers to the cDNA. The 40 µL reaction initially contained 4 µL of the cDNA reaction, 6 µg of Random Primers (hexanucleotides Gibco Catalogue # 48190) and 1X Fill in Buffer (10X provided with Klenow Polymerase) heated to 95 °C for 2 minutes then let cool at room temperature. A dNTP mix was added (to a final concentration of 0.025 mM each except dTTP which was at 0.009 mM (Promega)), then 10 Units of Klenow polymerase (USB Catalogue # 70057Z) and 25 nmol of either Cy3 dUTP or Cy5 dUTP (to either test or reference sample cDNA) and the reaction carried out in darkness at 37 °C for 3 hours. The reaction was again purified using Amicon filters and dried down in a SpeedVac. The probe was taken up in 10 µL of TE (pH 8.2) buffer and the total reaction used for the following hybridisation procedure.

For the hybridisation, 25 % formamide, 5 X SSC, 0.1 % SDS, and 30 µg of Salmon Sperm DNA were added to the probe to a final volume of 45 µL. The probe solution was heated to 95 °C for 5 minutes and spun to cool before adding to a prehybridised slide. Slides were prehybridised by incubating at 42 °C for 45 minutes in 25 % formamide, 5 X SSC, 0.1 % SDS and 0.1 mg/mL Bovine Serine Albumin (BSA) as per the CMT-GAPS Coated slides protocol for DMSO printings. Slides were rinsed with distilled water and dried by spinning in a slide spinner. Once the probe was added to the slide, a coverslip (Sigma Hybri-slips 22X 40mm) was added and the slide placed into a hybridisation chamber and incubated at 42 °C for 16 hours. The slide chamber was disassembled and the coverslip removed by washing the slide at 42 °C in 2 X SSC and 0.1 % SDS before the placing the slide in fresh 2 X SSC and 0.1 % SDS at 42 °C for 5 minutes. This was followed by a 10 minute room temperature wash of 0.1 X SSC/ 0.1 % SDS and 4 room temperature washes of 0.1 X SSC each for 1

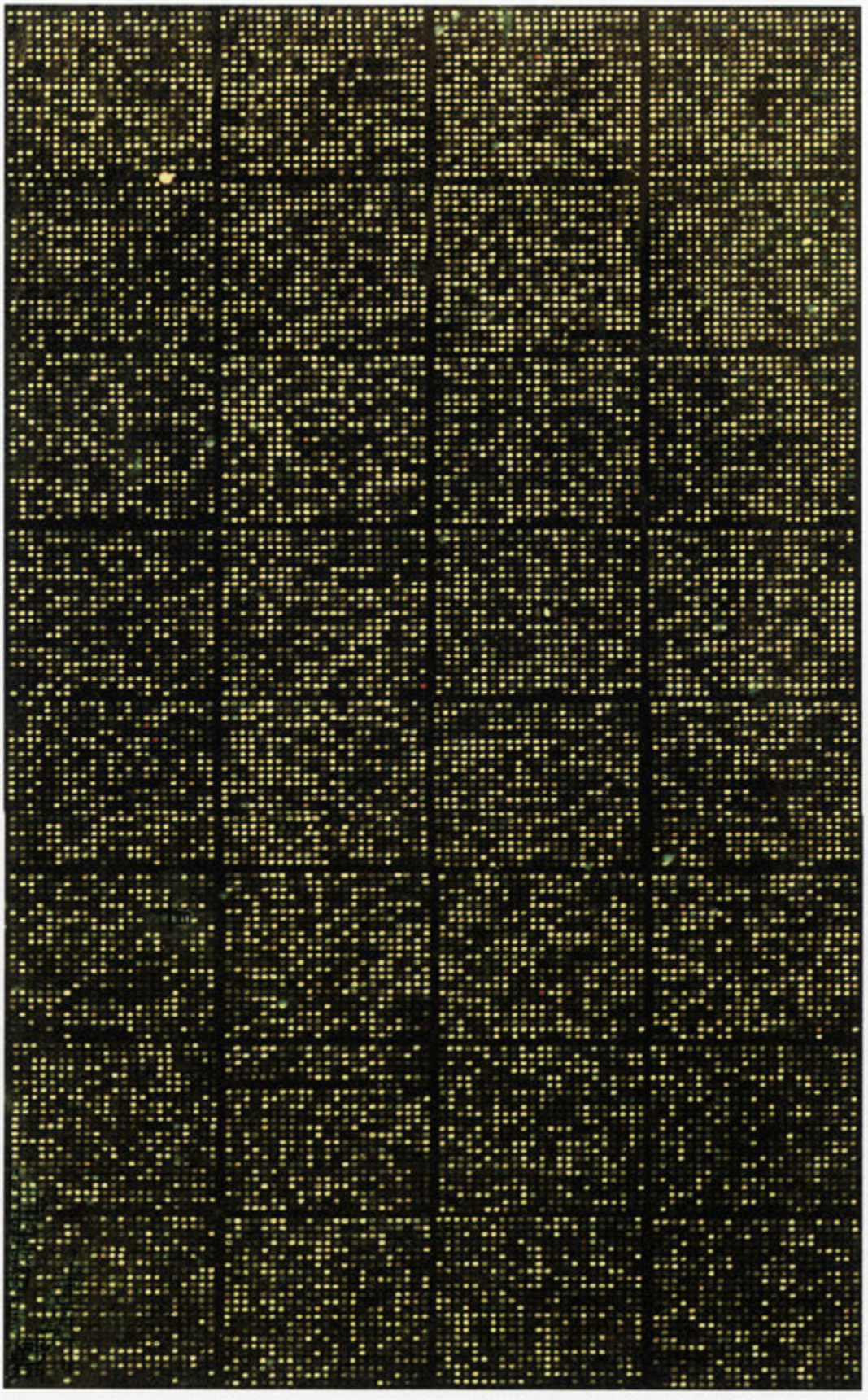


Figure 2.8 (Previous Page): A typical microarray slide from the experiments performed on the 17,000 *P. somniferum* cDNA library.

minute.

Slides were then rinsed for 10 seconds in distilled water and rinsed with EtOH (95 % drum ethanol) before spinning upside down in a centrifuge to dry.

Slides were scanned (Figure 2.8) with a GenePix 4000A microarray scanner (Axon Instruments, Union, CA), and spots were analysed using GenePix Pro 3 software. Manual adjustment of any spots that were poorly read by the software was undertaken to ensure only high quality microarray data was used. This involved manually checking the circled areas that the software had calculated as being a fluorescent spot. Cy3 and Cy5 labels were interchanged between sample and control cDNA in complementary experiments to reduce any of the effects from differential photobleaching of the two fluorescent dyes and cDNA incorporation inequalities and any inequalities in the incorporation of the two dyes into cDNA.

2.6 Microarray Analysis

Statistical functions written in R code, (Ihaka and Gentleman, 1996; for review see Ellner, 2001) freely available at http://www.r-project.org/, were used in a new statistical microarray analysis package, tRMA (tools for R Microarray Analysis; Wilson et. al., 2003). Details and a manual of tRMA are available online (http://www.pi.csiro.gena.trma.au). Normalisation was undertaken to eliminate possible biases in fluorescence due to a difference in label incorporation during cDNA synthesis and differences in the stability of the dyes. A spatial normalisation algorithm in tRMA helps to remove possible spatial fluorescence-based biases. The data were normalised using spatial-normalisation protocol from the tRMA package (Yang et. al., 2001). Median values were calculated for each gene from the total number of biological and technical replicate slides for each experiment. Biological replicates consisted of repeat experiments of groups

of six plants grown under the same conditions and material harvested from each of the plants at the same time and RNA pooled from each individual within each group. This level of replication was repeated for both the control C052 and each test sample line.

The biological replicates are necessary to try to avoid random gene expression due to variations in conditions and genetic material. While plant genetic material was matched as closely as possible, plants still vary in expression patterns of particular genes. As well as biological replicates, technical replicates were employed of each biological replicate to remove inconsistencies between the labelling and fluorescing efficiencies of each of the dyes.

Using the relevant function in tRMA, differentially expressed genes were extracted from the complete list of ~17000 clones' median values. Firstly the data is normalised and then rescaled with a mean of 0 and SD of 1 so that differentially expressed genes can be found statistically using standard Gaussian distribution. As the statistically defined genes always have a cut-off of less than 2-fold we applied a more stringent cut-off of 2 to select the differentially expressed genes.

2.7 Sequencing of Differentially Expressed Genes

Differentially expressed clones were taken from –80 °C storage and grown overnight in LB with 200 µg/mL Ampicillin and the next day plasmids were purified using the Qiagen QIAprep Spin Miniprep Kit (Catalogue # 27104).

Primers to the vector at the 5' and 3' regions of the inserted cDNA were used to get sequence data of the differentially expressed genes. The standard Big Dye Version 3.1 sequencing method was used and the precipitated product sent to the Australian Genome Resource Facility (AGRF) (University of Queensland St Lucia, QLD. Australia) and run on an ABI Prism Model 377 sequencer. Sequences were BLASTed against the NCBI database for both BLAST x and BLAST n results. For those clones that needed internal sequencing runs, internal

specific primers were designed which had an annealing temperature ≥65 °C and made by Sigma.

Sequencing was analysed on BioManager by ANGIS (University Of Sydney NSW, Australia), (http://www.angis.org.au/) first by running CodonCode BaseCaller (http://www.codoncode.com/oem/basecaller) version 0.000918 (Ewing *et al.*, 1998; Ewing and Green 1998) by CodonCode Corporation to determine correct coding of bases. Vector trimming was undertaken using CodonCode Matcher with vector (http://www.codoncode.com/oem/matcher) version 0.000918 by CodonCode Corporation. The resulting sequence was assembled using Contig Express supplied with Vector NTI 8 (Frederick, Maryland U.S.A.). Resulting sequence data was then screened for homology to database sequences using Biomanager by ANGIS for tBLAST x, BLAST x, BLAST n (Altschul, *et al.*, 1997). A cut off of an E value to power –60 was used to assign function or similarity to known genes in the database. tBLAST x values were used to help assign probable function rather than BLAST n values as there are few poppy genes sequenced in the database. It was generally found that the sequence similarities between: rice and poppy and *Arabidopsis* and poppy, were not close enough to give reliable BLAST n results and assign function. Chromatogram data was checked for poor sequence data and any that was found to be unreliable was discarded and repeated. The cDNAs were also screened for redundancy by alignment on Biomanager by ANGIS using ClustalIW (accurate) (Thompson *et al.*, 1994).

2.8 HPLC Analysis of Alkaloids

Latex from the plants was harvested by lancing the unripe capsules and adding approximately 0.5 µL of the exuding latex immediately to 250 µL of buffer (0.2 M ammonium phosphate, 0.25% SDS, 20% ethanol, pH 4.5). The sample was vortexed then centrifuged. 200 µL of the supernatant was decanted into a 40 x 8 mm auto analyser tube and an additional 250 µL of buffer was added to allow the sampling needle to reach the solution. The samples were run on a 96-tube carousel Waters HPLC system with a standard sample containing morphine, codeine, oripavine and thebaine. The HPLC phase was aqueous methanol

(approximately 30%) containing ammonium acetate buffer (0.08-0.12 M), pH 4-5. The flow rate of the mobile phase was 0.8-1.5 mL/minute. A Whatman Partisphere SCX column (4.6 x 125 mm) was used at a temperature of 40 °C. A Waters 440 UV detector was used to detect the peaks at 254 nm. The data was interpreted and collated on a Waters Millenium Data Station.

2.9 HPLC Analysis of Carotenoids

Approximately 100 µL of latex was isolated from each of the mutant *top1* and the control C052 after lancing the unripe capsule at 4 days post petal drop for possible carotenoid extraction and determination by HPLC (a detailed description of the HPLC procedure can be found in Norris *et al.*, 1995). 500 µL of acetone/ethyl acetate (60:40) was added and the solution vortexed. 400 µL of H_2O was added and the solution mixed again and spun fro 5 minutes at 13,000x rpm. The top layer containing the ethyl acetate (and colour from the *top1* latex) was transferred to a new tube (taking care not to get any of the water) and this was spun for 3 minutes. The samples were run on a standard C18 reverse phase column HPLC with ACN (Acetonitrile)/H_2O/TEA (triethylamine) background and an ethyl acetate gradient and run for 15 mins.

Saponification of samples was undertaken using a method based on Granado *et al.* (2001), the latex was extracted in ethanol (equal volume) and KOH added to a final concentration of 6% (v/v). The samples were left overnight under nitrogen in the dark and the next day diethyl ether was (approximately 100 µL) added to the sample. The carotenoid pigments partition into the ether phase. The sample was washed several times with distilled water to remove the KOH. The ethyl layer was removed and evaporated under oxygen-free nitrogen. The sample was reconstituted in ethyl acetate and injected onto the HPLC.

2.10 TLC Analysis of Alkaloids

Thin Layer Chromatography (TLC) was used to analyse alkaloids. Latex was collected from the mid vein of a torn leaf and adding to 200 µL of latex extraction buffer (20 % EtOH, 0.25 % SDS and 0.2M $NH_4H_2PO_4$). The latex-buffer solution was vortexed and allowed to sit for 30 minutes then spun and 5 µL taken for loading onto a TLC plate (Merck Silca gel 60). The plates were run according to Wagner (1984) using a chromatography solvent of toluene: acetone: ethanol: concentrated ammonia (40:40:6:2).

Standards used for the TLC were dissolved in 50 % EtOH/ 50 % NH_4OH in the following concentrations 2.5 mg/mL oripavine, 1.25 mg/mL thebaine, and 5 mg/mL codeine base, 5 mg/mL and morphine base. All alkaloids were kindly provided by Tasmanian Alkaloids Pty Ltd.

2.11 Radioactive Feeding Studies

[NB. This section (Section 2.12) describes experiments conducted by a collaborator in Germany. They are included here, rather than the introduction, as there has been no other publication of the method for this data for the mutant *top1*].

Plants used in the feeding studies were germinated in pots (30-50 cm in diameter) and allowed to grow for 5-6 weeks (10-12 cm total length) in the greenhouse before cutting the plant just above the root and placing in a 1.5 mL microcentrifuge tube filled with 300 µL of feeding solution. These plants were then placed in a transparent plastic lidded box containing enough Millipore water to obtain a relative humidity of 80-90% to stop the plants from wilting.

Feeding solutions consisted of: 200 000 cpm ^{14}C-oripavine and 9 nmol of oripavine; 215 000 cpm ^{14}C-thebaine and 15 nmol of thebaine; 500 000 cpm ^{3}H-codeine and 10 nmol codeine; 500 000 cpm ^{3}H-salutaridinol and 16 nmol salutaridinol; 260 000 cpm ^{3}H-*R*-reticuline and 6 nmol *R*-reticuline; each made

up to a total volume of 300 µL with Millipore water. The duration of each of the feeding studies was 24 hours for oripavine, thebaine, salutaridinol and *R*-reticuline. The codeine solution was fed for a period of 72 hours. To prevent plants from running dry Millipore water was used to top up the microcentrifuge tube.

After the feeding assay the plants were cut into small pieces and extracted in 80% Ethanol for 15-20 minutes. The crude extract was lyophilised and dissolved in a small quantity of methanol and run on TLC. A two dimensional TLC was used with Dimension 1 solvent system consisting of chloroform: acetone: diethylamine (5:4:1) and dimension 2 consisting of toluene: acetone: ethanol: ammonia (45:45:7:3).

2.12 Semi Quantitative Real time PCR

For analysis of the differentially expressed cDNA clones gene-specific oligonucleotides were made by Sigma. RNA was extracted using the Qiagen RNeasy Plant kit (Catalogue # 74903) as per the method supplied. Primer sets were designed using the Primer3 software package (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi) (Table 2.1). These primers were designed to have a Tm of 59-63 °C (with the optimal being =60 °C), yield a PCR product around 225-275 bp (optimal being 250 bp), GC content from 40-60% (with the optimal being 50 %), a Max self-complementarity equal to 4 and a Max 3' self-complementarity equal to 3. Qiagen Quantitect™ SYBR® Green Real Time PCR kit (Catalogue # 204243) was used in half reactions as per the manufacturer's instructions, for authentication of differentially expressed genes. A typical reaction contained 10 µL Quantitect™ SYBR Green RT-PCR Master Mix, 0.5 µM each of forward and reverse primers, 0.2 µL of Quantitect™ RT mix, 10 ng of RNA and RNase-free water to 20 µL. The reactions were carried out on a RotorGene 2000 (Corbett Research), and data was analysed by 'Relative quantitation', a comparison of two or more genes to each other and the result calculated as a ratio. (For further information see http://docs.appliedbiosystems.com/pebiodocs/04303859.pdf). See Appendix 2

Table 2.1: List of primers and the respective sequence used for the qRT-PCR, RT-PCR reactions and for further analysis.

Primer Name	cDNA and Primer direction	Size (bp)	Sequence of Oligonucleotide
AGM261*	221C9 5'	24	ACGGTCCTGAAAGATAGAGACTTG
AGM262*	221C9 3'	24	TCCAGAGAAAAGGAAAAACAAAAC
AGM263*	221C9 3'	24	ACCAGTGAAAGTAAACATCCCAAT
AGM264*	221C9 3'	24	GATAGGTTTCCTCCAGGATCTACA
AGM265*	221C9 5'	24	GTTTACTTTCACTGGTTTGAATAA
AGM180	221C9 5'	21	GCTCAGGATCAATATGGGCTA
AGM181	221C9 3'	20	TTAATAGCATCCGCAAGGTG
AGM249	27D12 5'	21	GGCTGAAGAGGTGATATTTGG
AGM250	27D12 3'	20	TATGCCTTGTCCACCAACTC
AGM256	116D7 5'	20	AAAAATACCGACGCCTGATG
AGM257	116D7 3'	20	ATGCAAAAGCCGAGAAAATG
AGM188	157C4 5'	20	ATGTCGCCCTCTTGGTTTAT
AGM189	157C4 3'	20	GTAGTCCACCTCTCGGGTGT
AGM182	91G3 5'	20	CCTTGTTCGCATTTTGACTC
AGM183	91G3 3'	20	AGGAATCTGGAGCCTACGAG
AGM64	SAT 5'#	20	ACAACAACAAATGCGGTCAA
AGM65	SAT 3'#	20	ATAATCCCATCCTGCACCAG
AGM196	111C7 5'	20	GCCATAACATTTGAGGGTGA
AGM197	111C7 3'	20	TTCCATTCCTCTCCAGGTTT
AGM259	91B6 5'	20	ATCAAAAATGCCCACATGGT
AGM260	91B6 3'	20	TTCAGCAACGACACATCACA
AGM137	54F9 3'	28	AGAATAAGCAAGGCGGCTCTAAGT TCAA
AGM138	54F9 5'	28	CTTACGCGTACCCAGCTTTCTTGTA CAA
AGM210	151H8 5'	20	TCCACCAGAAAAGAGGGTTT
AGM211	151H8 3'	20	TTTGTCTCGCCCAGGAGTAT

* these primers were extra primers designed to the cDNA 221C9 for further RT-PCR analysis see Chapter 5. # is the Salutaridine acetyl transferase (SAT) gene used as a control for the qRT-PCR analysis.

for the detailed qRT-PCR data.

2.13 Reverse transcriptase PCR

Several attempts were required to detect the appropriate number of cycles to show a difference in expression levels between the control and mutant *top1* RNA samples. Each reaction sample was repeated in triplicate to authenticate the results. RNA was run on a gel (See Chapter 2.3) to authenticate accurate spectral

quantification, as well as using an internal control of primers to the salutaridine acetyl transferase gene. The primers for this gene were also picked using the same program as the one used for the selection of qRT-PCR primers. Resulting reverse transcriptase-PCRs (RT-PCRs) were run on 0.5 X TBE agarose gel in 0.5 X TBE buffer at 120 mV and 0.8 Amps and visualised using the Biorad Gel Documentation System.

Chapter 3

Characterisation and Definition of the Different Germplasm Used in this Thesis

3.1 Introduction

To further the aims of this thesis a range of germplasm have been assembled. These include: several mutants of *P. somniferum* with altered morphinan alkaloid phenotypes (morphinan alkaloid yields of each of the mutants and their respective parental phenotypes are presented in Table 3.1); a low morphine cultivar of *P. somniferum*, Marianne; *Eschscholzia californica,* a distant member of the *Papaveraceae* family; and a transgenic *P. somniferum* with altered morphinan phenotype.

This chapter, where applicable, will extend the classical characterisation of the different genotypes, including HPLC and TLC analysis of the alkaloids and, in the case of the mutant *top1,* an investigation of the altered latex colour in order to determine if the colour change is related to the altered alkaloid phenotype. Further, feeding studies of the mutant *top1* will be presented to establish the blockage in the morphinan pathway and preliminary proteomic analysis will be presented to illustrate the extent of differences in protein expression in *top1*. A focus will be given to the alkaloid mutant *top1* due to its economic significance, its relevance to studying the last steps of morphine biosynthesis, and its genetic stability.

Table 3.1: Table showing the average dry straw alkaloid yields (%DW) from field trials of various mutants used in the study.

Mutant	Characteristic	Parent	Morphine	Codeine	Oripavine	Thebaine	Papaverine	Reticuline
C052	Parent line	-	1.87	0.09	0.02	0.08	0.00	0.00
W52-12-10	Parent line	-	2.1	0.11	0.00	0.16	0.00	0.00
D205	Low M	C052	0.92	0.00	0.01	0.00	0.00	0.00
D120	High P, low M	C052	1.76	0.08	0.00	0.01	0.45	0.00
D242	High O and T, high total morphinans*	C052	1.66	0.09	0.31	0.41	0.00	0.00
E40/41	M free, high R	W52-12-10	0.00	0.00	0.00	0.00	0.00	3.00
top1	M and C free. High T and O	C052	0.00	0.00	1.0	1.3	0.00	0.00
F075⌖ null	Null segregant of F075	C067#	11	6	1	82	0	0
F075⌖	M free, low O high T	C067#	0	0	0	100	0	0

* high total morphinans refers to the measured morphine, codeine, oripavine and thebaine combined yields. Field trials were undertaken by Tasmanian Alkaloids Pty Ltd. # C067 is the parental line used to generate the F075 mutant however the null segregating mutants were used as controls for this thesis. ⌖ As the F075 line was a segregating population the results from this trial are from individual plants. The results for the F075 trial are also given as a percentage of the total alkaloids measured (morphine, codeine, papaverine, thebaine, oripavine and reticuline), rather than as percentage dry weight yields. All three control lines (C067, W52-12-10 and C052) are morphine accumulating cultivars with small differences in yield characteristics. M=morphine, C=codeine, O=oripavine, T=thebaine, P=Papaverine, and R=reticuline (*R* and*S* enantiomer).

3.2 Reticuline Accumulating Germplasm

The first grouping of genotypes with altered alkaloid phenotypes is defined here as the reticuline accumulating germplasm. This class includes: *Eschscholzia californica* which, contains the alkaloid pathway to the reticuline branchpoint before continuing on to sanguinarine (Figure 3.1); the reticuline accumulating mutant *P. somniferum* (E40/41) (Figure 3.1); and the reticuline accumulating hairpin RNAi silenced transgenic *P. somniferum* (Figure 3.1).

3.2.1 *Eschscholzia californica*

Eschscholzia californica is a member of the *Papaveraceae* family and is commonly known as the California poppy. A native to North America it was named after Johann Friedrich von Eschscholtz a famous botanist, physician and naturalist. The plant grows as a perennial, regenerating from the roots in spring in colder climes, and flowers annually with a gold coloured flower. The American Indians, who ate the boiled leaves and sometimes smoked the leaves and petals, used *E. californica* in traditional medicine.

Recent advances in *E. californica* transformation techniques (Park and Facchini, 1999; 2000a; 2000b; 2001; Lee and Pederson, 2001) will allow genetic manipulation of the benzophenanthridine pathway in this plant species. Transformation of plant cell suspensions and root cultures with BBE or CYP80B1 (Park *et al.*, 2002 and 2003), illustrate that metabolic engineering of benzophenanthridine alkaloid accumulation is achievable. The stable genetic transformation of this species will help analysis of the regulation of the benzophenanthridine pathway by allowing newly isolated genes to be expressed or knocked out in stable cell or whole plant lines.

The main alkaloids found to accumulate in the *E. californica* plant are the pavine type, such as eschscholtzine and californidine and the benzophenanthridine alkaloids sanguinarine (Figure 1.6) and chelerytrine (Tomè *et al.*, 1999).

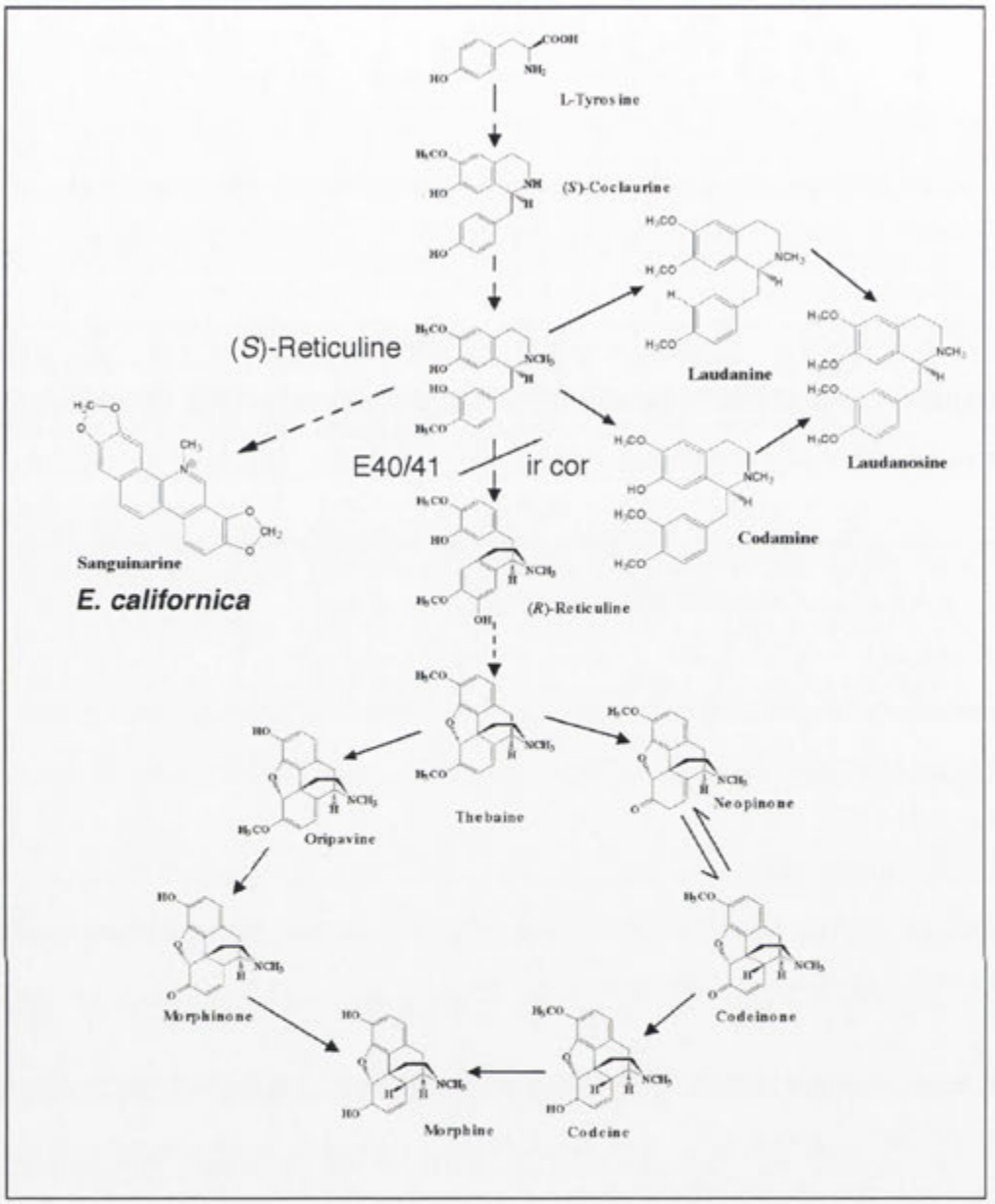


Figure 3.1: Abbreviated alkaloid pathway illustrating the accumulating products for each of the germplasm used in this subchapter of reticuline accumulating phenotypes. *E. californica* contains the benzylisoquinoline alkaloid pathway until the major branchpoint intermediate *S*-reticuline and from here progresses to sanguinarine, the major accumulating alkaloid. The morphinan alkaloids have not been found in *E. californica*. The mutant *P. somniferum* that accumulates *S*-reticuline also accumulates the intermediates laudanosine and laudanine. The hairpin inverted-repeat transgenic *P. somniferum* has the same alkaloid phenotype as the reticuline-accumulating mutant. Adapted from Kutchan 1998.

Eschscholzia cell suspensions have been used as a model system to study benzophenanthridine alkaloid biosynthesis since the accidental discovery that the medium changed colour when a penicillium contaminated *E. californica* callus (Schumacher *et al.*, 1987). The use of *Eschscholzia* cell suspensions has led to the elucidation of the entire enzymatic pathways of sanguinarine and macarpine synthesis (Reviewed in Kutchan and Zenk, 1993; Zenk, 1994; Kutchan, 1995 and 1996; Chou and Kutchan, 1998). The ability of this species to produce alkaloids in suspensions has made the enzymatic determination of these pathways possible. Its use has led to the discovery of some very complex reactions, which take place via the action of novel cytochrome P450 enzymes; genes cloned include those encoding CYP80B1 (Pauli and Kutchan, 1998) and *N*-methylcoclaurine 3'-hydroxylase. Elicited *E. californica* cell suspensions also enabled the berberine bridge enzyme (BBE) to be analysed and cloned (Dittrich and Kutchan, 1991; Kutchan and Dittrich, 1995; Hauschild *et al.*, 1998).

The morphinan alkaloid pathway has not been detected in either the plant or cell suspensions of *E. californica.* Both *Papaver* species share the pathway from tyrosine to the major branchpoint intermediate (*S*)-reticuline (Figures 1.2 and 1.3) and also the branchpoint berberine bridge enzyme leading to the production of the benzophenanthridine sanguinarine (reviewed in Kutchan 1996). As mentioned above the gene encoding the P450 CYP80B1 was cloned from *E. californica* suspension cultures (Pauli and Kutchan, 1998) and a gene coding for the equivalent enzyme in *P. somniferum,* 76% identical at the amino acid level, has since been found (Facchini and Yu, unpublished, NCBI # AF191772). The *Papaver* NADPH-dependent cytochrome P450 reductase was also characterised and cloned from *E. californica* cell suspensions and shows 63% homology to the *P. somniferum* cDNA at the nucleotide level (Rosco *et al.*, 1997).

The *cyp80b1* transcripts for both *P. somniferum* (Huang and Kutchan, 2000; Facchini and Park, 2003) *E. californica* (Pauli and Kutchan, 1998) are both induced upon addition of methyl jasmonate. Similarly BBE has also been shown to accumulate in response to elicitors such as methyl jasmonate in both *P. somniferum* (Huang and Kutchan, 2000; Facchini and Park, 2003) and *E. californica* (Dittrich and Kutchan, 1991; Hauschild *et al.*, 1998) suggesting that

the benzophenanthridine and benzylisoquinoline pathways are under similar transcriptional controls. While there are many similarities between the two species, the *bbe1* gene promoter isolated from *E. californica* and fused to a β-glucuronidase (gus) gene, was not functional in *P. somniferum* cell suspensions (Hauschild *et al.*, 1998). Since *E. californica* entirely lacks the morphinan branch after (*S*)-reticuline, substantial genetic differences between *P. somniferum* and *E. californica* are expected and this should be reflected in a microarray comparison by a large number of transcriptional differences. The object of including a global transcript comparison of *E. californica* to *P. somniferum* by microarrays was to obtain a broader scope of the transcriptional differences and determine if similar reticuline branchpoint accumulating germplasm had similar sets of genes differentially expressed. It may be expected that if these genes are specifically involved with the morphinan pathway they will also be relatively under-expressed in *E. californica* where the morphinan branch is also missing.

3.2.2 *P. somniferum* Reticuline Accumulating Mutant, E40/41

This is the second of three types of germplasm in the study classified as reticuline accumulating. The mutant, E40/41, was created using the EMS mutation process and discovered at Tasmanian Alkaloids. This mutant was produced in parental cultivar, W52-12-10, rather than cultivar C052. W52-12-10 and CO52 are said to be relatively close genetically (Dr. Fist, pers. comm.). TLC (Figure 3.2) and HPLC (Figure 3.3) analysis of latex alkaloids revealed a mutant that failed to accumulate any appreciable morphinan alkaloids and instead yielded high quantities of reticuline (See table 3.1). Further analysis showed that it was the *S*-enantiomer of reticuline that accumulated in this mutant and not the *R* form. The mutant also accumulated laudanine, laudanosine and codamine (Figure 3.1) in measurable quantities that are not usually seen in these quantities in the parental line. These compounds are simple methylated derivatives of (*S*)-reticuline.

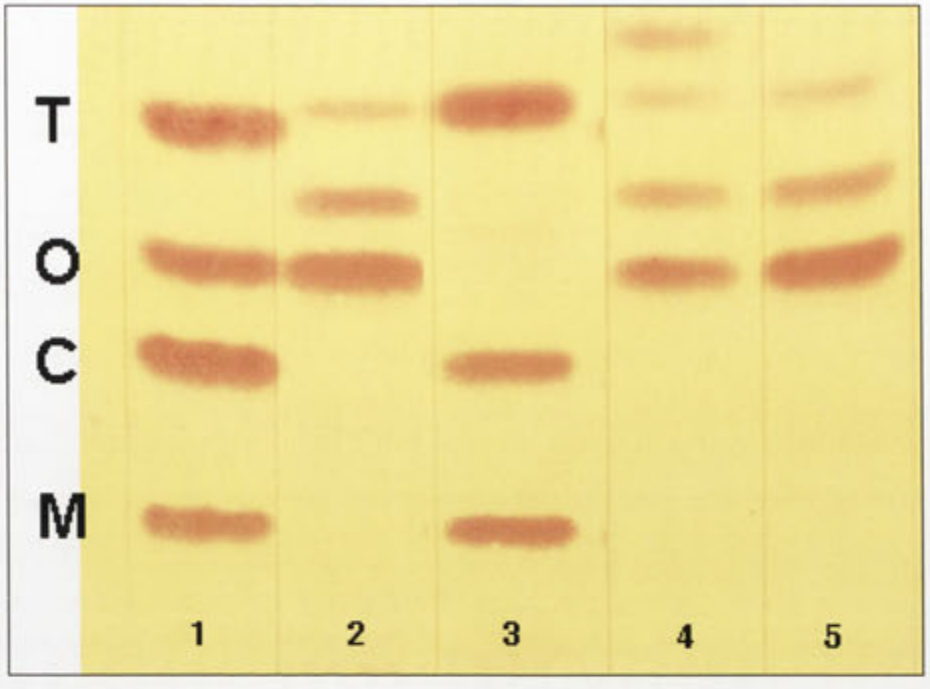


Figure 3.2: Picture of a TLC illustrating the difference in alkaloid accumulation between the parental control (3) (*P. somniferum* cv. W52-12-10) and the EMS mutant E40/41 (4). Lane 1 are the standards: morphine (M), codeine (C), oripavine (O) and thebaine (T). Lanes 2 and 5 are loadings of reticuline standards, which contain (from the bottom of the TLC to the top) reticuline, codamine and laudanine. Reticuline, codamine, laudanine and laudanosine are the dominant alkaloids found in mutant E40/41. The loading on this TLC was not enough to visualise the very low quantities of morphine and codeine, which can be seen on the following HPLC (Figure 3.3).

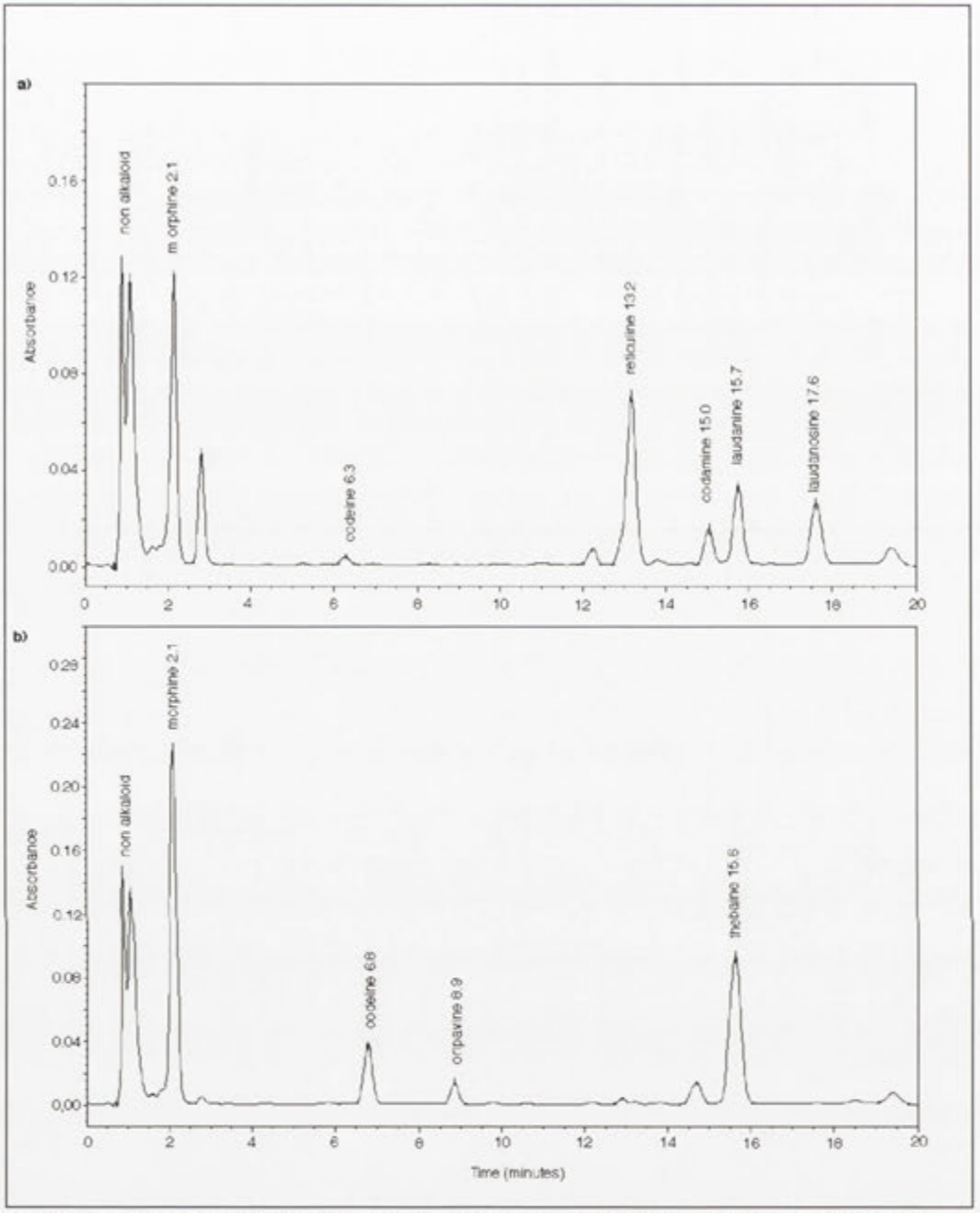


Figure 3.3: HPLC analysis of the reticuline-accumulating mutant, E40/41 (a) compared to the control W52-12-10 (b). The reticuline mutant still accumulates codeine and morphine but at much lower concentrations than the control. The reticuline mutant also accumulates codamine, laudanine, and laudanosine which are normally difficult to detect in the control phenotype.

3.2.3 Hairpin RNAi Silenced Codeinone Reductase *P. somniferum*

This is the third of three types of germplasm in the study classified as reticuline accumulating. It is a transgenic *P. somniferum* genotype in which codeinone reductase (Cor) has been silenced. Codeinone reductase is the penultimate enzyme of the morphine biosynthetic pathway in *P. somniferum* (Figures 1.2 and 1.3). The cloning of the codeinone reductase gene (Unterlinner *et al.*, 1999) and the development of a robust transformation system (Chitty *et al.*, 2003), have led to the first genetically transformed opium poppies with a modified morphinan alkaloid pathway (Allen *et al.*, 2004). This was achieved by transforming *P. somniferum* with a chimeric DNA-directed hairpin RNA construct designed to silence all members of the codeinone reductase multigene family through an RNAi mechanism. I am co-author on the paper describing this dramatic example of metabolic engineering (Allen *et al.*, 2004).

The efficacy of gene silencing in plants using inverted-repeat transgene constructs that encode a hairpin RNA (hpRNA) has been demonstrated for a range of genes (Wesley *et al.*, 2001; Wang and Waterhouse, 2002; Helliwell *et al.*, 2002; Waterhouse and Helliwell, 2003). This technology relies on using intron-containing constructs, where an intron separates the complementary sense and antisense sequences. RNAi silencing constructs have resulted in 90-100% of independent transgenic plants showing silencing. The degree of silencing with these constructs is much greater than that obtained using either co-suppression or anti-sense constructs (Smith *et al.*, 2000). The advantages of this technology are the ability to specifically silence one, some or all members of a multigene family by targeting unique or shared sequences.

Several isoforms of codeinone reductase were isolated from poppy, which have a high degree of sequence homology. The employment of RNAi silencing techniques to specifically knock out all members of the codeinone reductase gene family offers the potential to increase our understanding of the pathway. The codeinone reductase gene family contains at least 6 members (*cor1.1-1.6*) (Unterlinner *et al.* 1999). A closely related gene, *cor2* was also isolated, and although attempts to demonstrate that it encoded functional codeinone reductase

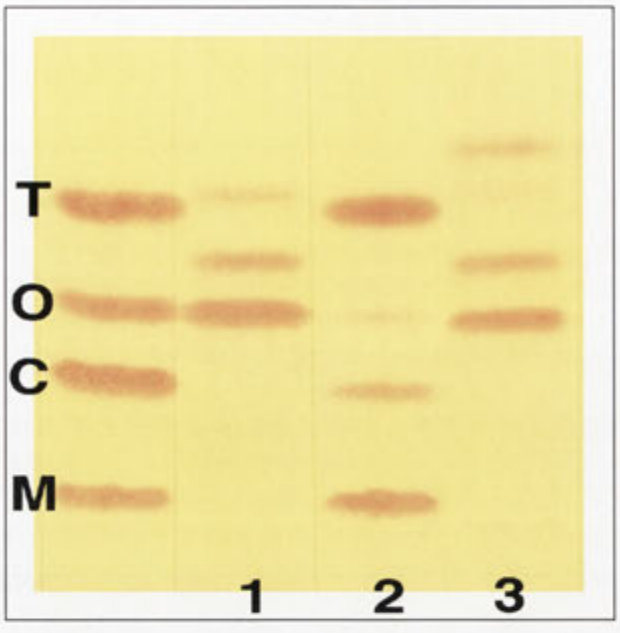


Figure 3.4: TLC illustrating the difference in the alkaloids accumulated by (1) the reticuline standard consisting of reticuline, laudanine and laudanosine from the bottom to the top of the TLC plate. (2) The control parental phenotype, (3) one of the codeinone reductase RNAi silenced transformed plants, which also has codamine (upper band) present on the TLC. Adapted from Allen *et al.*, 2004.

by expression in *E. coli* were unsuccessful (Unterlinner *et al* 1999), this gene was also silenced in case the Tasmanian cultivar C058-34, used for the transformation, did possess functional *cor2*. In order to silence all possible members of the *cor* family, regions of cDNA from both *cor*1.1 and *cor*2 were incorporated into a hpRNA based silencing vector, driven by a strong constitutive promoter (Allen *et al.*, 2004). The region chosen from *cor1.1* was identical to all the known members of the *cor1* family.

TLC (Figure 3.4), HPLC (Figure 3.5) and mass spectrometry showed that the precursor (*S*)-reticuline accumulates in transgenic silenced lines at the expense of morphine, codeine, oripavine and thebaine. (*S*)-reticuline is 9 enzymatic steps upstream of the substrate codeinone (Figure 1.3). Northern analysis and RT-PCR of transgenic lines verified loss of codeinone reductase gene transcript. Further analysis revealed the presence of 22-mer degradation products in silenced lines, and the corresponding enzyme activity was substantially reduced (Allen *et al.*, 2004). The methylated derivatives of (*S*)-reticuline also accumulated in the

codeinone reductase RNAi silenced plants, including: codamine, laudanosine and laudanine (Figure 3.1).

There was some variability in the degree of expression of the *cor* RNAi induced reticuline phenotype in the transgenic T1 plants analysed. For example lines 220-15-2, 215-25-1b, and 220-9-2 accumulated very little morphine, with over 90% of total alkaloid comprising reticuline and its methylated derivatives, whilst families 220-2-1 and 13-2 still accumulated morphine at up to 75% of total alkaloid. Most lines produced over 75% reticuline and its derivatives, which is profoundly different to the alkaloid profile of the commercial Tasmanian parental line C058-34 (Allen *et al.*, 2004). Screening 18 independent primary transgenic lines showed that reticuline and its derivatives accumulate in direct proportion to the decrease in morphine, codeine, thebaine and oripavine. The transformed plants showed stable inheritance of the reticuline phenotype (Allen *et al.*, 2004).

Chiral analysis was undertaken on methylated compounds, isolated by HPLC from transgenic latex of reticuline, codamine, laudanine and laudanosine, to determine if these compounds were the (*R*) or (*S*) enantiomer. Comparison with optically pure standards of (*R*)- and (*S*)-laudanosine, revealed these compounds were present in the (*S*) form. Allen *et al.*, (2004) showed that a reduction in the level of morphinan alkaloids in the *cor* RNAi transgenics also led to a concomitant reduction in the level of codeinone reductase activity and codeinone reductase message levels.

Silencing of codeinone reductase has led to the accumulation of the general intermediate (*S*)-reticuline at the expense of morphine, codeine, thebaine and oripavine. The accumulation of (*S*)-reticuline suggests that a negative feedback mechanism exists between the unique morphinan specific pathway, and the benzylisoquinoline secondary metabolic pathway, of which (*S*)-reticuline is a key intermediate. Such negative feedback seems the only cogent explanation for how the entire 9-enzyme morphinan pathway can be closed down in response to the loss of codeinone reductase. By reducing the production of codeinone reductase transcript and its corresponding protein product the blockage is somehow relayed back to the metabolic junction point at (*S*)-reticuline. This intermediate is shared

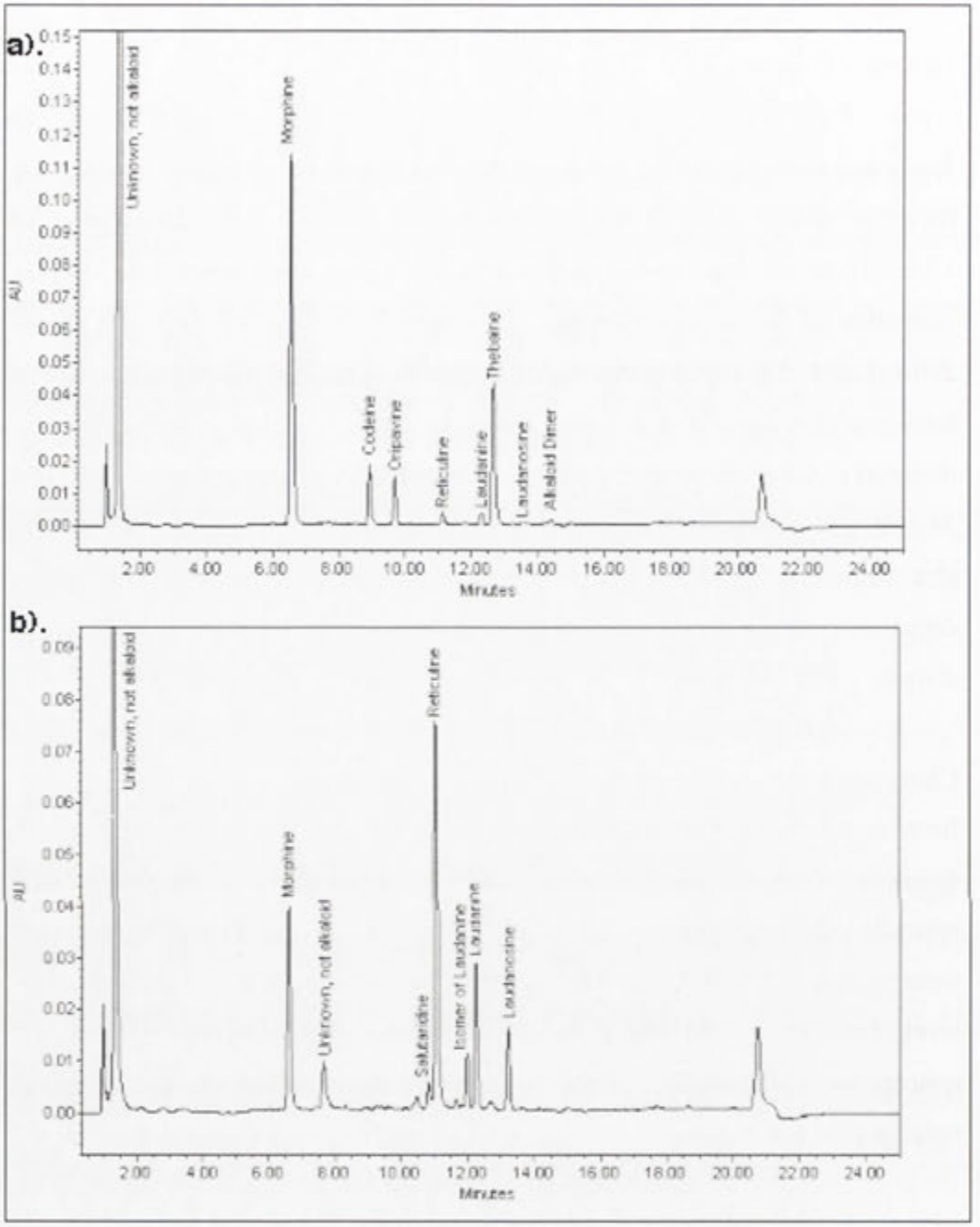


Figure 3.5: HPLC profiles of the *cor* RNAi transgenics (b) and the parental line C058-34 (a) used for the transformation showing differences in the alkaloids accumulated. The strongest expressing transgenics no longer accumulate morphine, codeine, thebaine, and oripavine, instead accumulating (*S*)- reticuline, laudanosine, laudanine and codamine. From Allen *et al*., 2004.

---

with other alkaloid pathways in *Papaver* and other species, such as those responsible for berberine, benzophenanthridines (for example sanguinarine),

phthalide isoquinolines (for example noscapine) and benzylisoquinolines (for example codamine, laudanine and laudanosine) (Figure 3.1).

A number of explanations for the resulting phenotype of the *cor* RNAi transformed plants. These explanations include a slight build up of the substrates codeinone, morphinone or neopinone which may act as inhibitors of one of the early enzymes of the morphinan branch, such as reticuline oxidase (=1,2-dehydroreticuline synthase) (Hirata *et al*, 2004). The substrate feedback may act to inhibit transcription of genes encoding earlier enzymes or inhibit the earlier enzymes themselves. Another possibility is that codeinone reductase enzyme may form a larger inter-dependent enzyme complex; the loss of codeinone reductase may disable the entire complex resulting in a failure of the first steps of the morphinan branch (Allen *et al*., 2004).

There is evidence in flavonoid biosynthesis in Arabidopsis endoplasmic reticulum for such a multienzyme complex (Burbulis and Winkel-Shirley, 1999); chalcone synthase, chalcone isomerase, and dihydroflavonol 4-reductase interact in an orientation-dependent manner. Similar evidence exists for multi-enzyme complexes involved in polyamine synthesis (Panicot *et al*., 2002), where metabolites are thought to canalise down a particular pathway as they pass from one enzyme to the next in the complex. This multienzyme complex is termed a *metabolon* and has been invoked to explain some features of *O*-methyltransferase associations in isoflavonoid biosynthesis (He and Dixon, 2000). It has been suggested metabolons may be associated with morphine biosynthesis (Facchini *et al*., 2003) but without evidence.

Compartmentalisation in alkaloid synthesis and accumulation has been demonstrated (St-Pierre *et al*., 1999, Bird *et al*., 2003), and it is possible that the transgenic codeinone reductase RNAi plants have resulted in changes to the normal transport and storage architecture, resulting in the failure of (*S*)-reticuline accumulating in the correct compartment for further processing into the morphinan alkaloids. (*S*)-reticuline is transported into vesicles preferentially to (*R*)-reticuline (Deus-Neumann and Zenk, 1986). Perhaps the substrate feedback inhibition in *cor* silenced plants acts on the transport or compartmentalisation

processes. RNAi silencing of codeinone reductase has revealed an unforseen level of control over the morphinan alkaloid branch.

The alkaloid profiles of the E40/41 mutant and codeinone reductase RNAi silenced plants are nearly identical, however the molecular mechanisms behind the phenotype are presumably quite different. The inclusion of both in the microarray experiments and comparison of which genes are differential in both cases should be interesting and potentially revealing.

3.3 Thebaine Accumulating Germplasm

The second grouping of genotypes to be considered consists of those that accumulate a profile of morphinan alkaloids different to wild type. These include a mutant *P. somniferum* accumulating thebaine (F075) and a *P. somniferum* mutant accumulating thebaine and oripavine (*top1*) (Millgate *et al.*, 2004) (Figure 3.6).

3.3.1 *P. somniferum* Thebaine Accumulating Mutant F075

The mutant F075 accumulates thebaine (Figure 3.6) as the dominant alkaloid without a corresponding accumulation of oripavine that occurs in the mutant, *top1*, (see section 3.3.2). The F075 mutant was similar to a spontaneous mutant reported by Nyman (1978, 1980) and like the spontaneous mutant the F075 mutant is unstable. Both Tasmanian Alkaloids Pty Ltd and I have tried unsuccessfully to fix the mutation by selfing mutant phenotypic individuals; each time thebaine individuals are selfed very few of the progeny show the thebaine phenotype. The phenotype does not follow simple Mendellian genetics nor even reproducible non-Mendelian ratios. Seed was planted from one F075 mutant line that accumulated thebaine but no detectable morphine, oripavine or codeine. The seed was grown as per Chapter 2 and a number of plants were taken through to the prehook stage of development before screening for progeny that expressed the thebaine phenotype.

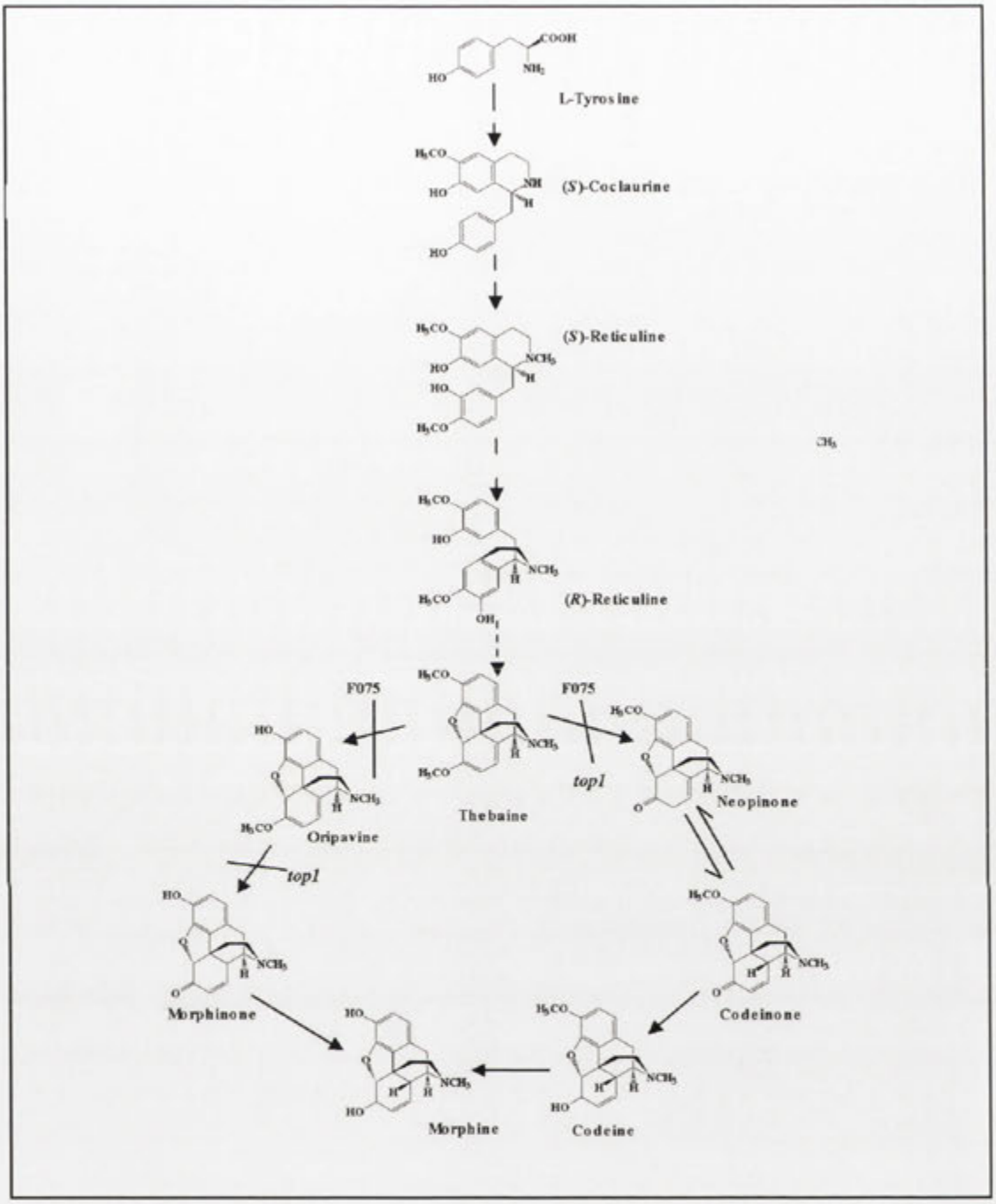


Figure 3.6: Abbreviated pathway illustrating where the two mutants discussed in this subchapter. The thebaine-accumulating *P. somniferum* accumulates only thebaine suggesting some form of a blockage in the pathway after the production of thebaine. The *top1* mutant accumulates the morphinan alkaloids thebaine and oripavine. The blockages are represented by lines through the enzymatic steps. Adapted from Kutchan 1998.

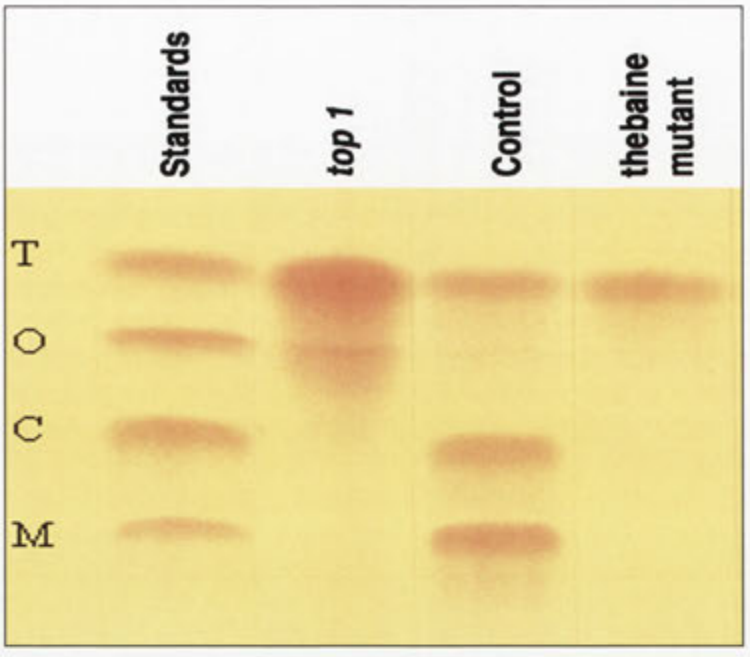


Figure 3.7: TLC showing the morphinan alkaloids accumulating in the latex of the F075 thebaine mutant compared to the control phenotype and *top1* mutant (see next subchapter 3.3.2).

Six out of a total of 160 plants tested were found to accumulate thebaine as the dominant morphinan alkaloid. The remaining plants were similar to the parental alkaloid phenotype. TLC (Figure 3.7) and HPLC (Figure 3.8) data was undertaken to authenticate the segregating individuals used for the microarray studies in the next chapter. The difficulty in obtaining a stable mutant F075 line hindered further work on this mutant and feeding studies, such as was undertaken with *top1*, have not been used to determine the enzymatic blockage in the pathway.

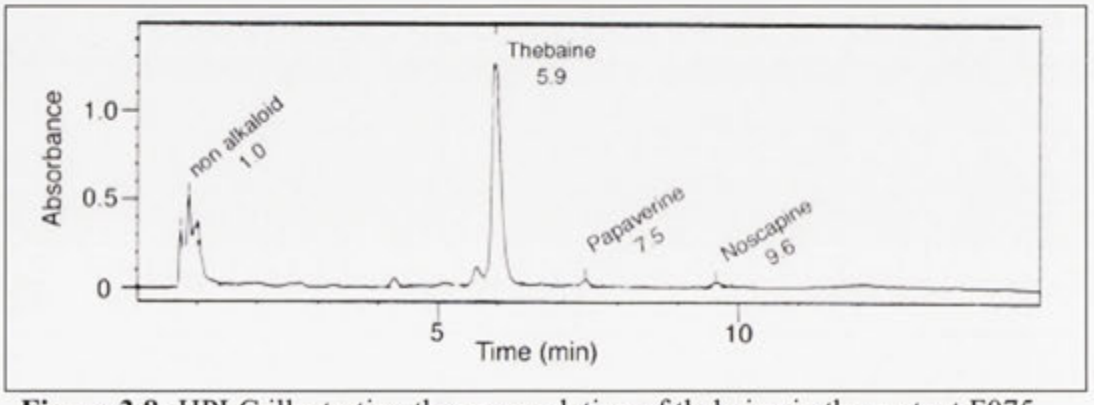


Figure 3.8: HPLC illustrating the accumulation of thebaine in the mutant F075.

3.3.2 *P.somniferum* Thebaine and Oripavine Accumulating Mutant, *top1*

The second mutant of the grouping consisting of an altered morphinan profile is a mutant called *top1*. It was an EMS induced mutant isolated at Tasmanian Alkaloids Pty Ltd. and shown to accumulate thebaine and oripavine but neither morphine nor codeine (Millgate *et al.*, 2004) (Figure 3.9). This morphinan alkaloid difference is further illustrated on the TLC plates (Figure 3.10). The mutation was named '*top1*' for thebaine and oripavine accumulating poppy. A true-breeding homozygous *top1* line was developed and used for further characterisation of the phenotype. Two independent crosses were advanced to F2 generations. The segregation of 375 individuals was Mendelian and consistent with a single gene (Table 3.2). The performance in the field revealed that there was no change in the total morphinan alkaloid yield of *top1* compared to its parent line CO52. It was difficult to accurately determine the heterozygote individuals from the homozygous wildtype or mutant phenotype as morphine was

Table 3.2: Segregation data for the *top1* phenotype

	Cross 1[a]	Cross 2[b]	Total
Morphine Phenotype	189	83	272
Thebaine Phenotype	81	22	103

[a,b] Two separate crosses between *top1* and normal morphine lines were made and advanced to F2.

As it was not always possible to distinguish heterozygotes from homozygous wild type, these two categories are combined for the analysis. Chi squared for equality of segregation ratio in the 2 crosses was 3.106, P between 5 and 10%, therefore acceptable that they represent the same genetic ratio.

Chi squared for the totals against a 3:1 ratio with Yates' correction, was 0.44 (P= 0.5068) and therefore conforms to a 3:1 ratio. (Millgate *et al.*, 2004)

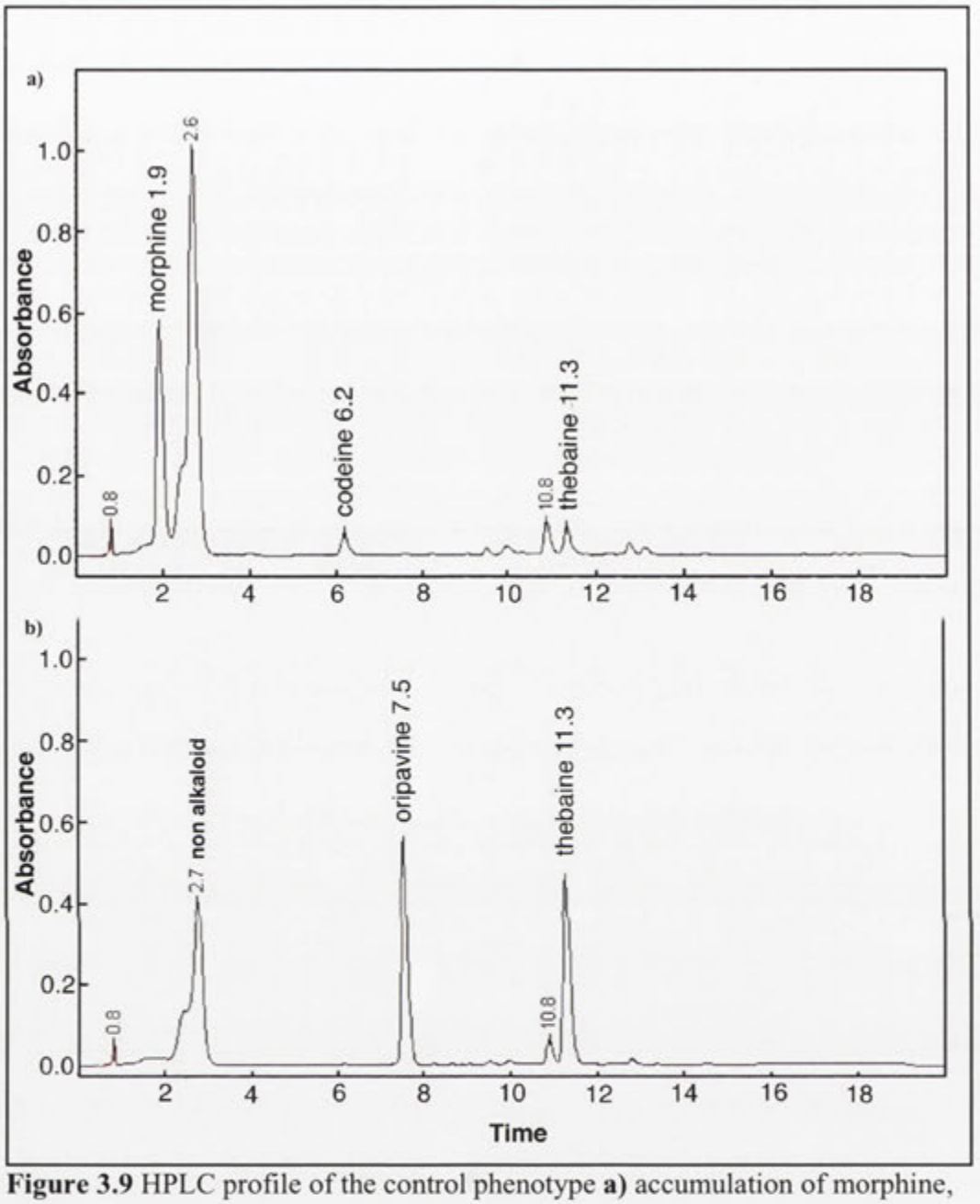


Figure 3.9 HPLC profile of the control phenotype **a)** accumulation of morphine, codeine and some thebaine, compared to the *top1* phenotype **b)** accumulating oripavine and thebaine only (Millgate *et al.*, 2004).

produced in the heterozygotes, suggesting that the mutation is neither dominant nor recessive, but rather semi-dominant (Millgate *et al.*, 2004). *Top1* developed similarly to the control parental type and only two apparent phenotypic differences were observed: The first being the altered alkaloid profile of thebaine and oripavine accumulation; and the second being the appearance of pigmented latex rather than the usual white colour seen in the control.

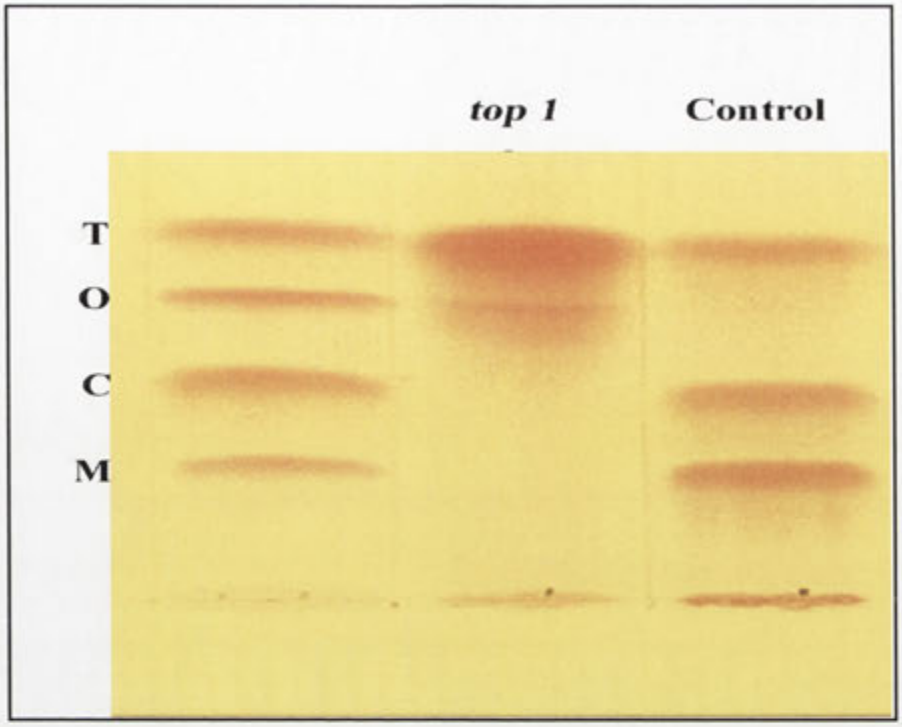


Figure 3.10: TLC showing *top1* accumulating oripavine (O) and thebaine instead of the control phenotype of thebaine (T), codeine (C) and morphine (M). The standards are run on the left lane (Millgate *et al.*, 2004).

3.3.2.1 Mutant Latex Colour Characterisation of *top1*

The only visible phenotypic change in the mutant was pigmented latex compared to the white coloured latex of *P. somniferum* (Figure 3.11). The latex colour change co-segregates with the thebaine phenotype and has become a reliable visual marker for the determination of this plant. HPLC analysis was undertaken to determine if the compound/s responsible for the colour change were carotenoids, in particular lycopene, a red coloured compound. The pink to orange colour of the latex lends itself to the possibility of a carotenoid being responsible as the colour absorbs around the same wavelength as the red coloured carotenoid lycopene. It is possible that the mutation, or a secondary mutation, may be perturbing the carotenoid pathway in some way. If it is a secondary mutation then it is tightly genetically linked to the mutation causing the alkaloid phenotype. While no strict quantification of this linkage is available, many thousands of segregants from crosses between *top1* and morphine poppies have been produced and examined at Tasmanian Alkaloids and during this study, and

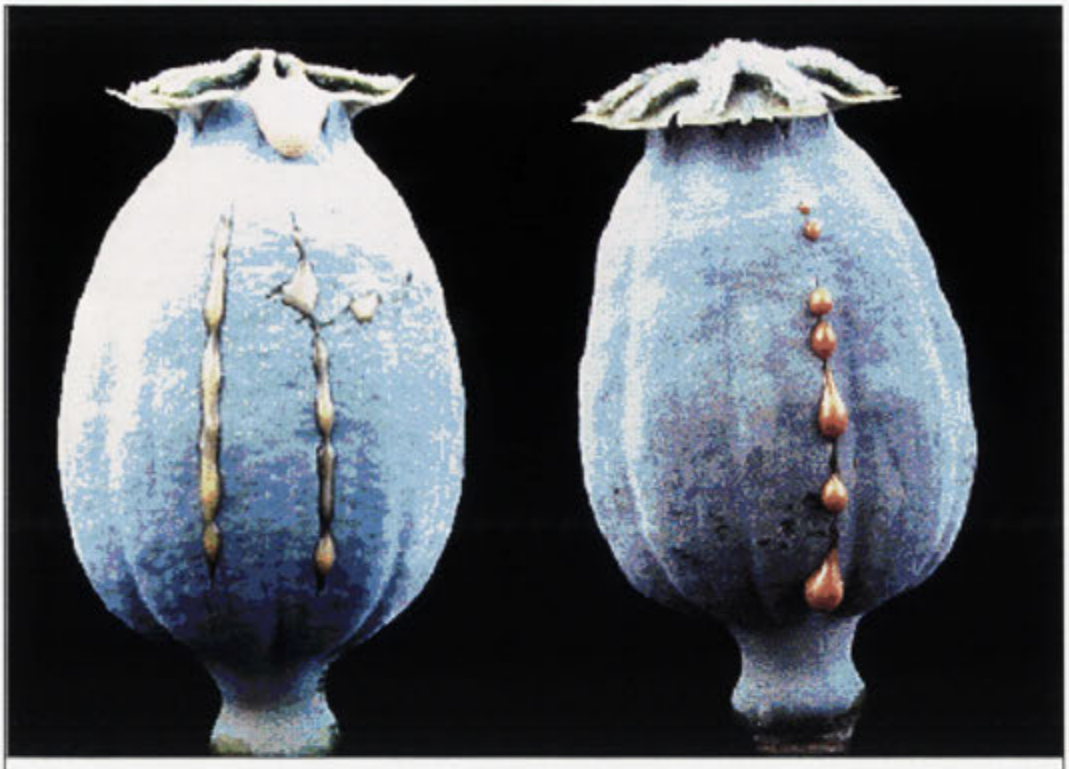


Figure 3.11: Picture of lanced mature capsules from Control (C052) and the *top1* mutant illustrating the pigmented latex from the mutant. The orange coloured latex of the *top1* mutant is a reliable indicator of its phenotype (Millgate *et al.*, 2004).

there has never been observed a dissociation of the thebaine/oripavine phenotype and latex colour. If the colour is a result of the defect in morphine production it could reveal something about the nature of the mutation.

The coloured compound may itself be an alkaloid derivative of the thebaine or oripavine, which are accumulating in *top1*. It may be a product which is normally degraded or utilised but is caused to accumulate due to a direct or indirect effect of the *top1* mutation. Similarly it could be the product of a new coloured compound that an altered enzyme can now produce. In any case the determination of the coloured compound may be revealing of some aspect of the *top1* mutation or the regulation of the pathway.

If the compound is a new alkaloid that is not normally produced in *Papaver* or any species for that matter it suggests that it is possible to mutate some of the enzymes in the pathway to produce and accumulate new alkaloids in opium poppy. There is also the possibility that the mutation has created a change in the transport process allowing alkaloids to be in a compartment in which they are not normally found allowing an enzyme to create a new coloured alkaloid.

While there is some carotenoid (possibly small amounts of the red coloured carotenoid lycopene) present in *top1* latex, it is in similar amounts in the wild type (C052) latex (Figure 3.12). The pigment colour eluting from the HPLC column at approximately the 5 minute mark; by comparison lutein eluted at around 12 minutes. The *top1* coloured product had a peak absorption at 440 nm. As a carotenoid control, an extract of tomato skin was extracted in the same way and at the same time as the latex samples. This was to show a trace for lycopene (Figure 3.12). Carotenoids display 3 characteristic absorption peaks at 445 nM, 471 nM and 503 nM (Cunningham *et al.*, 1996); the *top1* coloured product did not have this triple absorption signature suggesting that the colour is not a carotenoid. .

Carotenoids are sometimes found as esters and saponification of the HPLC extraction was undertaken to see if it is possible to release esterified xanthophylls that may be contributing to the latex colour of the mutant. The resulting HPLC trace files of the saponified samples did not reveal any esterified carotenoids in either the mutant *top1* or control C052 latex thus eliminating the possibility that the pigmented latex of the mutant is due to a change in carotenoid profile.

The data demonstrates that the pigmentation of the *top1* mutants' latex is not from a carotenoid but could possibly be from an alkaloid or an alkaloid derivative. *P. bracteatum* accumulates thebaine as the dominant alkaloid and dry-weight yields of thebaine in the capsule have been recorded as high as 3.6% (Lalezari *et al.*, 1974; Küppers *et al.*, 1976) or even 4.5% (Corrigan and Martyn, 1981). Triploidy-induced cultivars have reported yields as high as 8.8% thebaine accumulation in the whole plant in the first growing season, however this yield

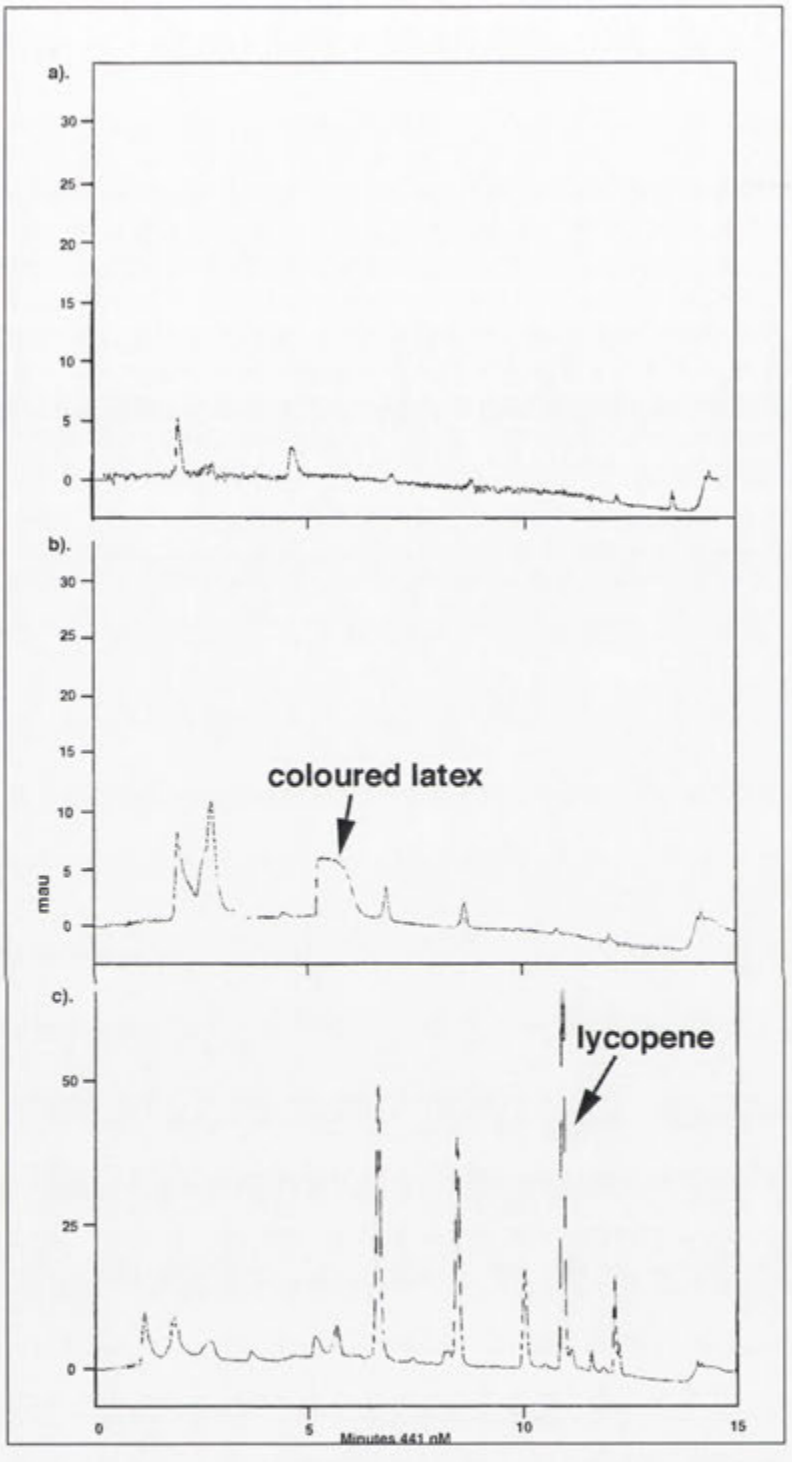


Figure 3.12: HPLC run of: a) C052 latex; b) *top1* latex; c) tomato skin run as a lycopene standard (Note the carotenoid sample is not on the same scale). Lycopene appears at approximately the 12-minute mark in the tomato sample. There is an unknown peak at around the 6 minutes mark. These peaks were not carotenoids and did not have the carotenoid signature three peaks.

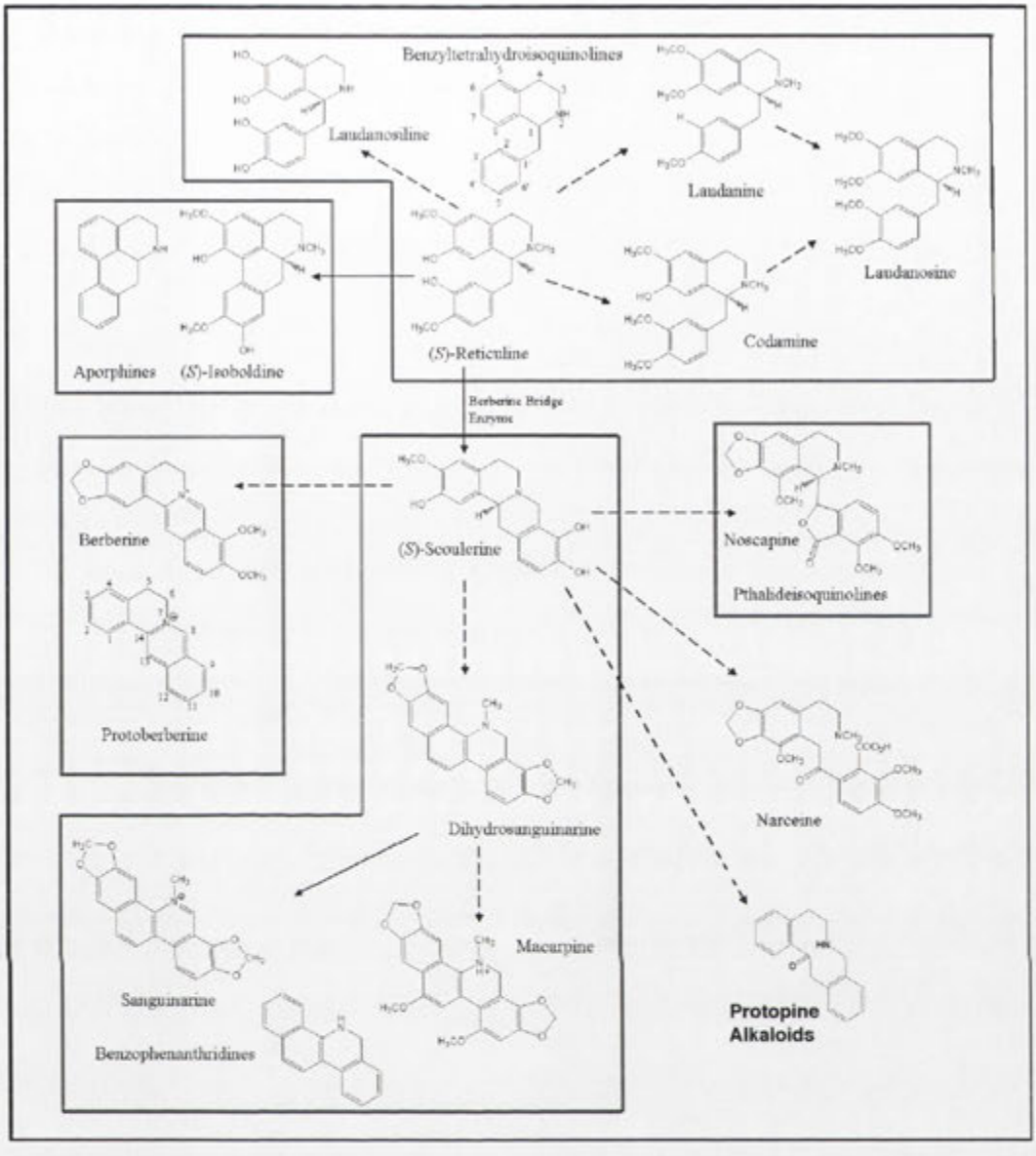


Figure 3.13: Protopine alkaloids are derived from (*S*)-scoulerine, and feeding studies have shown incorporation of labelled protopine into rhoeadine in *Papaver rhoeas* L. (Shamma and Moniot, 1978). Figure adapted from Facchini *et al.*, 1997 and Facchini *et al.*, 2000.

dropped as the plants entered their second year of growth (Milo *et al.*, 1987). *P. bracteatum's* latex is also highly pigmented.

Review of the literature revealed that there are several known alkaloids that occur in nature to give coloured compounds. *Sanguinaria canadensis* L., commonly known as the bloodroot plant, is a member of the *Papaveraceae* family and its roots produce an intensely red-pigmented juice, attributed to the accumulation of

the dominant alkaloids, sanguinarine and berberine. Sanguinarine is known to produce an intense red colour (Dostal and Slavik, 2000). Similarly, *E. californica* accumulates sanguinarine with other alkaloids in the roots, which have a characteristic yellow colour. There was no detectable increase in sanguinarine by TLC or HPLC analysis in the latex of the *top1* mutant as can be seen in Figures 3.9 and 3.10 respectively.

Rhoeadine and papaverrubine alkaloids are known to be responsible for the red colour developed by opium when treated with strong acids (Montgomery *et al.*, 1983). The rhoeadine alkaloids are derivatives of the protopine alkaloid structure (See figure 3.13) and two separate biosynthetic routes have been proposed (Shamma and Töke 1973; 1975). Rhoeadine alkaloids have been found in all *Papaver* species studied to date but the type and quantity of rhoeadine alkaloids can vary between taxonomic sections (Montgomery *et al.*, 1983). Table 3.3 lists the occurrence of various rhoeadine alkaloids in *P. somniferum* and *P. bracteatum*.

There are several methods of using colour-based reactions to detect thebaine or to differentiate thebaine from other alkaloids. These reactions usually involve altering the thebaine structure, require acidification and are not generally stable. The determination of the weak base thebaine often involves heating the sample with mineral acids to cleave the ether linkage, resulting in a phenolic hydroxyl group in position four of the morphinan alkaloid structure. This new phenol is amenable to the nitroso colour reaction and with the addition of 1% sodium nitrite solution gives a pigmented compound with an absorption peak near 450 nM (Ikonomovski, 1981). The colour fades after 20 minutes however suggesting the reaction is not stable. Frohde's reagent (0.5 gm of ammonium or sodium molybdate in 100 mL of concentrated H_2SO_4) added to thebaine gives an orange colour, however this reagent reacts with all of the morphinan alkaloids to give specific colours for the respective alkaloid (Vincent *et al.*, 1976).

Meconic acid (Figure 3.14) is a very strong acid that has long been known to be in the latex of *P. somniferum* and its presence usually signifies the presence of the morphinan alkaloids (Fairbairn and Williamson 1978). This strong acid can be

detected by spotting latex onto ferric chloride solution to yield a red-coloured complex. It is possible that this acid may be involved in alkaloid transport in and out of the latex vacuoles similar to how chelidonic acid is thought to work (Hauser and Wink, 1990) in trapping protoberberine and benzophenanthridine alkaloids in *Chelidonium majus* latex vacuoles. Meconic acid is possibly derived from chelidonic acid or synthesised from the same pathway as chelidonic acid. Tracer studies have suggested that chelidonic acid is derived from the condensation of phosphoenolpyruvate with a pentose phosphate. This in turn undergoes dehydration, dehydrogenation, ring closure and decarboxylation (Shen *et al.*, 2001).

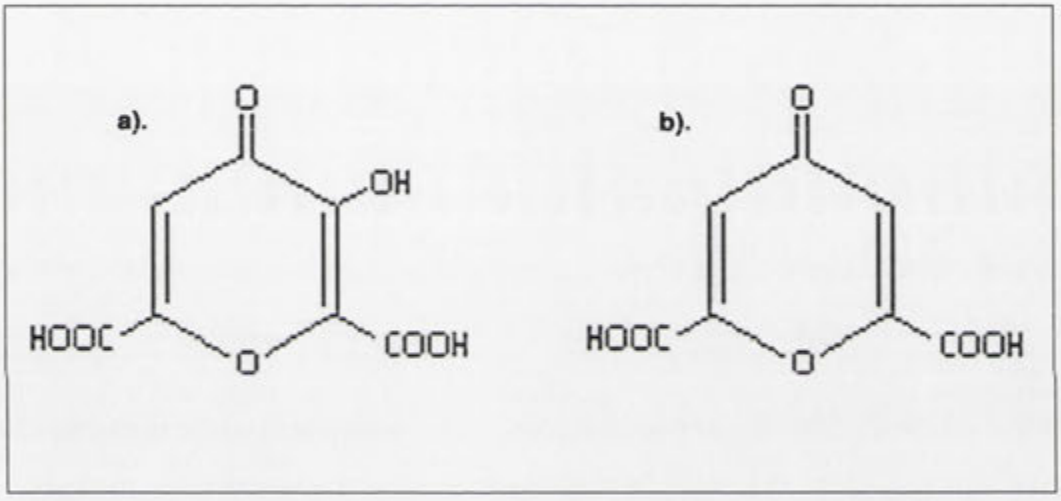


Figure 3.14: Structures of a) meconic acid and b) chelidonic acid.

Another possible basis for the pigmented latex could be the formation of dimeric alkaloids that could be initiated by the high concentration of thebaine. At least two dimeric alkaloids have been found in *P. somniferum*, pseudomorphine or 2,2'bimorphine (Bentley and Dyke 1959) and somniferine (Figure 3.15). Pseudomorphine is obtained by the oxidation of two morphine molecules and has been isolated from latex. It is highly unlikely that this is the compound responsible for the pigmented colour in the *top1* latex as morphine has been shown to be absent in the mutant. While this molecule is highly fluorescent (Darwin and Cone, 1980) it is unlikely it is responsible for the absorbance at around 450 nm. Somniferine contains a thebaine and a 14-hydroxymorphinone

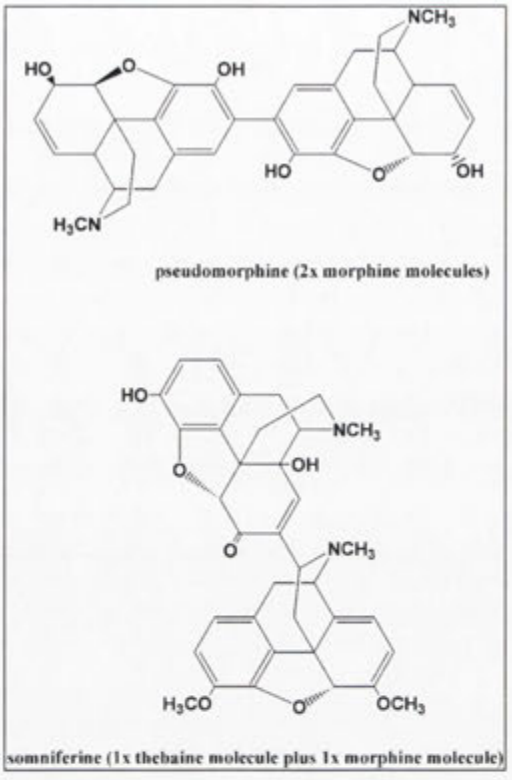


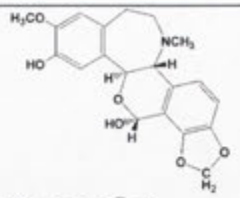
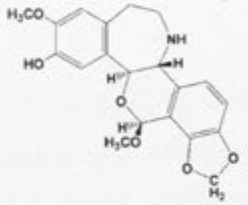
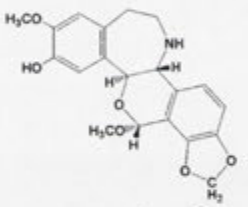
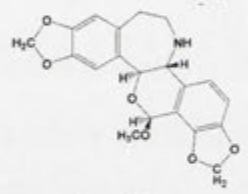
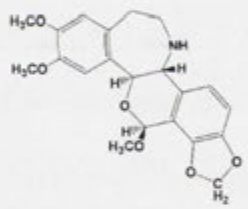
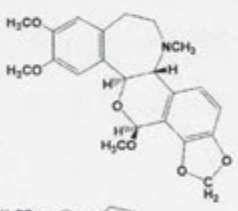
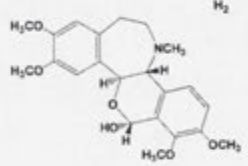
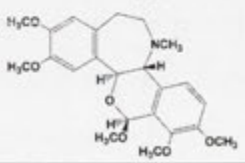
Figure 3.15: Dimeric morphinan molecules that have been found in *P. somniferum*.

---

unit, obtained via synthesis from morphine. The linkage is found between the C-16 carbon of thebaine and C-7 carbon of 14-hydroxymorphinone. The methyl ether has also been found (Dragar and Bick, 1988). This molecule was specifically isolated from the Tasmanian *P. somniferum* cultivars.

It is unlikely that this molecule is responsible for the colour change in the latex, as it has been found in the latex of the control plant. Another possibility could be that the enzymes responsible for these dimeric morphinan alkaloid molecules now have more thebaine substrate available and in the right cellular compartment to act on. This in turn could give new dimeric thebaine compounds that are coloured. To determine if this is the case further work needs to be undertaken. This would involve using HPLC to isolate the coloured fraction from the latex and then purifying this to homogeneity using columns. The purified compound can then be subjected to GCMS techniques to determine if the compound has the desired molecular weight of one or any of these dimer compounds.

Table 3.3: Rhoeadine alkaloids found in *P. bracteatum* and *P. somniferum*.

Rhoeadine Alkaloid	***P. somniferum***	***P. bracteatum***	**Structure**
N-methyl-14-*O*-demethyl-epiporphyroxine	+ (a,b,f)	-	
Papaverrubine D (Porphyroxine)	+ (a,b)	+ (a,c)	
Papaverrubine C (Epiporphyroxine)	+ (a,b)	-	
Papaverrubine E	-	+ (a,c)	
Papaverrubine B	-	+ (a,c)	
Glaudine	+ (a,b)	-	
Alpinigenine	-	+ (a,c,d,e)	
Alpinine	-	+ (c)	

Labels refer to references: (a)-Santavy 1970; (b)-Preininger *et al.*, 1981; (c)-Santavy 1979; (d)-Lavie *et al.*, 1981; (e)-Meshulam and Lavie 1980; (f)-Brochmann-Hanssen *et al.*, 1968.

3.3.2.2 Feeding Studies

[NB. This section describes experiments conducted by a collaborator in Germany. They are included here, rather than the introduction, to emphasise the logical and sequential characterisation of the mutant *top1*]

Professor Zenk at the Biozentrum Universität Halle, Halle, Germany, undertook feeding studies with the mutant *top1* (Millgate *et al* 2004). By feeding radiolabelled alkaloid precursors to the control C052 and the mutant *top1* plants it is expected to resolve where the blockage occurs in the pathway. By feeding radioactive intermediates prior to the blockage at thebaine and oripavine and measuring the resulting labelled end products it is expected to show which enzyme or enzymes are affected by the mutation. Using radioactive products fed to the plants after the blockage will yield the end products. If labelled compounds that occur before the blockage are fed to the plants then labelled intermediates between these precursors and the blockage should be found but not those occurring after the blockage.

The mutant *top1* had comparable activity (Table 3.4) of codeinone reductase to wild type poppy, C052 (Zenk, unpublished). When it was fed radiolabelled *R*-reticuline, *top1* accumulated label only in thebaine, whereas control poppy (C052) accumulated label in thebaine, codeine, oripavine and morphine. When fed radiolabelled thebaine, *top1* accumulated label only in oripavine, whereas control poppy (C052) accumulated label in codeine, oripavine and morphine. There was no difference when fed labelled codeine, both accumulating label in morphine. While control poppy (C052) could convert labelled oripavine into morphine, *top1* could not convert it at all. Taken together these feeding studies suggest a double block in the bifurcated pathway (Figure 3.6) between thebaine and neopinone and between oripavine and morphinone. Neopinone is not detected as a radiolabelled substrate as the conversion of neopinone to codeinone proceeds non-enzymatically (Gollwitzer *et al.*, 1993).

Table 3.4: Incorporation of isotopes after radiolabelled intermediate feeding.

Treatment Applied	Compound Measured	Munich Line	C052	*top1*
^{3}H-Reticuline	Thebaine	33	39	23
	Codeine	10	7	0
	Oripavine	0	4	0
	Morphine	7	19	0
	Compound X	0	4	0
^{3}H-Salutaridinol	Thebaine	35	14	20
	Codeine	9	6	0
	Morphine	14	11	0
^{14}C-Thebaine:	Codeine	64	10	0
	Oripavine	0	2	5
	Morphine	5	17	0
^{14}C-Oripavine:	Morphine	18	7	0
^{3}H-Codeine[a]:	Morphine	21	27	26

Compound X is an unknown labelled compound seen on the TLC plates.
[a] All treatments were left for 24hr except the ^{3}H-codeine treatment which was left for 72 hours. See above.

As the *top1* phenotype is semi-dominant, the *top1* mutation: may be in the gene responsible for the demethylation of thebaine and oripavine in the 6th position; or it may be a gene that controls the expression of the 6-*O*-demethylase; or possibly, an altered structural or transport component preventing the substrates, thebaine and oripavine, arriving at the appropriate sub-cellular compartment for the demethylation reaction. It is possible that the same enzyme performs both 6-*O*-demethylation reactions, since both pathways from thebaine to morphine are similar in the reaction process except in a different order. The only enzyme cloned from the morphinan pathway (Unterlinner *et al.*, 1999) is codeinone reductase and it has been shown to play a role in both pathways by stereo-selectively converting codeinone to codeine and morphinone to morphine (Lenz and Zenk, 1995b; Lenz and Zenk 1995c).

3.4 Germplasm with Altered Morphinan Alkaloid Yields

The last grouping of genotypes to be considered consists of those with altered morphinan alkaloid yields compared to the high morphine controls. These include: the *P. somniferum* cv Marianne (low morphine); the low alkaloid accumulating *P. somniferum* mutant D205; the high total morphinan mutant *P. somniferum*, D242; and the low morphinan but high papaverine alkaloid accumulating, *P. somniferum*, D120 (Figure 3.16).

3.4.1 *P. somniferum* cv. Marianne

P. somniferum cv. Marianne has been used previously in Canada for the study and isolation of several genes in the alkaloid pathway including TYDC (Facchini and De Luca, 1994), BBE (Facchini *et al*., 1996), 6'-*O*MT, 4'-*O*MT, *N*MT (Facchini and Park, 2003) and norcoclaurine synthase (Samanini and Facchini; 2001, 2002). Marianne is a high oil seed and low morphine producing variety, with much lower concentrations of morphinan alkaloids than any commercial varieties grown for morphine. This characteristic probably assisted in obtaining regulatory permission to grow the poppy and consequently much of the early work on morphinan alkaloid synthesis was undertaken on this variety. However *P. somniferum* cv. C052, a high morphinan alkaloid yielding cultivar grown commercially in Tasmania in recent years, and the *P. somniferum* cv. Marianne can be expected to differ in pathway transcript profiles.

3.4.2 *P. somniferum* Low Morphine Accumulating Mutant D205

D205 is an EMS generated mutant with low morphinan accumulation. HPLC analysis was initially undertaken by Dr Tony Fist at Tasmanian Alkaloids Pty Ltd to analyse the morphinan alkaloid yields of this mutant compared to the parental phenotype C052 on a capsule dry weight basis. Subsequent quantification of this mutant has confirmed that morphine is virtually the only detectable alkaloid, and the dry weight yield is half that of the CO52 progenitor (Table 3.1).

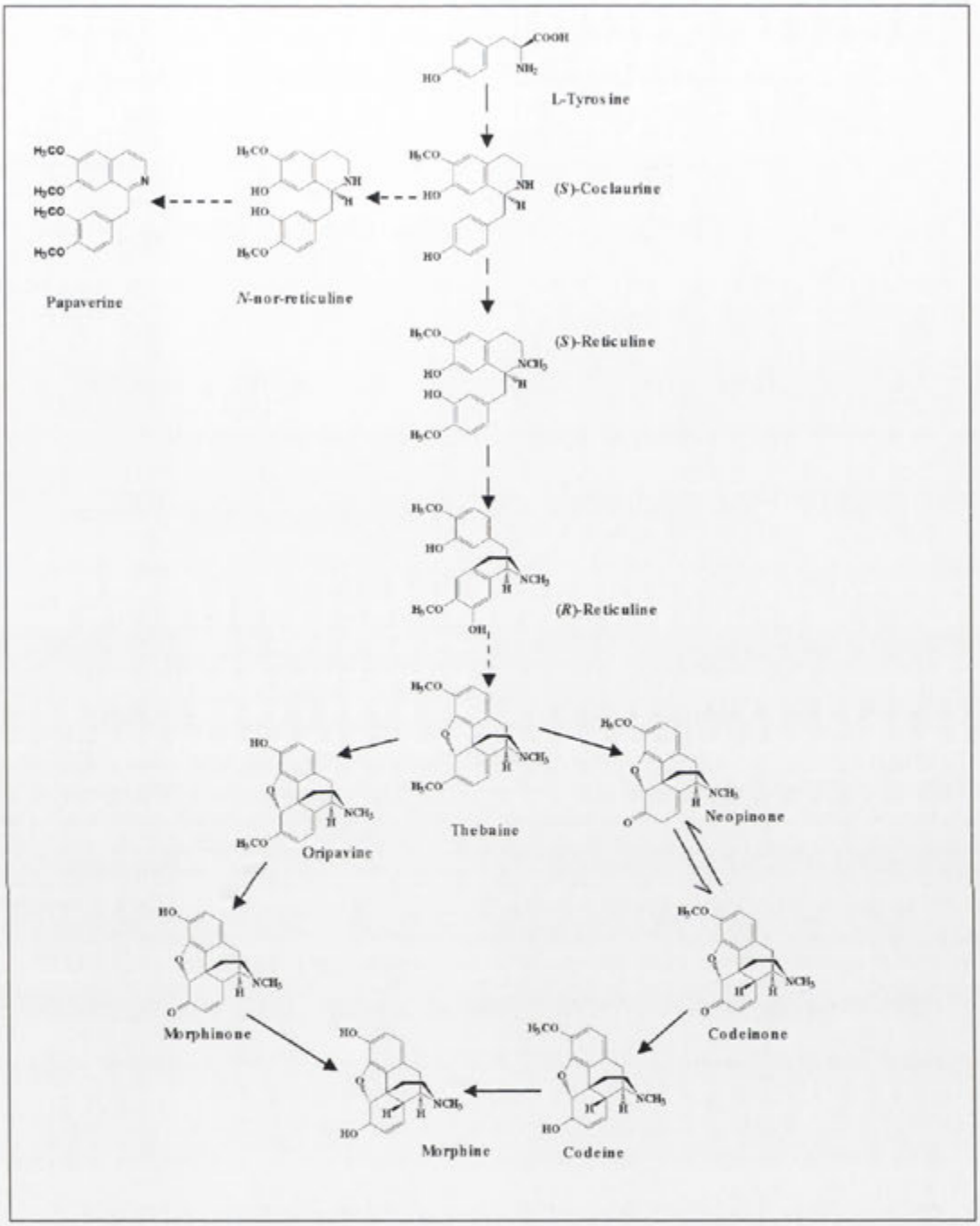


Figure 3.16: Abbreviated alkaloid pathway showing the accumulating alkaloids in the germplasm with altered morphinan alkaloid yields. The *P. somniferum* cultivar Marianne accumulates morphine and codeine but at much lower quantities than does the commercial *P. somniferum* cultivar used as control for this thesis. The low morphine mutant *P. somniferum* D205 similarly has lower levels of morphine and codeine to the wildtype *P. somniferum*. The mutant D242 accumulates thebaine and oripavine at higher levels than the wildtype whilst accumulating similar quantities of morphine and codeine. The last of the mutants in this chapter is that of D120 which has lower yields of morphine and codeine

but also accumulates papaverine at much higher yields, which are usually at very low levels in the parental commercial *P. somniferum* strain. Adapted from Kutchan 1998.

The mutation evidently imposes a rate-limiting constraint on morphine biosynthesis. It could be speculated that the mutation reduces the efficiency of an enzymatic step or affects the transport of intermediates between cell compartments of between different cell types. A positive regulator of the pathway could have been down-regulated or knocked out completely.

3.4.3 *P. somniferum* High Total Alkaloid Accumulating Mutant D242

Another mutant to be used in the microarray analysis is D242 that has been shown to accumulate larger quantities of thebaine and oripavine while still accumulating morphine and codeine. It is difficult to illustrate this yield difference using TLC plates due to the difficulties of obtaining accurate quantities of latex (Figure 3.17). However this can be shown by using small scale trials and measuring the content of the morphinan alkaloids using HPLC analysis and comparing this to the dry weight of the capsule. Such analyses have shown D242 accumulates 5x more thebaine and 15x more oripavine than the parental CO52, while still accumulating similar quantities of morphine and codeine (see Table 3.1).

This mutant has not been characterised any further than this preliminary analysis, and the cause of the phenotype is unknown. It is possible that a negative regulator of the pathway, before thebaine and oripavine accumulation, has been turned off or down so that it increases the flux through the pathway. The build up at the thebaine and oripavine step could be a compartmentalisation problem such that these intermediates can build up in higher concentrations but codeine and morphine cannot.

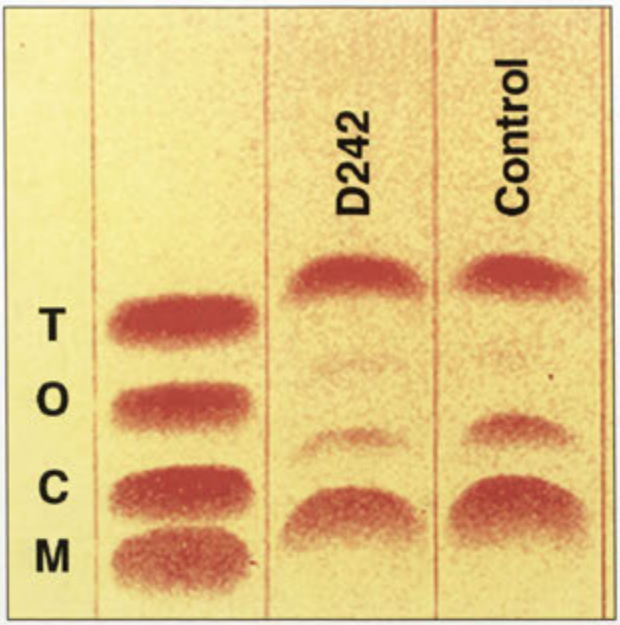


Figure 3.17: TLC showing the phenotype of the high oripavine and thebaine mutant D242. This mutant still accumulates morphine and codeine but not to the same extent as the parental control, C052. The overall yield of morphine, oripavine, codeine and thebaine is higher for the mutant than in the parent line. This is not evident in the TLC however HPLC analysis undertaken by Tasmanian Alkaloids Pty Ltd has shown that this mutant reliably accumulates more thebaine (about 5x) and oripavine (about 15x) than the parental control phenotype while still accumulating similar quantities of morphine and codeine (see Table 3.1).

3.4.4 *P. somniferum* Low Morphinan Alkaloid Mutant High Papaverine Accumulating Mutant D120

The papaverine-accumulating mutant (D120) produced much higher concentrations of papaverine than is ordinarily found in the wildtype phenotype and there is enough produced in the latex of these mutants to visualise the papaverine on a TLC plate (Figure 3.18). The resulting phenotype does however produce lower concentrations of morphine and thebaine (Table 3.1) suggesting that the mutation redirects metabolites toward papaverine but does not entirely block the pathway to morphine. Other than HPLC and TLC analyses, no other characterisation has been undertaken on this mutant.

It is possible that the mutation is affecting the enzyme responsible for the

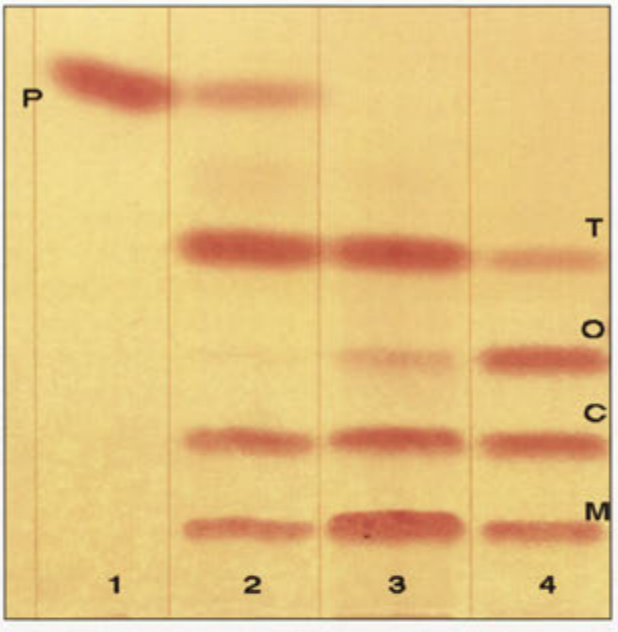


Figure 3.18: TLC showing the accumulation of the various alkaloids in the papaverine-accumulating mutant D120. Papaverine (P) standard (Lane 1) runs much higher up the TLC plate than the other alkaloids that accumulate in the opium poppy. D120 (Lane 2) still accumulates thebaine, codeine and morphine similar to the control plant (Lane 3). Lane 4 shows the morphine (M), codeine (C), oripavine (O) and thebaine (T) standards from the bottom to the top of the TLC plate.

---

methylation of the intermediate (*S*)-coclaurine, (*S*)-Coclaurine *N*-methyltransferase. This could be happening in a number of ways. Firstly, the mutation may be structural and affect the catalytic activity of the enzyme leading to a greater availability of (*S*)-coclaurine for the production of papaverine and less substrate, (*S*)-*N*-methylcoclaurine, available for the production of the morphinan alkaloids. The mutation could be in a promoter region of the enzyme or a transcription factor affecting the amount of message and thus enzyme available for the reaction. The mutation could be in a transport process for either the enzyme or the substrate.

3.5 Summary

This chapter has highlighted and described three groups of germplasm, which are to be further examined in subsequent chapters by global transcript analysis. These genotypes differ in some aspect of alkaloid profile and include mutants, natural variants, a related species and a transgenic.

By using the mutagen EMS it has been possible to develop several lines of *P. somniferum* that have different morphinan alkaloid profiles. Several of the mutants have altered morphinan alkaloid profiles such as the mutant F075, which only accumulates the morphinan thebaine and the mutant *top1*, which only accumulates thebaine and oripavine. Three other mutants have been characterised using HPLC and TLC analysis in this thesis that include: one that produces papaverine in greater quantities than the parent (D120); another that has a higher total morphinan alkaloid yield with an increase in thebaine and oripavine production (D242); and a low morphinan alkaloid yielding mutant (D205). These mutants also include the mutant E40/41 that accumulates the benzylisoquinoline alkaloid reticuline and contains a blockage to the morphinan alkaloids. The hairpin RNAi silenced codeine reductase transgenic lines were shown to accumulate the same benzylisoquinoline alkaloids as the reticuline mutant (E40/41) suggesting that there are complicated levels of control upon the pathway and that much is still to be learned from studying this pathway.

The characterisation of the different germplasm using HPLC and TLC analysis enables us to clearly see where the blockages are occurring in the morphinan pathway and where there are alkaloids accumulating at much higher concentrations due to blockages further down the pathway.

The three groupings are intended to emphasise at a gross level in which portion of the morphine pathway they differ. It was necessary to characterise several of the mutants and the hairpin RNAi *P. somniferum* plants in advance of the transcript analysis. This initial characterisation has shown that a number of different perturbations can occur in the morphinan alkaloid pathway and suggests, not surprisingly, that there could be several layers of control operating. It was hoped that differential transcript comparisons across these different groupings, and the

branchpoints they may represent,might add interpretive power to the transcript analyses and help eliminate inconsequential changes and highlight shared changes of significance. The next chapter will now follow up on this classical characterisation and use cDNA microarrays to determine global gene expression changes between the different germplasms and a control *P. somniferum* plant. Those genes showing altered expression levels will then be sequenced and further analysis, where possible undertaken to determine gene function.

Chapter 4

Gene Expression Analysis Using cDNA Microarrays Of Germplasm With Different Alkaloid Profiles To *P. somniferum*

4.1 Introduction

A number of *P. somniferum* EMS generated mutants have been developed and made available by Tasmanian Alkaloids Pty Ltd. HPLC analysis was used to screen the M2 population for altered morphinan alkaloid profiles (Table 3.1). As mentioned in Chapter 1 only a few morphinan biosynthetic genes have been cloned, thereby limiting options for direct molecular analysis of mutants. It is difficult to predict whether a mutation is in: a regulatory pathway; transport process; structural changes in particular cell types; or in the gene coding for the enzyme immediately after the blockage itself. Using the combination of mutants and different alkaloid accumulating germplasm described in chapter three, this chapter will employ global transcript analyses to explore differences between the different germplasm and the control *P. somniferum* plant.

Initial analysis of the *top1* phenotype using 2D Gel electrophoresis has shown that there are more than eight regions of protein changes in the mutant's latex compared to wildtype *P. somniferum*. These results give reason to believe that there would also be transcriptional changes in the mutant phenotype. It is anticipated that a mutational blockage will have a cascade of pleiotropic effects on other genes, which may include those encoding transcriptional regulators, transporters and enzymes. By looking at the mutant phenotype through global transcript analysis, and by looking for differential genes shared between the various genotypes, it is hoped to create a greater understanding of the morphinan alkaloid pathway.

Microarray analysis using cDNA clones has been selected to study the different germplasm. This method has been chosen due to its mass screening capabilities

to study a large number of transcripts expressed in the poppy genome. There are few annotated sequences in the database from opium poppy making it difficult to utilise oligo arrays. By printing the cDNAs on a microarray and studying their relative expression levels between the *top1* mutant phenotype and the control phenotype a large amount of information can be obtained relatively quickly. Another advantage of this method is that the cDNAs are stored and readily available for subsequent sequencing and further expression studies.

The only studies of alkaloid gene expression using mutants have been in tobacco (Wang *et al.*, 2000). The most recent is that of the gene family encoding arginine decarboxylase. Normally these genes show an increase in expression after removal of the floral meristem and upper leaves and stem, however the low alkaloid-producing mutant did not (Wang *et al.*, 2000). It is possible that the mutation can affect the RNA expression of a gene involved in the morphinan pathway either directly, with a mutation in the promoter, or indirectly by affecting a regulator of the gene. *Arabidopsis tt* and *td* (Abrahams et al., 2003a,b) mutants affecting anthocyanin [and tannins] synthesis have shown that the detection in changes of expression levels of one or several of the genes involved in anthocyanin synthesis is possible. This is important as the alkaloid genes may be at low levels of expression in the plant cell and by screening larger numbers of expressed sequences the chances of missing any lowly expressed genes or transcription factors can be minimised.

There is no publicly available gene library of any type from *P. somniferum.* Therefore a cDNA library needed to be produced for the microarrays. To do this a high yielding morphinan alkaloid inbred commercial cultivar was used for the source material in order to create a cDNA library enriched in morphinan alkaloid structural and/or regulatory genes. Several studies have shown that alkaloid production is highest in the plant just prior to flowering and capsule development, so the aerial part of the plant of this stage of development was used for the cDNA library synthesis. A non-subtractive library approach was taken. A subtractive library can be useful in enhancing genes expressed at a medium level, but the process risks losing low copy genes. On balance it was judged

safest to make a high quality non-subtractive library and increasing the number of cDNAs on the slides.

See chapter 2 for a detailed description of the development of the microarrays and the subsequent analysis. 17, 000 cDNA clones were selected randomly from a directionally cloned non-normalised cDNA library prepared from the prehook stage of *P. somniferum* (Figure 1.9). Additionally, the slides included several *P. somniferum* alkaloid gene controls, some with non-intergenic (such as the BBE intron) sequences and several non-poppy genes as negative controls (See Subchapter 4.2). No study had previously been undertaken to determine if the enzymes or transcript encoded by these genes would be differentially regulated in the mutant phenotype compared to the parent.

A high degree of rigor was applied to all aspects of the data collection and statistical analysis. First, RNA was extracted from tissue and pooled from six plants for each of the biological replicates to ensure consistency of the biological material. Second, the microarray hybridisations were repeated, including technical replicates for each, which involved a fluorescent dye swap. Third, the fluorescence data derived from the microarray images were normalized using tRMA software (see Chapter 2.6). Normalisation of the data is used for a number of reasons. These include: to remove labelling differences between the two dyes; fluorescing differences between the two dyes; spatial biases that occur between spots in the middle of the slide compared to those on the outside; hybridisation and washing inconsistencies across the slide; and differences in spotting of the DNA on the slide.

This chapter will first discuss the controls printed on the array and then describe each mutant or treatment used in the microarray experiments. The resulting differentially expressed genes will be listed and several of the sets of genes will be sequenced and compared to the database. Again these experiments on various genotypes will be placed in the same groupings as listed in Chapter 3 for cross-referencing and comparisons of the differential genes. Their potential involvement in the morphinan pathway will be discussed.

4.2 Microarray Controls

Various controls were printed on the microarray slides at the same time as the *P. somniferum* cDNA library clones. They were replicated and placed at randomised positions. Several different heterologous alkaloid-encoding genes were PCR amplified for use as positive specificity controls (Table 4.1A). These included: the BBE and cyp80b1 coding genes from *E. californica*; the 6' and 4' *O*MT coding genes from *C. japonica*; and genes coding for TYDC, BBE (genomic and cDNA), cytochrome P450 reductase isoforms 1 and 2, and codeinone reductase 1.1 and 2.1 were all from *P. somniferum*.

These were the only alkaloid gene controls that could be obtained at the time the arrays were printed. These control genes are not necessarily positive controls in the sense that they may not be differentially expressed in the mutants and genotypes relative to wild type poppy, As reviewed in Chapter 1.10: BBE (Dittrich and Kutchan; 1991); cyp80b1; TYDC; 6' *O*MT; and 4' *O*MT are all inducible upon elicitor treatment in opium poppy cell suspensions (Facchini and Park; 2003), while codeinone reductase is not (Huang and Kutchan; 2000). However even these inducible genes are not necessarily expected to be differential in any of the microarray experiments. Little is known about these genes' expression patterns, other than their varying expression levels in root, stem, leaf and flower (Facchini and Park; 2003). The intergenic region of the BBE gene and its promoter will be beneficial as negative control in assuring us that genomic DNA is not contaminating the RNA used for labelling the arrays. The alkaloid gene controls were also used to give reference points for grid orientation.

P. somniferum, *E. californica* and *C. japonica* alkaloid gene controls were PCR amplified for printing from cDNAs cloned using primers designed to Genbank sequence data, except for the *P. somniferum* BBE full-length clone with promoter region, kindly provided by Professor Facchini (Facchini et. al., 1996). The cDNAs for *Coptis japonica S*-adenosyl-L-methionine: norcoclaurine 6-*O*

Table 4.1A: The list of alkaloid genes used on the microarray.

Alkaloid Pathway Genes	Accession Number	Reference
BBE (*E. californica*)	AF005655	Hauschild et. al., 1998
BBE (*P. somniferum)* (whole genomic clone including intron)	AF025430	Facchini et. al., 1996
BBE promoter (*P. somniferum*)	AF025430	Facchini et. al., 1996
BBE (*P. somniferum*)	AF025430	Facchini et. al., 1996
TYDC (*P. somniferum*)	U16804	Maldonado-Mendoza et. al., 1996
Cytochrome P450 reductase (CPR) (*E. californica*)	U67186	Rosco et. al., 1997
Cytochrome P450 reductase Isoform 1 (*P. somniferum*)	U67185	Rosco et. al., 1997
Cytochrome P450 reductase (CPR) Isoform 2 (*P. somniferum*)	-	
6' *O*-methyl transferase (*C. japonica*)	D29811	Morishige et. al., 2000
Cyp80b1 (*E. californica*)	AF014801	Pauli and Kutchan, 1998
4' *O*-methyl transferase (*C. japonica*)	D29812	Morishige et. al., 2000
Codeinone reductase (cor 2.1) (*P. somniferum*)	AF108438	Unterlinner et. al., 1999
Codeinone reductase (cor1.1) (*P. somniferum*)	AF108432	Unterlinner et. al., 1999

Table 4.1B: The list of clones used as negative controls on the microarray slides.

Negative Controls	Accession Number
Bacillus Thuringiensis Cry 1a BT protein	U89872
Phosphinothricin acetyl transferase	X17220
Green Fluorescent Protein	AF078810
Hemoglobin, beta (HBB)	NM_000518
Beta-glucuronidase (uidA)	A00196
Hygromycin phosphotransferase	K01193
Luciferase	X65316
Kanamycin (neomycin phosphotransferase)	V00618
B-cell receptor	AF126021
Myosin Heavy Chain	X13988
Myosin regulator Light chain 2	M21812
Growth Factor Insulin-like	X0607868

methyltransferase (D29811: Sato et. al., 1994) and *S*-adenosyl-L-methionine: 3'-hydroxy-*N*-methylcoclaurine 4'-*O* methyltransferase (D29812: Morishige et. al., 2000) were kindly provided by Professor Sato.

Negative controls, other than the BBE promoter and intergenic regions, were added several times randomly across the array. These included several reporter genes not found in poppy and some human gene sequences (Table 4.1B). The inclusion of several non-poppy sequences on the array was intended to give confidence in the stringency of the hybridising process; they should demonstrate that there was no unspecific cross hybridisation. The neomycin phosphotransferase II (*nptII*) gene placed on the array will served as a non-poppy sequence negative control for all of the experiments except for the *nptII* containing transgenic *P. somniferum* plant. In this instance the *nptII* served as a positive control; it was expected to be differential in the transformed poppy. Other negative controls placed on the array randomly where DMSO and water controls to also check for blocking efficiency and unspecific hybridisation.

4.3 Reticuline Accumulating Germplasm

4.3.1 *Eschscholzia californica*

E. californica plants were grown under similar conditions to *P. somniferum* plants and plants were harvested pre-flowering, as close as possible to the equivalent stage in *P. somniferum*. RNA extraction and labelling experiments were undertaken as described in Chapter 2. Two microarray slides were used for the analysis. This included a technical replicate for the one biological sample. After analysis of these two experiments, a large number of genes (approximately 16% of the array) were differentially expressed between the *P. somniferum* and *E. californica* plant materials (Figure 4.1) using the cut-off of 2. The entire set consisted of 2692 differentially expressed genes. Remembering that only one biological replicate was included in the analysis, several of the randomly spotted control genes were also found to be differential between the two species (Table 4.2). These differentially expressed clones included five of the negative controls

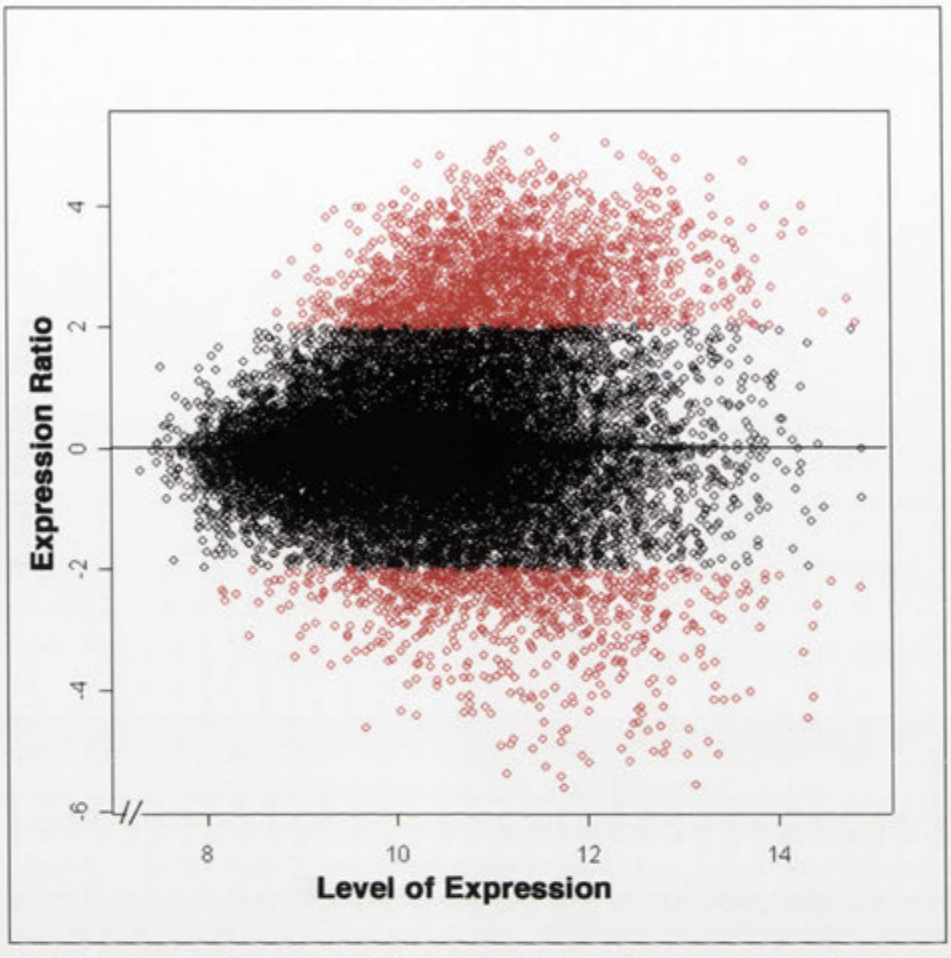


Figure 4.1: Graphical representation of the differentially expressed genes between *E. californica* and *P. somniferum*. A total of 2692 genes were differentially expressed (red circles). Out of this total, 1907 genes were more highly expressed in the *E. californica* plant material compared to the *P. somniferum* plant. In comparison only 785 genes were found to be upregulated in the *P. somniferum* control plant.

placed on the array. Each of these clones were spotted at least 7-8 times across the arrays, however only one spot for each of these cDNAs showed up as being differential. This was the case for controls encoding: neomycin phosphotransferase; Insulin-like Growth Factor; Myosin Heavy Chain, and the Green Fluorescent Protein. This result suggests that there must be some labelled transcripts with slight similarities to bind to these sequences even at the stringency of the hybridisation employed. It may be assumed that an increased number of biological replicates would have eliminated these spurious control results and also would have reduced the total number of differential genes.

Table 4.2: Control genes that were found to be differential between *E. californica* and *P. somniferum*.

Control Gene Name	Accession Number	Expression Ratio	More Highly Expressed In
Bacillus Thuringiensis Cry 1a BT protein	U89872	4.8	*P. somniferum*
		4.1	*E. californica*
		3.2	*E. californica*
Kanamycin (neomycin phosphotransferase)	V00618	3.2	*E. californica*
Growth Factor Insulin-like	X0607868	2.5	*E. californica*
Myosin Heavy Chain	X13988	2.6	*P. somniferum*
Green Fluorescent Protein	AF078810	3.5	*E. californica*
6' *O*-methyl transferase (*C. japonica*)	D29811	2.5	*E. californica*
		3.1	*P. somniferum*
TYDC (*P. somniferum*)	U16804	3.9	*P. somniferum*
		3.3	*P. somniferum*
Cytochrome P450 reductase Isoform 1 (*P. somniferum*)	U67185	3.0	*P. somniferum*
		2.9	*P. somniferum*
		2.5	*P. somniferum*
		2.5	*P. somniferum*
		3.2	*P. somniferum*
BBE promoter (*P. somniferum*)	AF025430	3.0	*P. somniferum*
		2.8	*P. somniferum*
		2.5	*P. somniferum*

The merged cells of the Table show how many replicate spots of the cDNA proved to be differentially expressed in the experiment. Since there were eight replicate spots of these controls, not all of the replicate spots were differentially expressed in the experiment.

Another negative control, the cDNA coding for the *Bacillus Thuringiensis* Cry 1a BT protein, was differentially expressed between the two species also. In this case three of the eight cDNAs randomly spotted on the array were differentially expressed. This differential expression was not found to be uniform, as two of the spots were more highly expressed in *E. californica* than in *P. somniferum* and the third cDNA spotted on the array was more highly expressed in *P. somniferum* (Table 4.2). This is further evidence that two slides were insufficient for reliable conclusions. One of the positive controls, 6' *O*-methyl transferase (*C. japonica*), also showed this type of differential expression where one cDNA was morehighly expressed in *P. somniferum* and the same cDNA spotted in a different region of the slide appeared more highly expressed in *E. californica* (Table 4.2). Another positive control, TYDC (*P. somniferum*), was also

differentially expressed in this experiment (Table 4.2), however it was only one of the eight randomly spotted cDNAs.

Despite the limitations of the *Eschscholzia* experiments there are other indications of its general validity. Two isoforms of the Cytochrome P450 reductase have been isolated from *P. somniferum* (Rosco et. al., 1997) while only one has been isolated from *E. californica* (Rosco et. al., 1997). The Cytochrome P450 reductase isoform 1 from *P. somniferum* seems to be unique in this *Papaver* species and the majority of cDNA spots of this gene (five of eight) placed randomly across the array were more highly expressed in *P. somniferum* than in *E. californica* (Table 4.2). The remaining 3 spots of this control cDNA may have been in a region with generally poor hybridisation. This result gives confidence in the arrays that detecting differentially expressed transcripts is possible even in an experiment where the number of replicates and experiments are fewer than desirable.

One anomaly that is difficult to explain is the differential expression of the *P. somniferum* BBE promoter region in half the replicate spots. Four of the eight genomic promoter regions spotted on the array were more highly expressed in *P. somniferum* than in *E. californica* (Table 4.2). This suggests that there was possible DNA contamination in the extracted RNA used in the experiment, from *P. somniferum* cv. C052. PCR analysis has been used to check the RNA extracted for this experiment using primers designed to the gene codeinone reductase and to the BBE promoter and neither set of primers gave a positive band in the RNA sample, while *P. somniferum* genomic DNA was positive as expected. Closer inspection of the BBE promoter region confirms it does not seem to have an internally transcribed region as there are several stop codons in all three frames in both directions suggesting that while an RNA transcript may be produced it is unlikely that functional protein is. The remaining potential explanation is that there is another poppy coding region with sufficient homology to the BBE promoter.

4.3.2 *P. somniferum* Reticuline Accumulating Mutant E40/41

Plants were grown and RNA prepared for microarray experiments of the reticuline accumulating mutant E40/41 and the parental genotype W52-12-10 with the same procedures and conditions detailed in Chapter 2. The experiments consisted of two biological replicates each with their own technical replicate.

Analysis of the microarray experiments revealed a relatively small subset of genes that were differentially expressed between the two cultivars (Figure 4.2). Only two genes were significantly under-expressed in the control W52-12-10 line compared to E40/41. Only 16 cDNAs were significantly under-expressed in the reticuline mutant compared to the parental line. The expression ratios were relatively low for all of the clones (Table 4.3). Only one clone, 11F5, was found to have a differential expression ratio greater than 2. When sequenced it was found to match a gene encoding a chlorophyll a/b binding protein (Light harvesting complex a gene 3.1). This gene encodes the LHCI-680 protein of the light-harvesting complex I (LHCI) (Wang *et al.*, 1994). Such a gene is very unlikely to be responsible for the reticuline accumulating phenotype. It's apparent differential expression is more likely a secondary effect of the mutation or due to another mutation in E40/41 unrelated to the reticuline accumulating phenotype. The differential expression ratios and resulting BLAST x values are listed in Table 4.3.

Several of the differential genes reticuline mutant E40/41 (11F5, 24B2, 11B2, 153C2, 209F1 and 81A11) were homologous to photosynthetic enzymes. There were four differentially expressed cDNAs that (33F5, 91B6, 112B7, 27D12) had no substantial homology to any sequences in the NCBI databases. Two differentially expressed cDNAs (86B7 and 86B2) had reasonable homology to two different *Arabidopsis* clones from the database, however the Arabidopsis clones have no known function.

One cDNA (7E1) was homologous to the *Arabidopsis* nicastrin precursor which has a known function in *Caenorhabditis elegans* as a transmembrane glycoprotein that is part of a multimeric complex responsible for

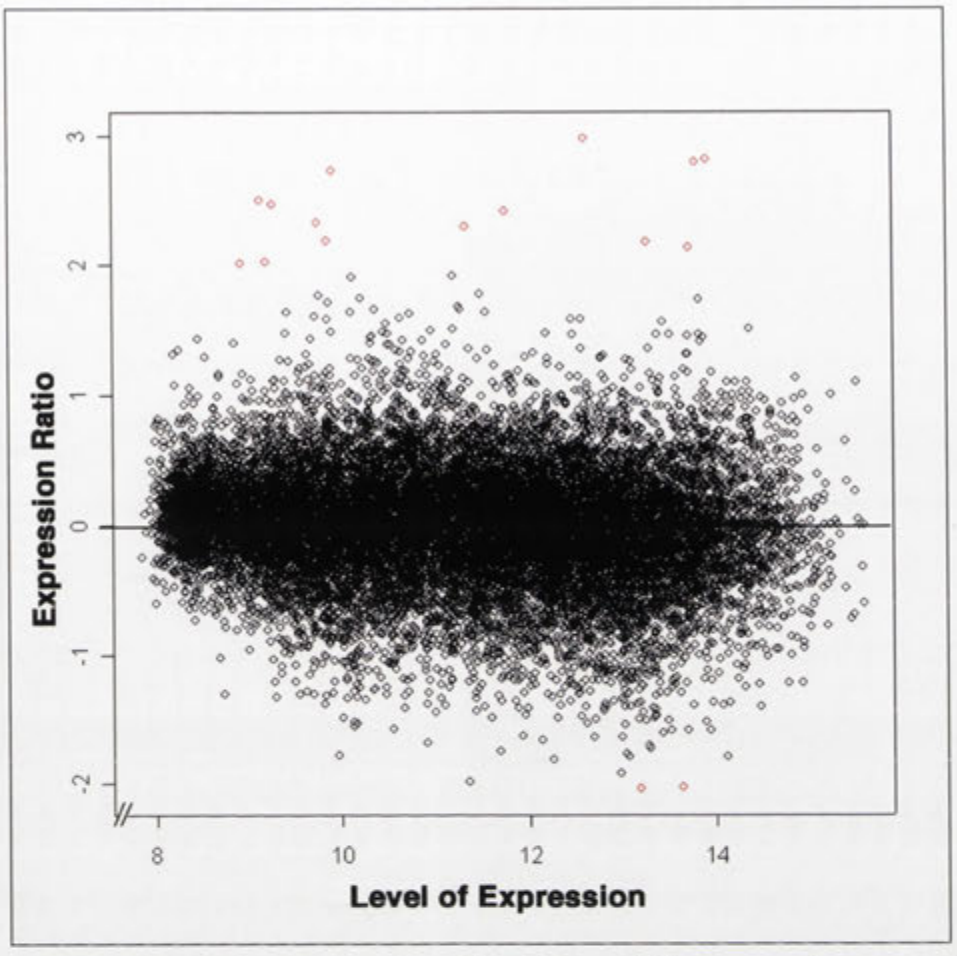


Figure 4.2: Graphical representation of the differentially expressed genes (red circles) of the high reticuline-accumulating mutant, E40/41 compared to the parental control W52-12-10.

intramembraneous proteolysis of proteins such as Notch/GLP-1 and beta amyloid precursor protein (Nishimura *et al.*, 2000). However its role in plants is yet to be determined. Another interesting cDNA from these experiments is 199E8, which was homologous to an *Arabidopsis* gene with accession number AF410286. The *Arabidopsis* clone is a transporter-related cDNA that has several transmembrane domains and is possibly involved in the vesicular pathway of protein sorting. It is possible that the poppy cDNA (199E8) may be involved in determining where one or more of the enzymes involved in the alkaloid pathway are located. It could be speculated that under-expression of 199E8 in the mutant could prevent the enzyme 1,2-dehydro reticuline synthase to be sorted to the correct sub-cellular compartment and resulting in an accumulation of (*S*)-reticuline.

Table 4.3: Clones differentially expressed in the reticuline mutant E40/41 compared to the control.

Name	Expression Ratio	Accession Number	Description	E-value
11F5	2.1	AF139465	*Vigna radiata* LHCII type III chlorophyll a/b binding protein (CipLhcb3) nuclear gene for chloroplast product.	1E-161
7E1	1.9	NP_190832	*A. thaliana* nicastrin precursor.	1E-167
86B7	1.9	BT002529	*A. thaliana* putative protein (At3g53720) mRNA.	3E-56
199E8	1.8	AF410286	*A. thaliana* At3g14410/MLN21_19 mRNA. *Transmembrane transporter*	1E-132
24B2	1.8	AF014052	*N. tabacum* Mg protoporphyrin IX chelatase (Chl H) mRNA.	0
33F5	1.7	-	No Significant hits	-
11B2	1.7	Z50099	*S. tuberosum* mRNA for transketolase.	0*
86B2	1.5	NP_200634	*A. thaliana* expressed protein	1E-77
91B6	1.1	-	No Significant hits	-
153C2	1.1	AF196292	*Apium graveolens* nonreversible glyceraldehyde-3-phosphate dehydrogenase mRNA.	0
112B7	1.0	-	No Significant Hits	-
27D12	1.0	-	No Significant hits	-
111C7	0.9	AY093119	*A. thaliana* Apospory-associated protein c (hypothetical 35.4 kDa protein).	4E-52
108H11	0.8	AY099862	*A. thaliana* Potential phospholipid-transporting ATPase 9 (ec 3.6.3.1)	0
209F1	-1.3	O04687	*Mesembryanthemum crystallinum* chlorophyll a/b-binding protein.	1E-121
81A11	-1.4	O04687	*Mesembryanthemum crystallinum* chlorophyll a/b-binding protein.	1E-119

All values are BLAST x values except for the * 11B2 which is a tBLAST x value.

4.3.3 Hairpin RNAi Silenced Codeinone Reductase *P. somniferum*

Three different transgenic lines (215-1, 220-1, 220-2) (Allen *et al.*, 2004) were grown and RNA harvested as per Chapter 2. RNA extraction and labelling experiments were undertaken as described in Chapter 2. The RNA from the three transgenics was used as three biological replicates and for each a technical replicate (dye swap) was employed. This gave a total of six slides however one

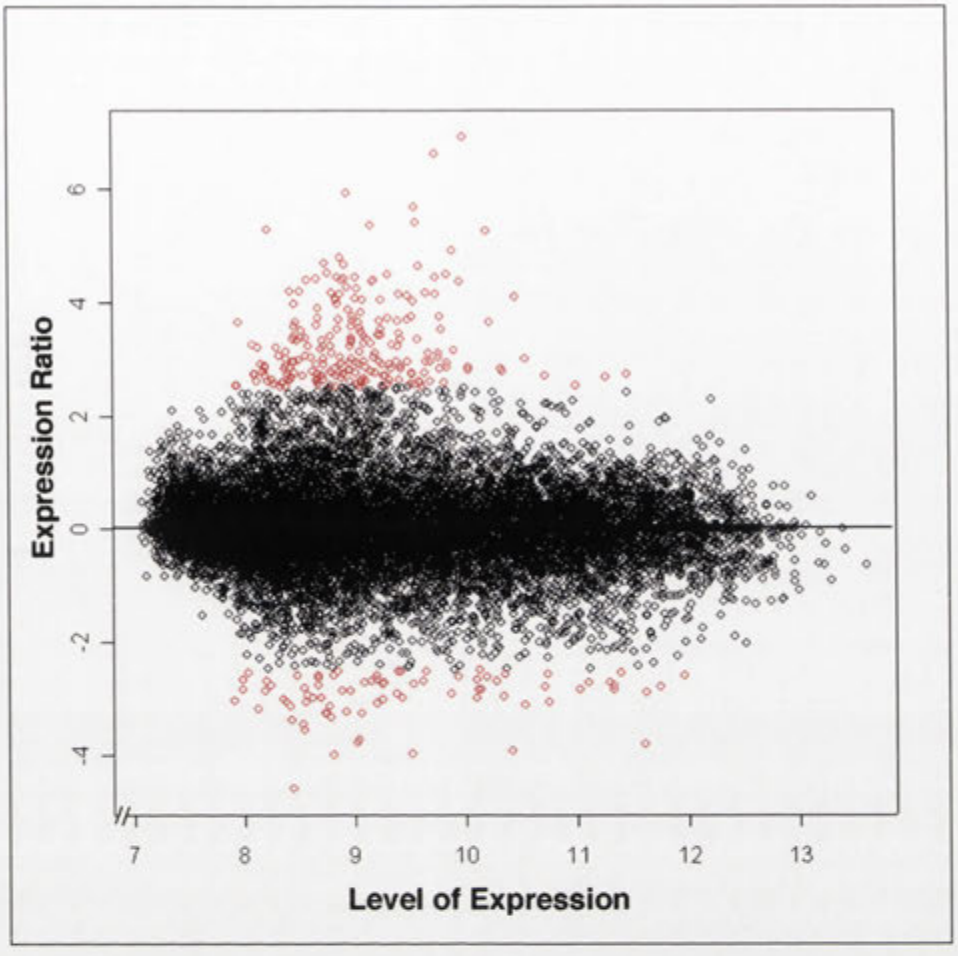


Figure 4.3: Graphical representation of the differentially expressed cDNAs in the transgenic codeinone reductase RNAi plants compared to the parental phenotype C058-34. The arbitrary cut-off of 2 was used to select the differential genes (red circles).

---

slide for the transgenic line (220-2) was lost due to a poor Cy3 labelling reaction and this left a total of five slides for the microarray analysis.

Analysis of the five-microarray experiments revealed a number of genes that were differentially expressed in the transgenic lines compared to the parental phenotype C058-34 (Figure 4.3). A total of 31 clones were under-expressed in the transgenic and 15 clones were differentially over-expressed in the mutant using the cut-off value of 2 (Table 4.4). All 46 differential clones were sequenced from both ends in duplicate and used for the BLAST analyses (Table 4.4). A good positive control that was used in this experiment was the NPTII gene that was used as a selectable marker in the transformation of the RNAi-cor

Table 4.4: Clones that were differentially expressed between the transgenic codeinone reductase RNAi plants and the parental phenotype *P. somniferum* cv. C058-34.

Name	Expression Ratio	Accession Number	Description	E-value
12B4	3.2	-	No Significant Hits	-
40D12	2.9	PSAF_FLATR	*F. trinervia* photosystem I reaction centre subunit iii, chloroplast precursor	2E-59
55A5	2.7	-	No Significant Hits	-
34F10	2.6	-	No Significant Hits	-
37E1	2.4	RBS4_MESCR	*M. crystallinum* ribulose bisphosphate carboxylase small chain 4, chloroplast precursor (ec 4.1.1.39)	3E-76
30G3	2.3	-	No Significant Hits	-
117F9	2.3	-	No Significant Hits	-
30G3	2.3	-	No Significant Hits	-
87D1	2.2	O8136	*Prunus armeniaca* sulfite reductase	1E-150
52D5	2.2	PSU08599	*P. sominiferum* clone TYDC3 tyrosine/dopa decarboxylase mRNA.	0#
218E7	2.2	-	No Significant Hits	-
178D6	2.2	-	No Significant Hits	-
90B7	2.2	-	No Significant Hits	-
36D3	2.1	-	No Significant Hits	-
98C12	2.1	-	No Significant Hits	-
144C8	2.1	-	No Significant Hits	-
132H11	2.1	-	No Significant Hits	-
198B12	2.1	TBB2_ARATH	*A. thaliana* tubulin beta-2/beta-3 chain.	1E-74
81H6	2.1	PSAD_CUCSA	*Cucumis sativus* photosystem i reaction center subunit ii, chloroplast precursor (photosystem i 20 kda subunit)	8E-83
110C10	2.1	-	No Significant Hits	-
210D2	2.1	-	No Significant Hits	-
218C4	2.1	-	No Significant Hits	-
63E10	2.1	-	No Significant Hits	-
115E11	2.1	-	No Significant Hits	-
52H3	2.1	Q9XQB1	*Phaseolus aureus* (mung bean) lhcii type iii chlorophyll a/b binding protein.	1E-124
96B2	2.1	-	No Significant Hits	-
19C10	2.0	-	No Significant Hits	-
38A12	2.0	-	No Significant Hits	-
38D10	2.0	-	No Significant Hits	-
112E5	2.00	-	No Significant Hits	-
152H6	2.00	-	No Significant Hits	-
198C10	-2.0	-	No Significant Hits	-
88F4	-2.1	Q8S4F6	*A. thaliana* sulfolipid synthase (hypothetical protein at5g01220).	1E-65
77D7	-2.1	CB12_LYCES	*Lycopersicon esculentum* chlorophyll a-b binding protein 7, chloroplast precursor	1E-119

CD3	-2.1	V00618	Kanamycin (neomycin phosphotransferase)	X
110E3	-2.2	Q40185	*Lemna gibba* light-harvesting chlorophyll a/b protein precursor.	1E-129
161D8	-2.2	Q9SVW8	*A. thaliana* hypothetical protein.	9E-59
196A10	-2.2	Q9XQB6	*P. aureus* chlorophyll a/b-binding protein cp24.	1E-118
C14	-2.2	V00618	Kanamycin (neomycin phosphotransferase)	X
198H12	-2.2	Q9XQB2	*P. aureus* chlorophyll a/b binding protein cp29.	1E-130
198G11	-2.2	-	No Significant Hits	-
86D12	-2.3	METE_CATRO	C 5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase (ec 2.1.1.14)	1E-110
83H3	-2.4	PSBO_LYCES	L oxygen-evolving enhancer protein 1, chloroplast precursor (oee1)	1E-102
198C8	-2.4	AB027001	*N. paniculata* plastidic aldolase1.	1E-137
CE12	-2.6	V00618	Kanamycin (neomycin phosphotransferase)	X
37E12	-3.5	CB24_TOBAC	*N. tabacum* chlorophyll a-b binding protein 40, chloroplast precursor (lhcii type i cab-40) (lhcp).	1E-123

This is the BLAST n value for this cDNA as it was almost an exact match at the nucleotide level to the TYDC entry listed beside it. The + and – represent the direction of the differential expression in the experiment. + represents the clone was differentially underexpressed in the *cor* hairpin RNAi plants compared to the non transformed wildtype plant C052. – represents the clone was differentially overexpressed in the *cor* hairpin RNAi plants compared to the non transformed wildtype plant C052. X-represents sequence identity of an already known gene. NB: the entire set of differential genes for this experiment are listed, using a differential cut-off of 2. None of the possible homologous clones have been removed.

construct. The NPTII gene was differentially expressed between the parental phenotype and the transgenic lines with a greater than two fold increase in expression. While this is a good positive control it is still lower than one would expect possibly due to low expression of the NPTII transgene (kanamycin neophosphotransferase II gene, V00618) or the expression of genes in the poppy of similar nucleotide sequence.

There are many differential clones in this set, which have no significant homology in the database and deserve further attention beyond this thesis. Because the transgenic comparison is virtually isogenic, these differentially expressed genes are candidates for being affected directly or indirectly by the silencing of codeinone reductase and the accumulation of (*S*)-reticuline (see further discussion in Chapter 5).

4.4 Thebaine Accumulating Germplasm

4.4.1 *P. somniferum* Thebaine Accumulating Mutant F075

Microarray experiments were undertaken on pooled RNA from three of the six plants with the thebaine phenotype for the first biological replicate. The remaining three F075 thebaine phenotype plants were then pooled for the second biological replicate. In both cases a dye-swapping technical replicate was included, giving a total of four microarray experiments for the F075 mutant.

Microarray hybridisation results and analyses revealed a set of 46 genes that were differentially expressed between the parental control and the F075 mutant (Figure 4.4) using the expression ratio cut-off of 2. All of the significantly differentially expressed genes were found to be under-expressed in the F075 mutant compared to the parent. Several genes were differentially over-expressed in the mutant F075 compared to the parental wild-type plant however these cDNAs differential expression ratios were substantially lower than the cut-off of 2.

The differentially expressed genes in the F075 were sequenced in the 3' to 5' direction and BLAST x values are listed in Table 4.5.

One cDNA, 20A7, was a shared differential with cv. Marianne (see section 4.5.1). This gene had no significant hits in the database and may warrant further attention.

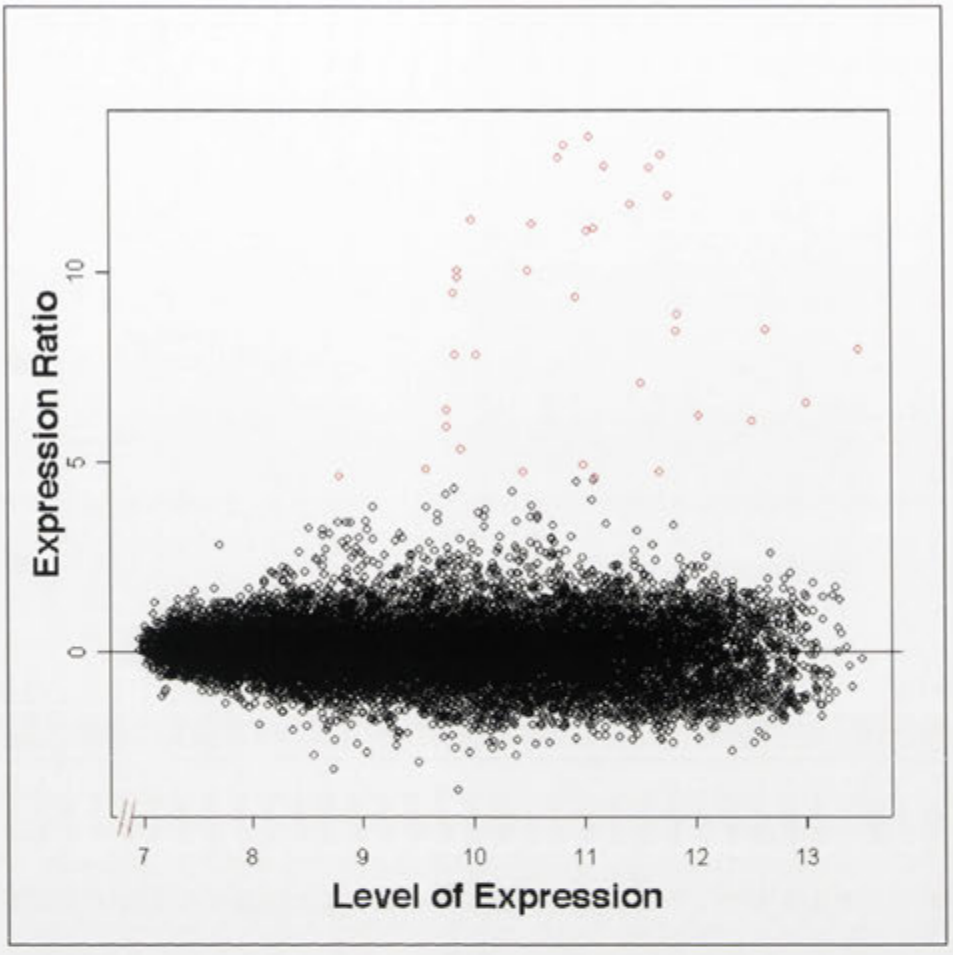


Figure 4.4: Graphical representation of the differentially expressed genes in the F075 mutant. A 2-fold expression differential was used as the cut-off. The red circles show the genes that were differentially down regulated in the F075 mutant.

Interestingly six cDNAs (156F9, 33H5, 71G5, 145E11, 46A6, 33F11) had similar BLAST x values to a tobacco P450-like enzyme in the database (Figure 4.5). This set of genes were not found to be differentially expressed in the *top1* mutant (see section 4.4.2) suggesting the possibility that these cDNAs are specifically associated with the thebaine accumulating phenotype but not the thebaine plus oripavine phenotype. It is thought that the plant enzyme responsible for the demethylation reactions of thebaine are P450s since mammalian liver P450 enzymes have been found which perform the demethylation of thebaine *in vivo* and *in vitro* (Kodaira and Spector 1988).

Table 4.5: Genes that were differentially expressed between the F075 thebaine accumulating mutant and the control C052. All of the genes were more highly expressed in the control than in the mutant. NB the entire set of differential genes is listed here. None of the possible homologous clones have been removed.

Name	Expression Ratio	Accession Number	Description	E-value
27D12	10.2	-	No Significant Hits	-
91G3*	9.0	AY372539	*A. thaliana* signal recognition particle receptor-like protein (at4g30600/f17i23_60).	1E-153
111C7	8.8	AY093119	*A. thaliana* apospory-associated protein c (hypothetical 35.4 kDa protein).	4E-52
91B6	8.7	-	No Significant Hits	-
108H11^	8.3	AY099862	*A. thaliana* potential phospholipid-transporting ATPase 9 (ec 3.6.3.1)	0
156F9†	7.9	AF389267	*Glaucidium palmatum* 26S ribosomal RNA gene, partial sequence	1E-178
151H8	7.7	-	No Significant Hits	-
28G4	7.7	AF133291	*Rhizoctonia solani* virus putative capsid protein mRNA, complete cds	6E-77
33H5†	7.6	AF389262	*Dicentra eximia* 26S ribosomal RNA gene, partial sequence.	1E-130
83G7	7.1	AY372539	*A. thaliana* signal recognition particle receptor-like protein (at4g30600/f17i23_60).	1E-153
71G5†	7.0	AF389267	*Glaucidium palmatum* 26S ribosomal RNA gene, partial sequence.	0
83F5*	6.8	AY372539	*A. thaliana* signal recognition particle receptor-like protein (at4g30600/f17i23_60).	1E-153
157C4	6.8	AB050471	*A. thaliana* phosphoenolpyruvate carboxykinase [atp] (ec 4.1.1.49) (pepck).	1E-164
221C9	6.6	AF419301	*Securigera parviflora*. Aspartate aminotransferase (fragment).	2E-54
52F2^	6.3	AY099862	*A. thaliana* potential phospholipid-transporting ATPase 9 (ec 3.6.3.1)	0
91E9*	6.2	AY372539	*A. thaliana* signal recognition particle receptor-like protein (at4g30600/f17i23_60).	1E-153
54F9	5.8	-	No Significant Hits	-
116D7	5.1	BT002807	*A. thaliana* psi type ii chlorophyll a/b-binding protein, putative.	6E-95
112B7	4.4	-	No Significant Hits	-
83E5	4.3	AF258810	*Lycopersicon esculentum* aldehyde oxidase (AO3) mRNA, complete cds	1E-157
148A9	3.6	-	No Significant Hits	-
145E11†	3.6	AF389270	*Xanthorhiza simplicissima* 26S ribosomal RNA gene, partial sequence.	0
144F10	3.4	-	No Significant Hits	-

153C2	3.1	AF196292	*Apium graveolens* nonreversible glyceraldehyde-3-phosphate dehydrogenase mRNA.	0
186F8	2.9	AF419301	*Securigera parviflora* aspartate aminotransferase (CRAAT1) mRNA,	6E-40
147A12#	2.7	ATR506156	*Amborella trichopoda* chloroplast genomic DNA, 23S ribosomal RNA	0
46A6†	2.7	AF274639	*Cercidiphyllum japonicum* 26S ribosomal RNA gene, complete sequence	0
194D4	2.7	NPU87848	*Nicotiana plumbaginifolia* non-phosphorylating glyceraldehyde dehydrogenase (GapN) mRNA	0
33 F11†	2.5	AF293760	*Sanguinaria canadensis* 18S ribosomal RNA gene, partial sequence	0
34D10	2.4	-	No Significant Hits	-
102D3	2.3	-	No Significant Hits	-
71H4	2.2	-	No Significant Hits	-
83H8	2.2	-	No Significant Hits	-
24C1	2.2	-	No Significant Hits	-
20A7	2.1	-	No Significant Hits	-
88H8	2.1	Q40185	*Lemna gibba* light-harvesting chlorophyll a/b protein precursor.	6E-70
139A8	2.1	-	No Significant Hits	-
104G9	2.1	-	No Significant Hits	-
83H7	2.1	-	No Significant Hits	-
209B1	2.1	Q8L6A3	*Vitis vinifera* putative ribosomal protein l2	7E-80
25D3	2.0	-	No Significant Hits	-
133D7	2.0	ATPA_ TOBAC	*Nicotiana tabacum* atp synthase alpha chain (ec 3.6.3.14).	1E-169
103B4	2.0	Q8H694	*Oryza sativa* osjnba0004i20.3 protein. Contains the phosphoglycerate mutase family phosphohistidine signature	1E-95
89C4	2.0	Q9LDY5	*A. thaliana* t10f20.17 protein (t10o22.14) (putative fkbp-type peptidyl-prolyl cis-trans isomerase protein)	4E-65

* These four clones are homologous and were the same four clones that were differentially expressed in the *top1* mutant microarray analysis.
^ These two clones are homologous and were the same two clones that were differentially expressed in the *top1* mutant microarray analysis.
†The best hits for these clones were taken from the BLAST n results. All of these clones align to various regions of the ribosomal DNA sequence.
#This is homologous to a chloroplast 23S ribosomal gene.

The tobacco P450 clone that the translated sequences of the 6 genes aligned to was isolated using a synthetic oligonucleotide to the heme-binding region of the avocado *CYP71A1* (Sugiura *et al.*, 1996). In this work 2,4-D induced cell

suspensions of tobacco were used to isolate Poly A^+ RNA to clone and express the coding sequence in both *E. coli* and *Saccharomyces cerevisiae.* The authors showed that when the microsomal located tobacco P450 was co-expression in *E. coli* and *S. cerevisiae* with a P450 reductase, the extracts had 7-ethoxycoumarin *O*-deethylase activity (Sugiura *et al.*, 1996). Without the presence of an *N*-terminal microsomal targeting sequence there was no detectable activity. Increased message levels of the P450-like gene could be detected by northern analysis when cell suspensions were induced with 2,4-D. Further to this, evidence was found that the enzyme was a P450 as it produced an absorbance spectrum of 450 nm and the enzyme was inhibited when carbon monoxide was introduced, a classic characteristic of a P450. Upon checking the BLAST n results for these differentially expressed genes I noted that all of these genes were highly homologous to plant ribosomal DNA sequences (Figure 4.6). Several of the genes gave better BLAST n hits to ribosomal DNA sequences than they did BLAST x values. The cDNA sequences in the 5'-3' direction were homologous to the region of DNA coding for the 26S ribosomal DNA. This was perplexing, as the characterisation of the tobacco P450-like gene did not mention its possible alignment to a 26S ribosomal DNA sequence in the complementary direction. Further analysis of the sequence of the tobacco P450-like gene showed that it did indeed align to the 26S ribosomal DNA coding region in the complementary direction. However the entire tobacco P450-like gene did not align, as the 3' end of the sequence was different to the 26S ribosomal DNA coding region (Figure 4.7). The cDNA library was directionally cloned (See Chapter 2) but it is possible the genes were cloned in the wrong direction if they had a similar sequence at the 5'end to the adaptor or an adenosine nucleotide rich region at the 5' end.

Further research of the literature revealed another two genes that also code for functional enzymes in one direction but have very close homology to the 26S ribosomal DNA sequence in the reverse direction. A nuclear-targeted cDNA called ribin was isolated from rat using Southwestern-binding assays and was shown to modulate ribosomal transcription (Kermekchiev and Ivanova, 2001). The complementary strand of the ribin cDNA was homologous to the 26S

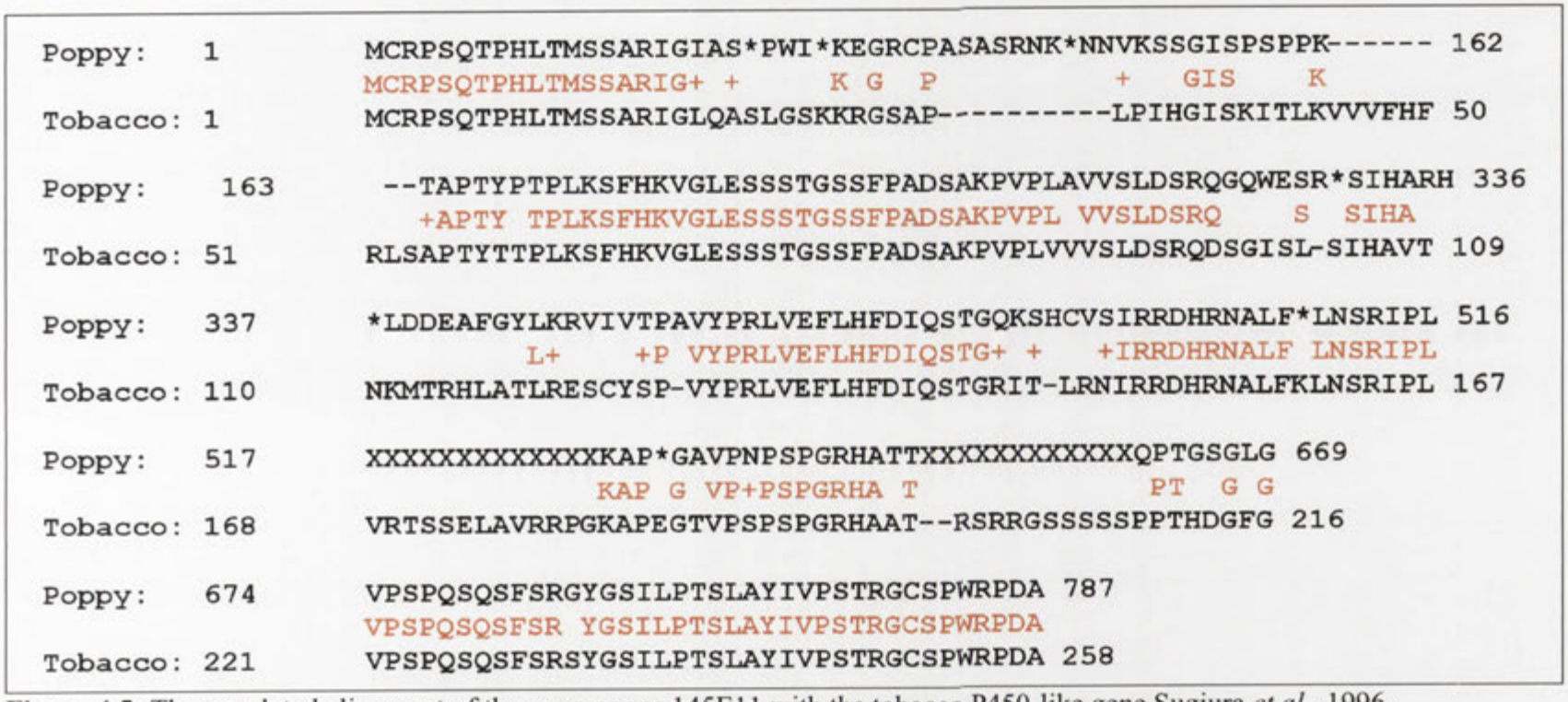


Figure 4.5: The translated alignment of the poppy gene 145E11 with the tobacco P450-like gene Sugiura *et al.*, 1996.

The numbers at the ends of the rows are indicative of the nucleotide number for the poppy sequence and the amino acid sequence number for the tobacco P-450 like gene respectively.

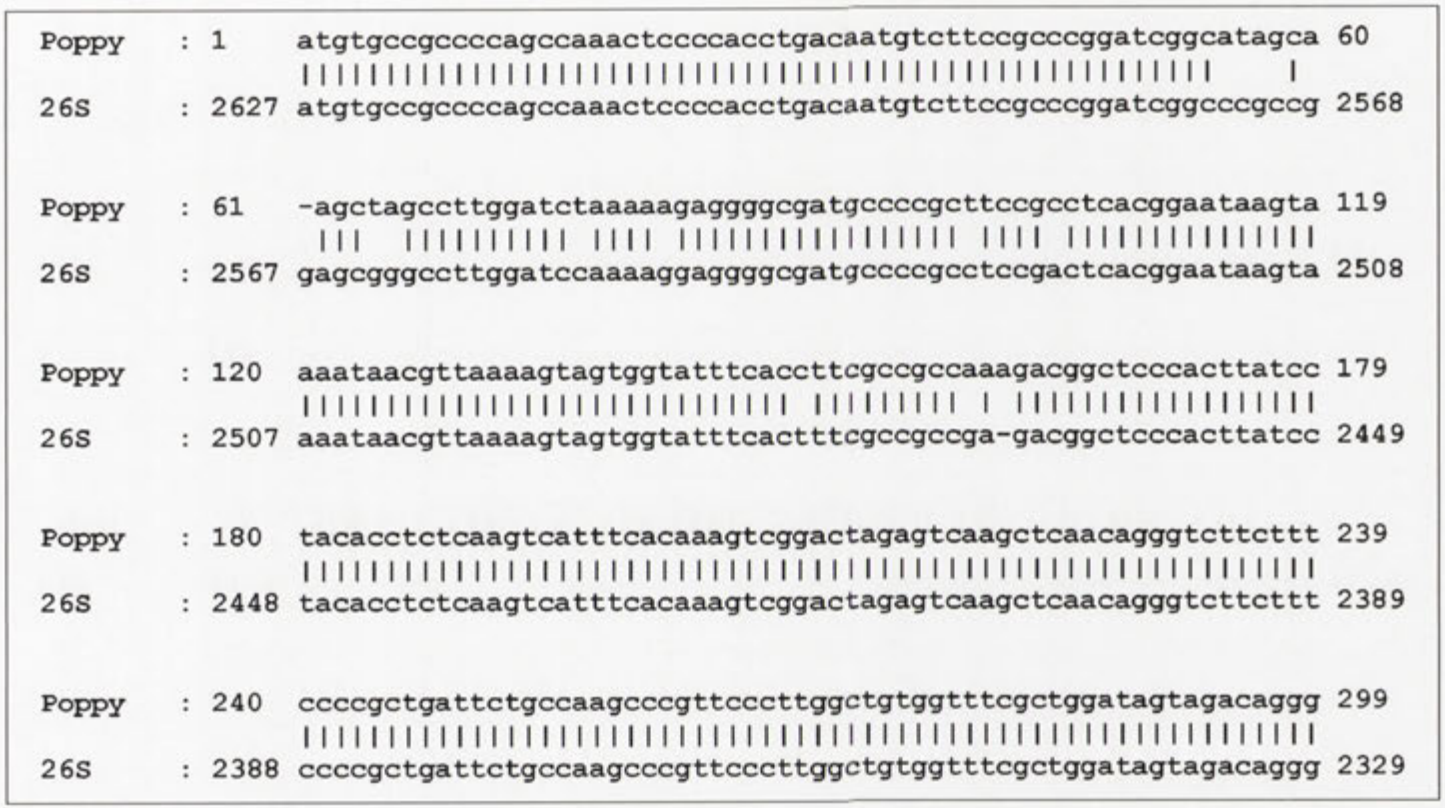


Figure 4.6: The BLAST n alignment of the cDNA 145E11 with a 26S ribosomal gene isolated from *Coptis japonica*. The homology suggests that the cDNA 145E11 codes for a 26S ribosomal gene. However when the BLAST x alignment is undertaken a striking similarity is obtained to the tobacco P450-like gene.

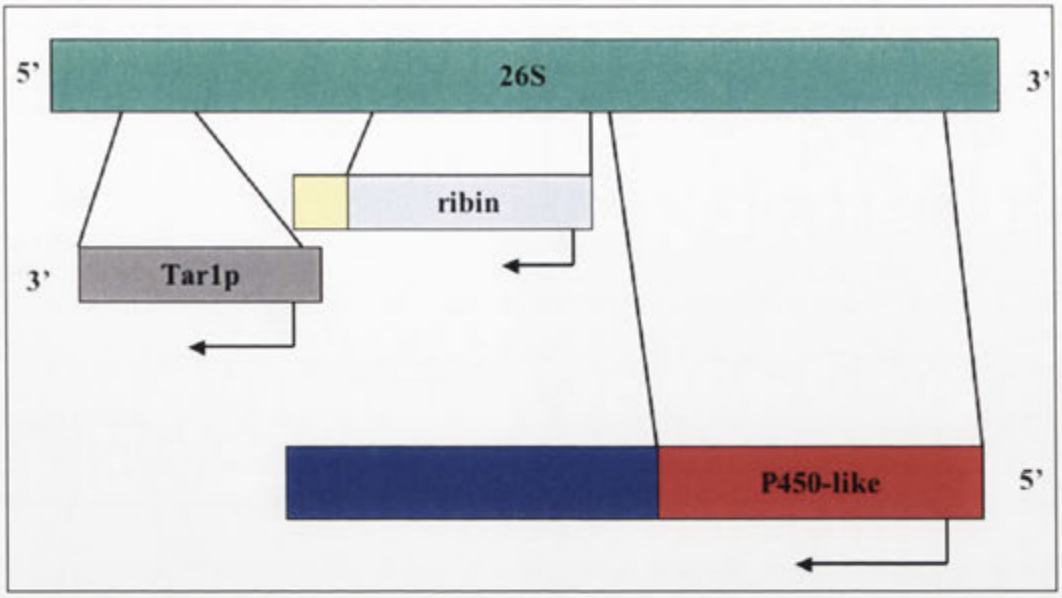


Figure 4.7: Diagram illustrating the homology between the tobacco P450-like gene, 3'-5' direction, and the 26S ribosomal DNA sequence, 5'-3' direction. The red shaded area shows the region of homology at the 5' end and the blue coloured area shows the non-homologous region. The diagram also shows where the other two genes, ribin and Tar1p, lie in comparison to the 26S ribosomal DNA sequence. Both of these genes also transcribe in the reverse direction to the 26S rDNA gene. The ribin gene also has a different 3' coding region to the 26S rDNA.

---

ribosomal DNA sequence. However, the 3'UTR of the ribin protein is different to the rat 26S ribosomal DNA (Figure 4.7).

The other gene encoded in the 26S ribosomal region was Tar1p, Transcript Antisense to rDNA. This cDNA was isolated from yeast cells and its overexpression rescued a mitochondrial RNA polymerase mutant (Coelho *et al.*, 2002). This gene was different to the ribin and tobacco P450-like gene (Figure 4.7).

This has led to several questions. Firstly is it possible to have two functional genes overlapping to this extent? Secondly, is the P450 a real P450 or is it a 26S ribosomal DNA? Thirdly, is it possible to have functional transcripts complementary to the 26S rDNA sequence? The evidence for the tobacco P450-

like gene being a P450 appears to be sound, since functionality has been demonstrated when the cDNA was co-expression in yeast with a P450-reductase.

One problem with this gene is that, if it is a membrane-bound P450, then why didn't Sugiura et al isolate the membrane-spanning coding region as well? The authors had to substitute a membrane anchoring coding region to the cDNA for its co-expression with the yeast P450-reductase system. Also, there was no mention of complementarity to the 26S rDNA and there have been no further follow-up publications. This suggests that either the gene analysed was not a P450 and follow up work was unrewarding, or the complementarity of the gene to 26S ribosomal DNA sequence has made it difficult to clone the genomic version. Since there are many copies of the ribosomal genes it may be that only one has acquired the enzyme functionality in the complementary strand; the strong residual homology to the other ribosomal genes would make it very difficult to clone the enzyme-encoding version. The 3' region of the tobacco P450-like gene has a very different coding region and if this was present in the 26S rDNA sequence it would more than likely lose ribosomal function.

It is now well established that the chloroplast genome has genes nested within genes and that the complementary strand can encode different genes (Rasmussen *et al.*, 1984). A recent study in yeast has found that 2% of the yeast genome can contain new gene transcripts that are transcribed from genes nested within larger open reading frames on either the same or complementarystrands (Kumar *et al.*, 2002). The occurrence of exons within introns is also now known with several 'accidental' discoveries of these genes. The *Tau* gene of humans contains a regulatory sequence within its intron that *Tau* exon-10 expression is dependent upon (D'Souza and Schellenberg, 2002). A conserved intronless gene, C18orf2, has been found nested within the GNAL gene on chromosome 18p11I in humans (Vuoristo *et al.*, 2001).

In plants the nuclear encoded 26S rDNA sequence is part of a larger DNA unit that encodes the 18S, 5.8S and 26S rDNAs in that order (Figure 4.8). The larger nuclear ribosomal DNA unit is repeated many thousands of times in higher plant genomes. This sequence is highly conserved across species. The 26S rDNA

coding region is the least variable and the 5.8 and 18S rDNAs are more variable. Separating each of the rDNA sequences is a more variable region (variable in comparison to the rDNA sequences but still highly conserved) called the internal transcribed spacer (ITS). Many studies have used this conservation as a means of establishing relatedness between closely related taxa (Baldwin 1992). With so many copies of this rDNA unit in the plant genome it is possible that mutation has occurred in one or a few copies resulting in new or altered functions. It is the rRNA that affects translation of protein, not a translated protein of this RNA. The peptidyl transferase reaction still takes place in *E. coli* when the 50S subunits were treated with proteinase K and SDS, which cleaves and denature proteins respectively (Noller *et al.*, 1992).

It is likely that there are genes transcribed on the complementary strand of particular member of the rDNA cluster such as the *Tar1p* gene, however the poppy cDNAs isolated from the microarray study that have homology to the 26S rDNA sequence in the complementary strand, all contain a stop codon approximately 22 codons in from the start methionine when aligned with the tobacco P450-like gene. It is possible that these cDNAs on the array are true ribosomal RNAs

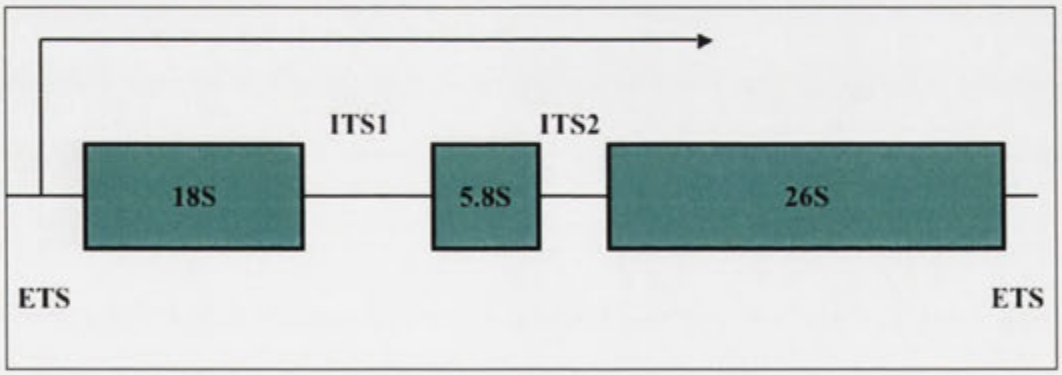


Figure 4.8: Illustration of the nuclear ribosomal DNA sequence. The ETS represents the external transcribed spacer that has the highest sequence variation in the ribosomal DNA sequence. The internal transcribed spacer regions (ITS1 and ITS2) are the next most highly variable sequence in the ribosomal DNA sequence. The arrow shows the direction of transcription of the ribosomal RNA. The order of the rDNA in plants is 18S, 5.8S and 26S. The S is a measure of the ribosomal RNAs size.

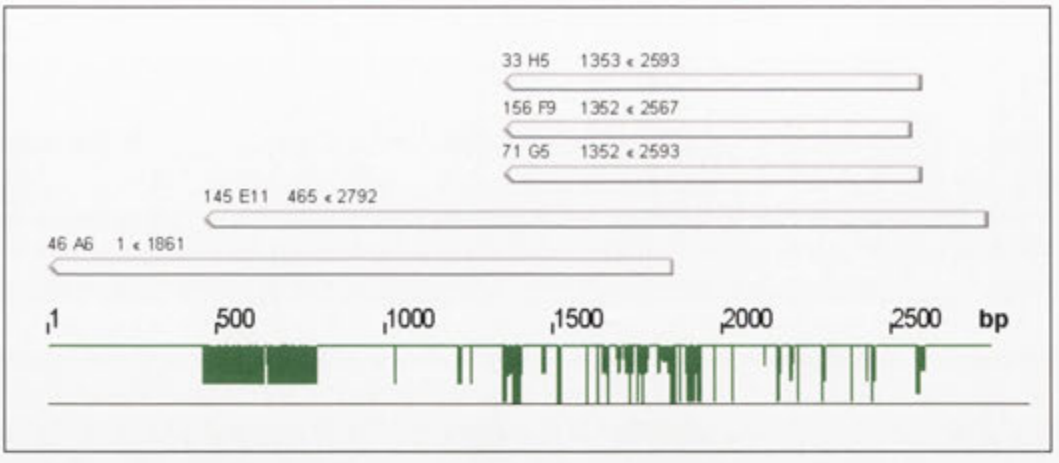


Figure 4.9: Alignment of the 5 separate 26S rDNA cDNAs that were differentially expressed in the F075 mutant microarray experiments. All of the cDNAs except for 145E11 had a Poly A^+ tail at the 5' end of the 26S rDNA sequence. The green line indicates the homology between the sequences. The narrower the green line the better the homology between the different 26S coding DNA sequences. For example at the 500 bp region the differences can be attributed to the Poly A^+ tail that has been left on the genes. Appendix 6 contains the actual nucleotide alignments.

---

that have been reverse transcribed within the cDNA library. The isolation of the Poly A^+ mRNA for the cDNA library was undertaken using the Promega Poly A^+ mRNA beads purification system© from total RNA isolated from *P. somniferum* (Chapter 2.2). Ribosomal RNA does not have a Poly A^+ tail and the *P. somniferum* cDNA library was reverse transcribed using an oligo dT primer on Poly A^+ mRNA (Chapter 2.2). The *P. somniferum* 26S rDNA coding cDNAs that have been sequenced from the thebaine-accumulating microarray experiment all contained, except for 145E11, a Poly A^+ tail. Further to this the total RNA isolated for labelling of the probes for the microarray experiments was treated with RNase H to degrade any RNA after the labelling process, which again used an oligo dT primer to the Poly A^+ tail. This is suggestive that either there was aberrant ribosomal RNA labelling with Poly A^+, which has been reverse transcribed, cloned and then printed upon the array or that there are genes that are transcribed and possibly functional with homology to the 26S ribosomal DNA. The microarray experiments of the thebaine-accumulating mutant are the only experiments in which the five 26S rDNA cDNAs are differentially expressed.

The potential significance of this result is underscored by all four cDNAs being consistently under-expressed in the thebaine-accumulating mutant.

The five cDNAs are not to identical regions of the 26S rDNA sequence, however they do share one region of homology to the 26S rDNA sequence (Figure 4.9). It is possible that there is a gene transcribed from the reverse strand of one copy of poppy 26S rDNA and it is contained within this region of homology or that there is an alternate gene in the poppy that is expressed that has high homology to this region. There still remains the question as to why this expressed gene is consistently under-expressed in the thebaine-accumulating mutant.

Transcription of the gene could be authenticated on a Northern blot using single stranded RNA probes for the complementary strand of the 26S rDNA-coding region. The use of Poly A^+ mRNA on the Northern could be used to enhance the results. However the presence of the cDNA with a Poly A tail and the fact that RNA is labelled by the microarray labelling process and presents a differential expression between the mutant and the parent cultivar indicates that there is transcription of this gene taking place.

Further work needs to be undertaken to isolate this gene and to test its function in the plant. One way to do this would be to try transforming *P. somniferum* with part of the sequence in an inverted repeat RNAi silencing construct. However it may not be possible to obtain transformed plants as the inverted repeat construct would also knock down transcription of the 26S ribosome. This would more than likely be deleterious to the plant as the inverted repeat technology is not strand specific and the 26S ribosome is essential as part of the transcription machinery.

Another way to find the gene encoding this message is to probe the cDNA library with the region of homology between the cDNA clones; the set of hybridising clones would be sequenced and screened to find one, which can be translated to a functional protein. Using a differential library between the thebaine-accumulating mutant and the *P. somniferum* parent lines could enhance success. The problem with these experiments is the large number of 26S ribosomal coding regions present in the genome.

Mutant F075 has considerable potential as a tool for deciphering the complexities of the morphinan alkaloid pathway. However the major shortcoming of F075 is that pure breeding lines cannot be obtained. Because the frequency of plants with the thebaine phenotype are rare in seed populations, it is difficult to acquire enough material for molecular analysis. It is possible that it is a chromosome rearrangement that is unstable rather than a simple point mutation causing the phenotype.

4.4.2 *P.somniferum* Thebaine and Oripavine Accumulating Mutant, *top1*

The differentially expressed genes for the *top1* microarray experiments are presented in Table 4.5. This set of experiments utilised a total of 12 microarray slides. This included 6 biological replicates, with each biological replicate consisting of RNA extracted from pooled material from 6 plants. Each of the biological replicates consisted of one dye swap each to give the total of twelve microarray slides. The large number of biological replicates yields a much greater statistical significance upon the resulting clones that are found to be differential between the mutant *top1* and the control *P. somniferum*. All of the 14 clones were differentially under expressed in the mutant compared to the control and no genes were found to be statistically significantly upregulated in the mutant (Figure 4.10).

Removal of the redundant clones from the set of 14 cDNAs that were differentially expressed in the *top1* experiments left a subset of 10 different cDNAs (Table 4.6).

Six of the ten clones showed very high sequence homology to known genes in the database of different species (Table 4.6). The most notable of these were four clones (83G7, 91E9, 91G3, and 91C9) that showed very high homology to the *Arabidopsis thaliana* signal recognition particle receptor-like protein (SRPR) (atg30600/f17i23_60) (Figure 4.11). These four clones are not full length to the coding region of the SRPRS in the database and are missing the 3' and 5' UTR

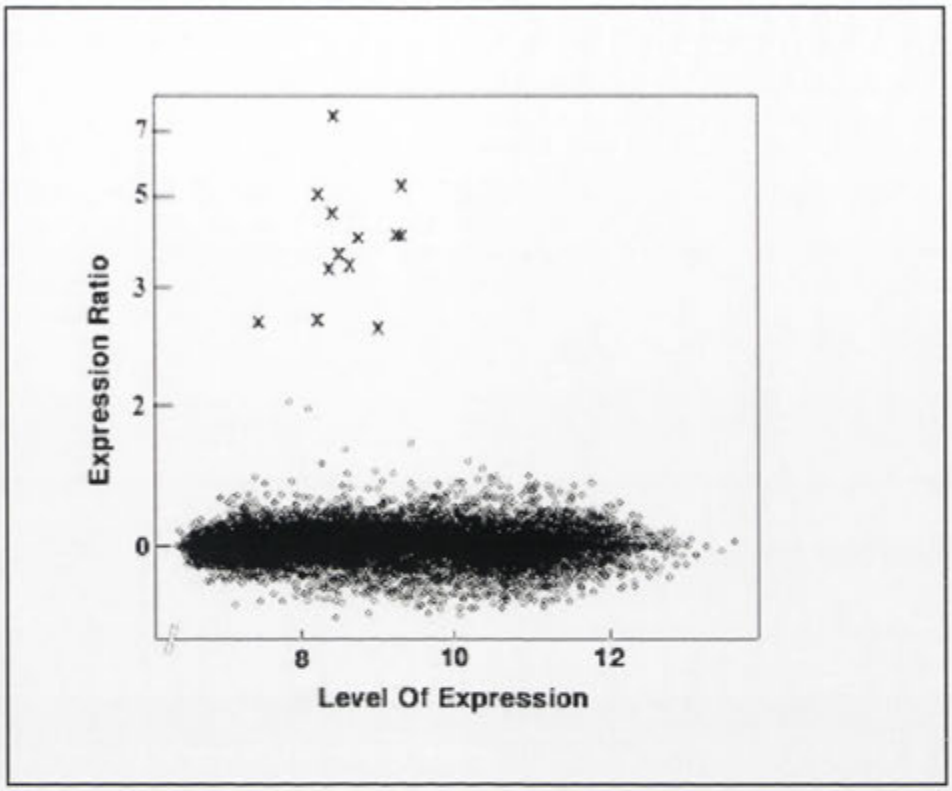


Figure 4.10: Graphical representation of the analysis of the 12 microarray slides from *top1* compared with the C052 control. The 14 differentially expressed genes are each represented by an x. A positive expression ratio shows the genes were differentially downregulated in *top1* compared to the control plant (Millgate *et al.*, 2004).

regions. The nucleotide alignment for these four clones (Figure 4.12) were the same indicating they are cDNAs from the same gene, or members of the same multigene family. The srpr protein is part of an oligomeric complex known as the signal recognition particle that mediates targeting and insertion of the signal sequence of exported proteins into the membrane of the endoplasmic reticulum.

Two clones 52F2 and 108H11 had homology to the *A. thaliana* potential phospholipid-transporting ATPase 9 (EC 3.6.3.1) (Figure 4.13) otherwise known as an amino phospholipid flippase 9. These ATP-dependent translocators (flippases) control transmembrane lipid asymmetry. The nucleotide alignment of these two clones (Figure 4.14), however showed that they are almost identical, except for clone 108H11, which contains an extra 130 nucleotides 323 bp from the start of the sequenced cDNAs. This may be the result of an unexcised intron

Table 4.6: Possible function of the 14 cDNA clones found to be differentially regulated between C052 and *top1* (Millgate *et al.*, 2004).

Clone	cDNA Size	Number of copies	Array Expression Ratio	Description Best Hit blast X	E-value
116D7	1098	1	7.9	*A. thaliana* psi type ii chlorophyll a/b-binding protein, putative.	6E-95
91G3	2152	4	5.1	*A. thaliana* signal recognition particle receptor-like protein (at4g30600/f17i23_60).	1E-153
27D12	498	1	4.7	No significant hits.	-
157C4	2402	1	4.4	*A. thaliana* phosphoenolpyruvate carboxykinase [atp] (ec 4.1.1.49) (pepck).	1E-164
91B6	527	1	4.4	No significant hits.	-
54F9	144	1	3.9	No significant hits.	-
52F2	2015	2	3.6	*A. thaliana* Potential phospholipid-transporting ATPase 9 (ec 3.6.3.1)	0
221C9	995	1	3.4	*Securigera parviflora*. Aspartate aminotransferase (fragment).	2E-54
151H8	785	1	2.9	No significant hits.	-
111C7	913	1	2.7	*A. thaliana* Apospory-associated protein c (hypothetical 35.4 kDa protein).	4E-52

Previous Table 4.6: The complete set of non-redundant differentially expressed genes in the *top1* mutant including identified genes and unidentified genes, the latter based on E values greater than E^{-50}. The table lists: clone identification (Clone Id); number of redundant genes (Number of Copies); accession number (Accession); gene name from BLAST x results (Description) and E values are shown. **a**. The E-value is the score calculated by NCBI BLAST x searches. **b**. The Microarray Ratio is the normalised ratio value (see Chapter 2.6) compared to the control *P. somniferum*. A positive value reveals higher expression in the control (reference sample) than in *top1* (test sample).

or there is more than one copy of the gene present. Several sets of primers were tried to differentiate between the two clones, 52F2 and 108H11, however multiple bands were obtained when primers to either gene were used. It is possible there is a gene family of varying sizes or that post-transcriptional modification, for example intron removal, may be taking place giving different length transcripts. As it was not possible to design quantitative real time PCR primers to distinguish between the two, this clone was not authenticated using qRT-PCR. Clone 153C2 was similar to the NADP-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.9). Other clones with homology to genes of known function were: 157C4 (*A. thaliana* phosphoenolpyruvate carboxykinase (pep carboxykinase EC 3.4.24)); and 221C9 (*Securigera parviflora* aspartate aminotransferase). Several other clones had sequence homology, however these were to sequences of unknown function. QRT-PCR analysis was used to further authenticate the non-redundant set of 10 genes/alleles.

It is possible that by using an array prepared from RNA of specialised cells, such as from the latex, the array may have been enhanced for alkaloid genes. For example many photosynthetic genes would have been eliminated. However it was judged necessary to use the whole above ground plant for the experiments because the evidence suggests cells neighbouring the laticifers are important sites for the transcripts and biosynthesis (Bird et al., 2003 and Weid et al., 2004). Many of the genes involved in alkaloid synthesis are possibly expressed in sieve element companion cells or phloem parenchyma. Further the use of the parent

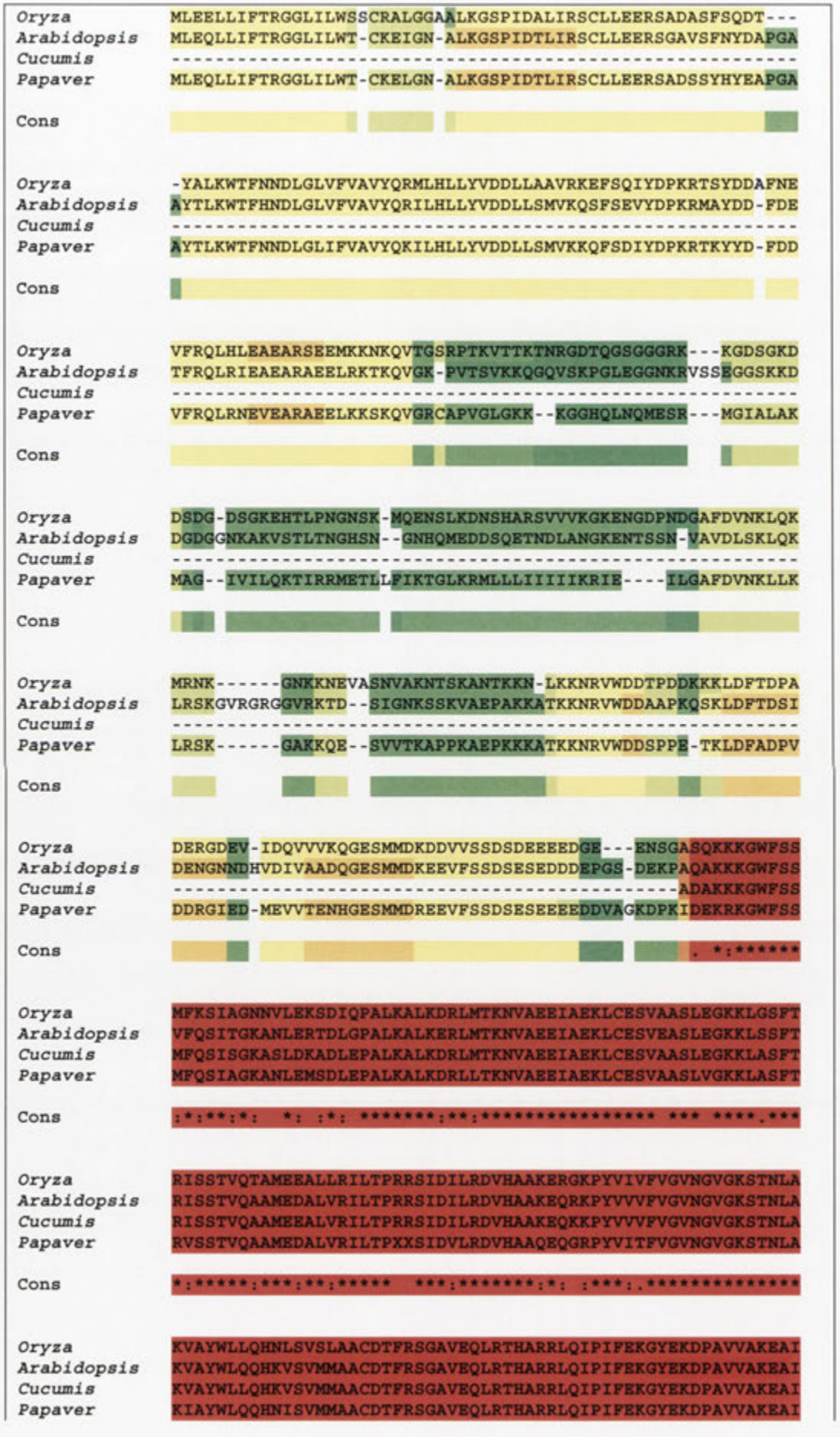
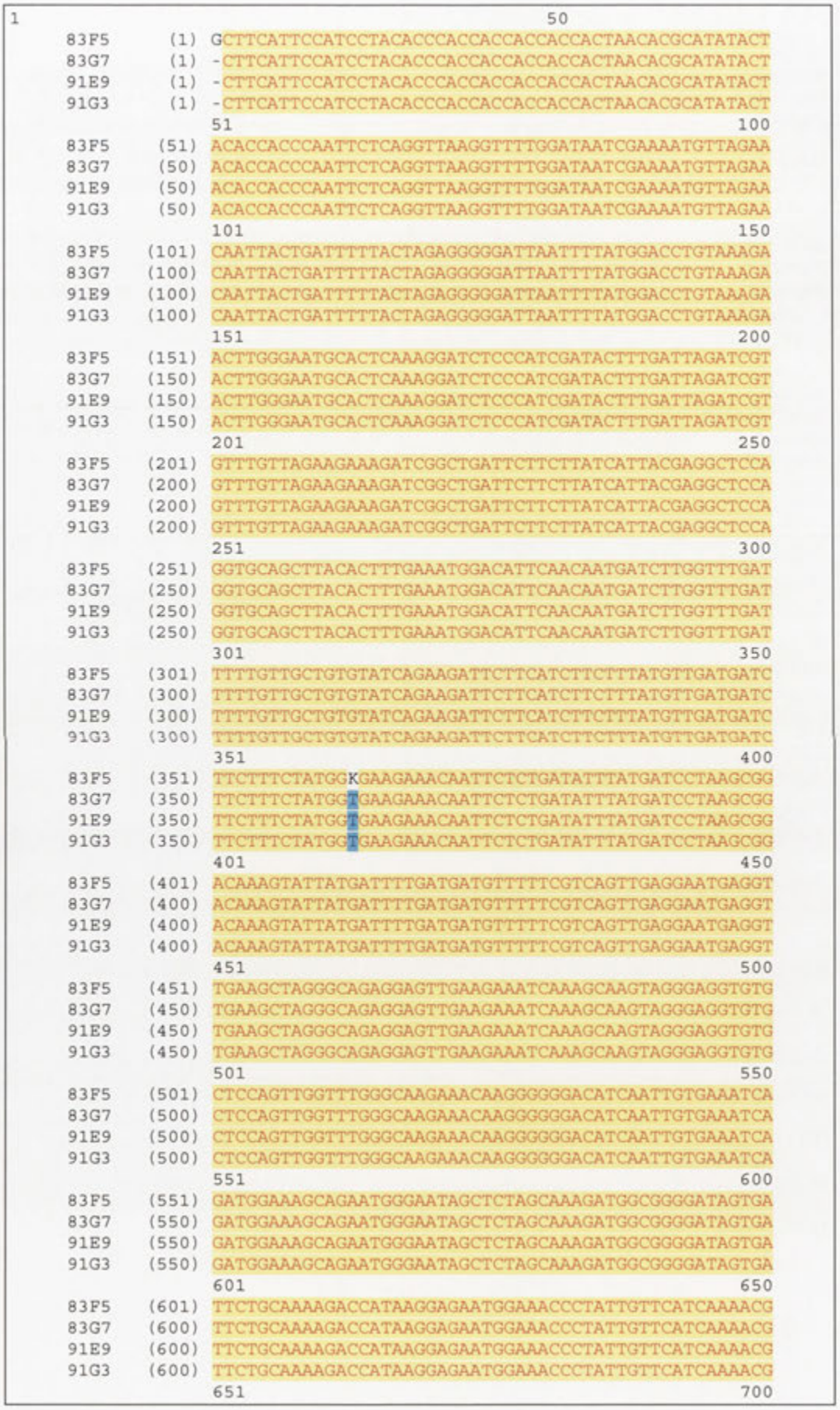
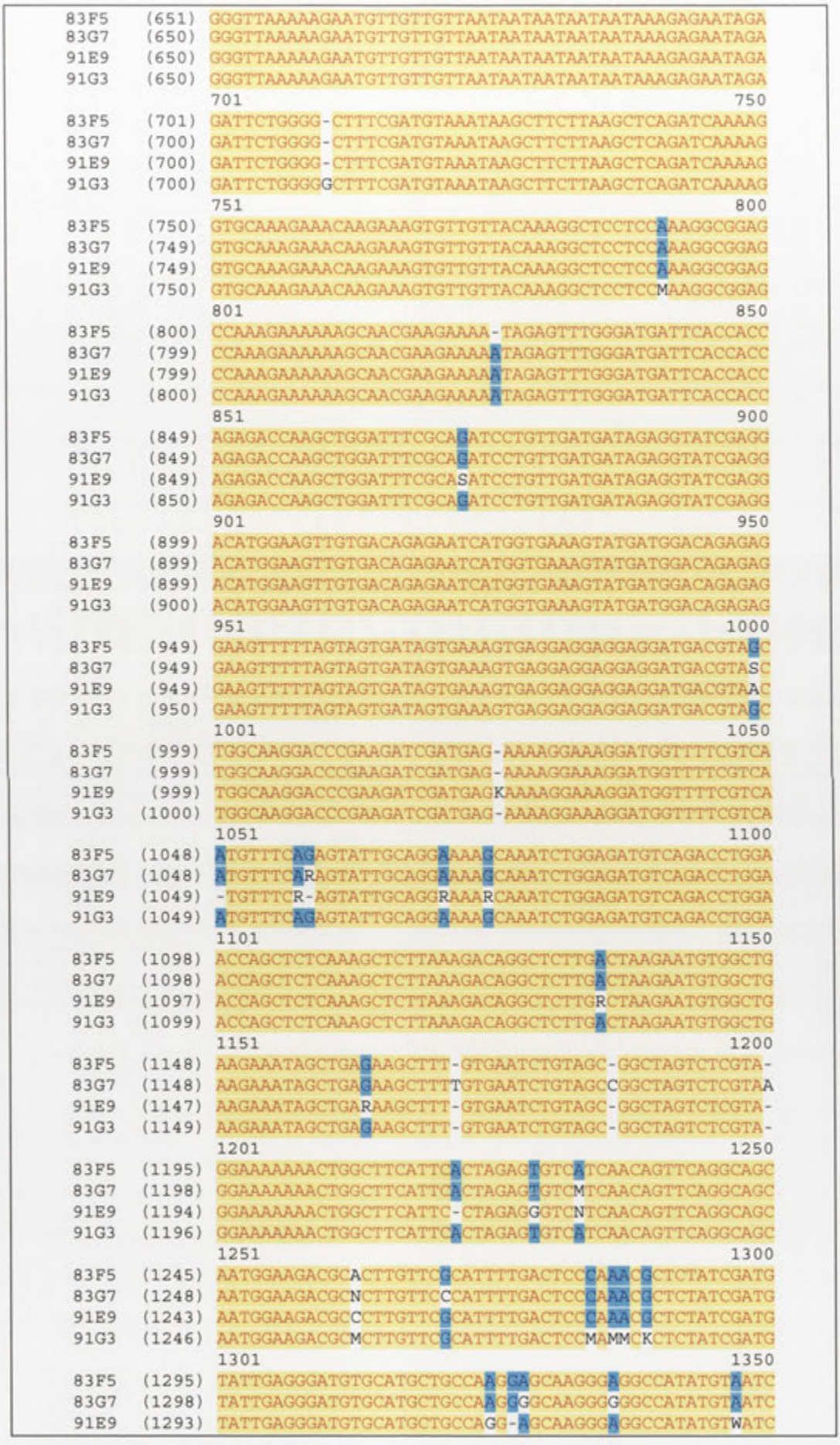
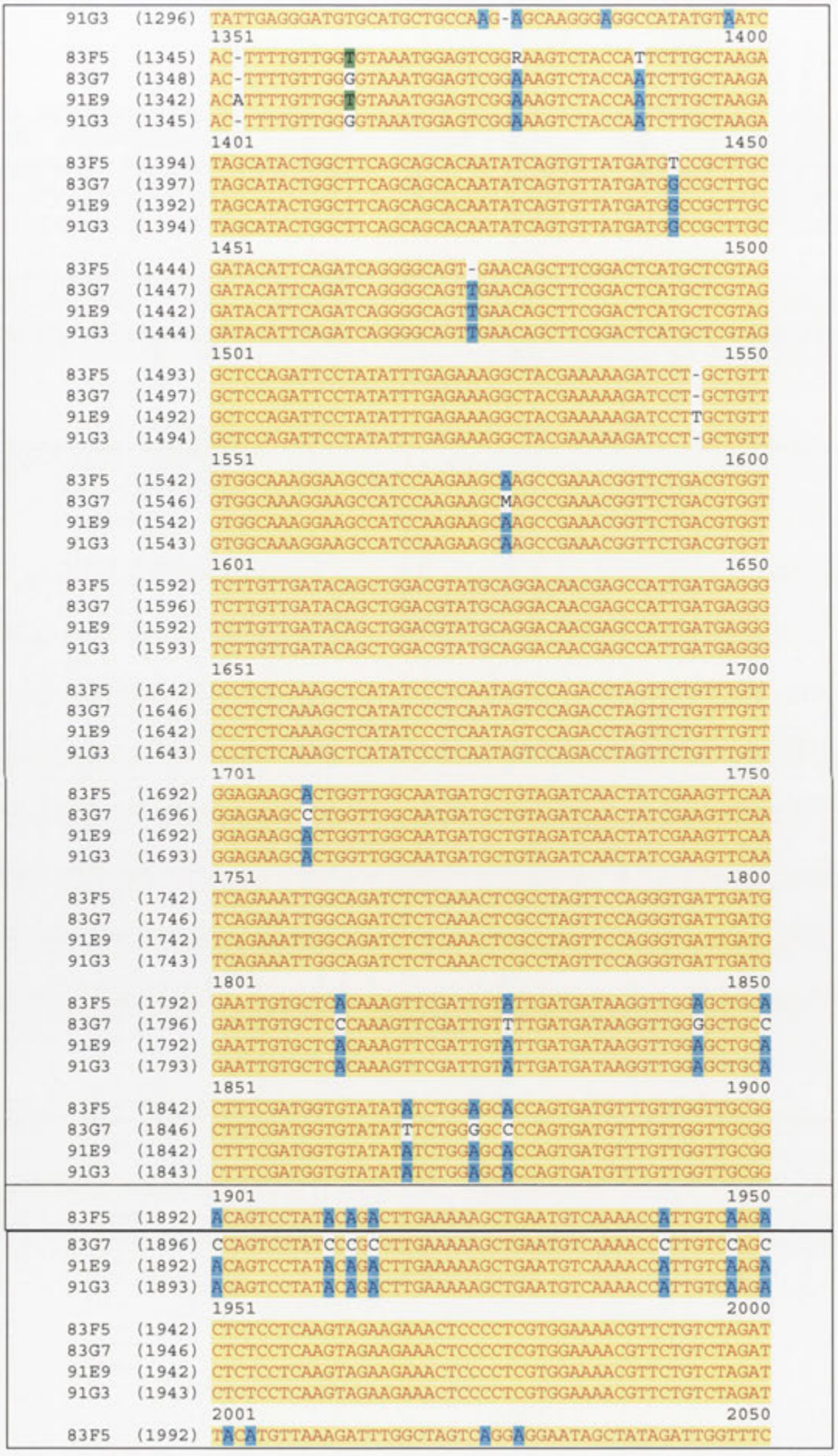

Oryza MLEELLIFTRGGLILWSSCRALGGAALKGSPIDALIRSCLLEERSADASFSQDT---
Arabidopsis MLEQLLIFTRGGLILWT-CKEIGN-ALKGSPIDTLIRSCLLEERSGAVSFNYDAPGA
Cucumis ----------
Papaver MLEQLLIFTRGGLILWT-CKELGN-ALKGSPIDTLIRSCLLEERSADSSYHYEAPGA
Cons
Oryza -YALKWTFNNDLGLVFVAVYQRMLHLLYVDDLLAAVRKEFSQIYDPKRTSYDDAFNE
Arabidopsis AYTLKWTFHNDLGLVFVAVYQRILHLLYVDDLLSMVKQSFSEVYDPKRMAYDD-FDE
Cucumis ----------
Papaver AYTLKWTFNNDLGLIFVAVYQKILHLLYVDDLLSMVKKQFSDIYDPKRTKYYD-FDD
Cons
Oryza VFRQLHLEAEARSEEMKKNKQVTGSRPTKVTTKTNRGDTQGSGGGRK---KGDSGKD
Arabidopsis TFRQLRIEAEARAEELRKTKQVGK-PVTSVKKQGQVSKPGLEGGNKRVSSEGGSKKD
Cucumis ----------
Papaver VFRQLRNEVEARAEELKKSKQVGRCAPVGLGKK--KGGHQLNQMESR---MGIALAK
Cons
Oryza DSDG-DSGKEHTLPNGNSK-MQENSLKDNSHARSVVVKGKENGDPNDGAFDVNKLQK
Arabidopsis DGDGGNKAKVSTLTNGHSN--GNHQMEDDSQETNDLANGKENTSSN-VAVDLSKLQK
Cucumis ----------
Papaver MAG--IVILQKTIRRMETLLFIKTGLKRMLLLIIIIIKRIE----ILGAFDVNKLLK
Cons
Oryza MRNK------GNKKNEVASNVAKNTSKANTKKN-LKKNRVWDDTPDDKKKLDFTDPA
Arabidopsis LRSKGVRGRGGVRKTD--SIGNKSSKVAEPAKKATKKNRVWDDAAPKQSKLDFTDSI
Cucumis ----------
Papaver LRSK------GAKKQE--SVVTKAPPKAEPKKKATKKNRVWDDSPPE-TKLDFADPV
Cons
Oryza DERGDEV-IDQVVVKQGESMMDKDDVVSSDSDEEEEDGE---ENSGASQKKKGWFSS
Arabidopsis DENGNNDHVDIVAADQGESMMDKEEVFSSDSESEDDDEPGS-DEKPAQAKKKGWFSS
Cucumis ----------ADAKKKGWFSS
Papaver DDRGIED-MEVVTENHGESMMDREEVFSSDSESEEEEDDVAGKDPKIDEKRKGWFSS
Cons
Oryza MFKSIAGNNVLEKSDIQPALKALKDRLMTKNVAEEIAEKLCESVAASLEGKKLGSFT
Arabidopsis VFQSITGKANLERTDLGPALKALKERLMTKNVAEEIAEKLCESVEASLEGKKLSSFT
Cucumis MFQSISGKASLDKADLEPALKALKDRLMTKNVAEEIAEKLCESVAASLEGKKLASFT
Papaver MFQSIAGKANLEMSDLEPALKALKDRLLTKNVAEEIAEKLCESVAASLVGKKLASFT
Cons
Oryza RISSTVQTAMEEALLRILTPRRSIDILRDVHAAKERGKPYVIVFVGVNGVGKSTNLA
Arabidopsis RISSTVQAAMEDALVRILTPRRSIDILRDVHAAKEQRKPYVVVFVGVNGVGKSTNLA
Cucumis RISSTVQAAMEEALVRILTPRRSIDILRDVHAAKEQKKPYVVVFVGVNGVGKSTNLA
Papaver RVSSTVQAAMEDALVRILTPXXSIDVLRDVHAAQEQGRPYVITFVGVNGVGKSTNLA
Cons
Oryza KVAYWLLQHNLSVSLAACDTFRSGAVEQLRTHARRLQIPIFEKGYEKDPAVVAKEAI
Arabidopsis KVAYWLQQHKVSVMMAACDTFRSGAVEQLRTHARRLQIPIFEKGYEKDPAVVAKEAI
Cucumis KVAYWLLQHKVSVMMAACDTFRSGAVEQLRTHARRLQIPIFEKGYEKDPAVVAKEAI
Papaver KIAYWLQQHNISVMMAACDTFRSGAVEQLRTHARRLQIPIFEKGYEKDPAVVAKEAI

Figure 4.11: The translated alignment of the poppy clone 83E5 with three signal recognition particle receptor proteins from the NCBI genbank database.
Sequence homology is represented by the colour code- BAD AVG GOOD
Oryza: putative signal recognition particle receptor [*Oryza sativa* (japonica cultivar-group)]. Accession number: BAD09759.
Arabidopsis: signal recognition particle receptor alpha subunit family protein [*Arabidopsis thaliana*]. Accession number: NP_194789.
Cucumis: signal recognition particle receptor protein [*Cucumis sativus*]. Accession number: AAQ76042.
Papaver: Clone 83E5 from the *top1* microarray results.
T-COFFEE, Version_1.41 was used for the amino acid alignments (Notredame *et al.*, 2000).

plant C052 to generate the cDNA for printing the microarrays may impose a bias in certain circumstances, e.g. if the gene of interest have been upregulated in the mutant phenotype and is normally not expressed in the control C052, the gene will not be found using the array. Several cloned morphine associated genes; TYDC, cyp 80b1, berberine bridge enzyme and codeinone reductase, did not show any differential expression between *top1* and its parent across the 12 slides (data not shown).


1 50
83F5 (1) GCTTCATTCCATCCTACACCCACCACCACCACCACTAACACGCATATACT
83G7 (1) -CTTCATTCCATCCTACACCCACCACCACCACCACTAACACGCATATACT
91E9 (1) -CTTCATTCCATCCTACACCCACCACCACCACCACTAACACGCATATACT
91G3 (1) -CTTCATTCCATCCTACACCCACCACCACCACCACTAACACGCATATACT
51 100
83F5 (51) ACACCACCCAATTCTCAGGTTAAGGTTTTGGATAATCGAAAATGTTAGAA
83G7 (50) ACACCACCCAATTCTCAGGTTAAGGTTTTGGATAATCGAAAATGTTAGAA
91E9 (50) ACACCACCCAATTCTCAGGTTAAGGTTTTGGATAATCGAAAATGTTAGAA
91G3 (50) ACACCACCCAATTCTCAGGTTAAGGTTTTGGATAATCGAAAATGTTAGAA
101 150
83F5 (101) CAATTACTGATTTTTACTAGAGGGGGATTAATTTTATGGACCTGTAAAGA
83G7 (100) CAATTACTGATTTTTACTAGAGGGGGATTAATTTTATGGACCTGTAAAGA
91E9 (100) CAATTACTGATTTTTACTAGAGGGGGATTAATTTTATGGACCTGTAAAGA
91G3 (100) CAATTACTGATTTTTACTAGAGGGGGATTAATTTTATGGACCTGTAAAGA
151 200
83F5 (151) ACTTGGGAATGCACTCAAAGGATCTCCCATCGATACTTTGATTAGATCGT
83G7 (150) ACTTGGGAATGCACTCAAAGGATCTCCCATCGATACTTTGATTAGATCGT
91E9 (150) ACTTGGGAATGCACTCAAAGGATCTCCCATCGATACTTTGATTAGATCGT
91G3 (150) ACTTGGGAATGCACTCAAAGGATCTCCCATCGATACTTTGATTAGATCGT
201 250
83F5 (201) GTTTGTTAGAAGAAAGATCGGCTGATTCTTCTTATCATTACGAGGCTCCA
83G7 (200) GTTTGTTAGAAGAAAGATCGGCTGATTCTTCTTATCATTACGAGGCTCCA
91E9 (200) GTTTGTTAGAAGAAAGATCGGCTGATTCTTCTTATCATTACGAGGCTCCA
91G3 (200) GTTTGTTAGAAGAAAGATCGGCTGATTCTTCTTATCATTACGAGGCTCCA
251 300
83F5 (251) GGTGCAGCTTACACTTTGAAATGGACATTCAACAATGATCTTGGTTTGAT
83G7 (250) GGTGCAGCTTACACTTTGAAATGGACATTCAACAATGATCTTGGTTTGAT
91E9 (250) GGTGCAGCTTACACTTTGAAATGGACATTCAACAATGATCTTGGTTTGAT
91G3 (250) GGTGCAGCTTACACTTTGAAATGGACATTCAACAATGATCTTGGTTTGAT
301 350
83F5 (301) TTTTGTTGCTGTGTATCAGAAGATTCTTCATCTTCTTTATGTTGATGATC
83G7 (300) TTTTGTTGCTGTGTATCAGAAGATTCTTCATCTTCTTTATGTTGATGATC
91E9 (300) TTTTGTTGCTGTGTATCAGAAGATTCTTCATCTTCTTTATGTTGATGATC
91G3 (300) TTTTGTTGCTGTGTATCAGAAGATTCTTCATCTTCTTTATGTTGATGATC
351 400
83F5 (351) TTCTTTCTATGGKGAAGAAACAATTCTCTGATATTTATGATCCTAAGCGG
83G7 (350) TTCTTTCTATGGTGAAGAAACAATTCTCTGATATTTATGATCCTAAGCGG
91E9 (350) TTCTTTCTATGGTGAAGAAACAATTCTCTGATATTTATGATCCTAAGCGG
91G3 (350) TTCTTTCTATGGTGAAGAAACAATTCTCTGATATTTATGATCCTAAGCGG
401 450
83F5 (401) ACAAAGTATTATGATTTTGATGATGTTTTTCGTCAGTTGAGGAATGAGGT
83G7 (400) ACAAAGTATTATGATTTTGATGATGTTTTTCGTCAGTTGAGGAATGAGGT
91E9 (400) ACAAAGTATTATGATTTTGATGATGTTTTTCGTCAGTTGAGGAATGAGGT
91G3 (400) ACAAAGTATTATGATTTTGATGATGTTTTTCGTCAGTTGAGGAATGAGGT
451 500
83F5 (451) TGAAGCTAGGGCAGAGGAGTTGAAGAAATCAAAGCAAGTAGGGAGGTGTG
83G7 (450) TGAAGCTAGGGCAGAGGAGTTGAAGAAATCAAAGCAAGTAGGGAGGTGTG
91E9 (450) TGAAGCTAGGGCAGAGGAGTTGAAGAAATCAAAGCAAGTAGGGAGGTGTG
91G3 (450) TGAAGCTAGGGCAGAGGAGTTGAAGAAATCAAAGCAAGTAGGGAGGTGTG
501 550
83F5 (501) CTCCAGTTGGTTTGGGCAAGAAACAAGGGGGGACATCAATTGTGAAATCA
83G7 (500) CTCCAGTTGGTTTGGGCAAGAAACAAGGGGGGACATCAATTGTGAAATCA
91E9 (500) CTCCAGTTGGTTTGGGCAAGAAACAAGGGGGGACATCAATTGTGAAATCA
91G3 (500) CTCCAGTTGGTTTGGGCAAGAAACAAGGGGGGACATCAATTGTGAAATCA
551 600
83F5 (551) GATGGAAAGCAGAATGGGAATAGCTCTAGCAAAGATGGCGGGGATAGTGA
83G7 (550) GATGGAAAGCAGAATGGGAATAGCTCTAGCAAAGATGGCGGGGATAGTGA
91E9 (550) GATGGAAAGCAGAATGGGAATAGCTCTAGCAAAGATGGCGGGGATAGTGA
91G3 (550) GATGGAAAGCAGAATGGGAATAGCTCTAGCAAAGATGGCGGGGATAGTGA
601 650
83F5 (601) TTCTGCAAAAGACCATAAGGAGAATGGAAACCCTATTGTTCATCAAAACG
83G7 (600) TTCTGCAAAAGACCATAAGGAGAATGGAAACCCTATTGTTCATCAAAACG
91E9 (600) TTCTGCAAAAGACCATAAGGAGAATGGAAACCCTATTGTTCATCAAAACG
91G3 (600) TTCTGCAAAAGACCATAAGGAGAATGGAAACCCTATTGTTCATCAAAACG
651 700


83F5 (651) GGGTTAAAAAGAATGTTGTTGTTAATAATAATAATAATAAAGAGAATAGA
83G7 (650) GGGTTAAAAAGAATGTTGTTGTTAATAATAATAATAATAAAGAGAATAGA
91E9 (650) GGGTTAAAAAGAATGTTGTTGTTAATAATAATAATAATAAAGAGAATAGA
91G3 (650) GGGTTAAAAAGAATGTTGTTGTTAATAATAATAATAATAAAGAGAATAGA
701 750
83F5 (701) GATTCTGGGG-CTTTCGATGTAAATAAGCTTCTTAAGCTCAGATCAAAAG
83G7 (700) GATTCTGGGG-CTTTCGATGTAAATAAGCTTCTTAAGCTCAGATCAAAAG
91E9 (700) GATTCTGGGG-CTTTCGATGTAAATAAGCTTCTTAAGCTCAGATCAAAAG
91G3 (700) GATTCTGGGGGCTTTCGATGTAAATAAGCTTCTTAAGCTCAGATCAAAAG
751 800
83F5 (750) GTGCAAAGAAACAAGAAAGTGTTGTTACAAAGGCTCCTCCAAAGGCGGAG
83G7 (749) GTGCAAAGAAACAAGAAAGTGTTGTTACAAAGGCTCCTCCAAAGGCGGAG
91E9 (749) GTGCAAAGAAACAAGAAAGTGTTGTTACAAAGGCTCCTCCAAAGGCGGAG
91G3 (750) GTGCAAAGAAACAAGAAAGTGTTGTTACAAAGGCTCCTCCMAAGGCGGAG
801 850
83F5 (800) CCAAAGAAAAAAGCAACGAAGAAAA-TAGAGTTTGGGATGATTCACCACC
83G7 (799) CCAAAGAAAAAAGCAACGAAGAAAAATAGAGTTTGGGATGATTCACCACC
91E9 (799) CCAAAGAAAAAAGCAACGAAGAAAAATAGAGTTTGGGATGATTCACCACC
91G3 (800) CCAAAGAAAAAAGCAACGAAGAAAAATAGAGTTTGGGATGATTCACCACC
851 900
83F5 (849) AGAGACCAAGCTGGATTTCGCAGATCCTGTTGATGATAGAGGTATCGAGG
83G7 (849) AGAGACCAAGCTGGATTTCGCAGATCCTGTTGATGATAGAGGTATCGAGG
91E9 (849) AGAGACCAAGCTGGATTTCGCASATCCTGTTGATGATAGAGGTATCGAGG
91G3 (850) AGAGACCAAGCTGGATTTCGCAGATCCTGTTGATGATAGAGGTATCGAGG
901 950
83F5 (899) ACATGGAAGTTGTGACAGAGAATCATGGTGAAAGTATGATGGACAGAGAG
83G7 (899) ACATGGAAGTTGTGACAGAGAATCATGGTGAAAGTATGATGGACAGAGAG
91E9 (899) ACATGGAAGTTGTGACAGAGAATCATGGTGAAAGTATGATGGACAGAGAG
91G3 (900) ACATGGAAGTTGTGACAGAGAATCATGGTGAAAGTATGATGGACAGAGAG
951 1000
83F5 (949) GAAGTTTTTAGTAGTGATAGTGAAAGTGAGGAGGAGGAGGATGACGTAGC
83G7 (949) GAAGTTTTTAGTAGTGATAGTGAAAGTGAGGAGGAGGAGGATGACGTASC
91E9 (949) GAAGTTTTTAGTAGTGATAGTGAAAGTGAGGAGGAGGAGGATGACGTAAC
91G3 (950) GAAGTTTTTAGTAGTGATAGTGAAAGTGAGGAGGAGGAGGATGACGTAGC
1001 1050
83F5 (999) TGGCAAGGACCCGAAGATCGATGAG-AAAAGGAAAGGATGGTTTTCGTCA
83G7 (999) TGGCAAGGACCCGAAGATCGATGAG-AAAAGGAAAGGATGGTTTTCGTCA
91E9 (999) TGGCAAGGACCCGAAGATCGATGAGKAAAAGGAAAGGATGGTTTTCGTCA
91G3 (1000) TGGCAAGGACCCGAAGATCGATGAG-AAAAGGAAAGGATGGTTTTCGTCA
1051 1100
83F5 (1048) ATGTTTCAGAGTATTGCAGGAAAAGCAAATCTGGAGATGTCAGACCTGGA
83G7 (1048) ATGTTTCARAGTATTGCAGGAAAAGCAAATCTGGAGATGTCAGACCTGGA
91E9 (1049) -TGTTTCR-AGTATTGCAGGRAAARCAAATCTGGAGATGTCAGACCTGGA
91G3 (1049) ATGTTTCAGAGTATTGCAGGAAAAGCAAATCTGGAGATGTCAGACCTGGA
1101 1150
83F5 (1098) ACCAGCTCTCAAAGCTCTTAAAGACAGGCTCTTGACTAAGAATGTGGCTG
83G7 (1098) ACCAGCTCTCAAAGCTCTTAAAGACAGGCTCTTGACTAAGAATGTGGCTG
91E9 (1097) ACCAGCTCTCAAAGCTCTTAAAGACAGGCTCTTGRCTAAGAATGTGGCTG
91G3 (1099) ACCAGCTCTCAAAGCTCTTAAAGACAGGCTCTTGACTAAGAATGTGGCTG
1151 1200
83F5 (1148) AAGAAATAGCTGAGAAGCTTT-GTGAATCTGTAGC-GGCTAGTCTCGTA-
83G7 (1148) AAGAAATAGCTGAGAAGCTTTTGTGAATCTGTAGCCGGCTAGTCTCGTAA
91E9 (1147) AAGAAATAGCTGARAAGCTTT-GTGAATCTGTAGC-GGCTAGTCTCGTA-
91G3 (1149) AAGAAATAGCTGAGAAGCTTT-GTGAATCTGTAGC-GGCTAGTCTCGTA-
1201 1250
83F5 (1195) GGAAAAAAACTGGCTTCATTCACTAGAGTGTCATCAACAGTTCAGGCAGC
83G7 (1198) GGAAAAAAACTGGCTTCATTCACTAGAGTGTCMTCAACAGTTCAGGCAGC
91E9 (1194) GGAAAAAAACTGGCTTCATTC-CTAGAGGGTCNTCAACAGTTCAGGCAGC
91G3 (1196) GGAAAAAAACTGGCTTCATTCACTAGAGTGTCATCAACAGTTCAGGCAGC
1251 1300
83F5 (1245) AATGGAAGACGCACTTGTTCGCATTTTGACTCCCAAACGCTCTATCGATG
83G7 (1248) AATGGAAGACGCNCTTGTTCCCATTTTGACTCCCAAACGCTCTATCGATG
91E9 (1243) AATGGAAGACGCCCTTGTTCGCATTTTGACTCCCAAACGCTCTATCGATG
91G3 (1246) AATGGAAGACGCMCTTGTTCGCATTTTGACTCCMAMMCKCTCTATCGATG
1301 1350
83F5 (1295) TATTGAGGGATGTGCATGCTGCCAAGGAGCAAGGGAGGCCATATGTAATC
83G7 (1298) TATTGAGGGATGTGCATGCTGCCAAGGGGCAAGGGGGGCCATATGTAATC
91E9 (1293) TATTGAGGGATGTGCATGCTGCCAGG-AGCAAGGGAGGCCATATGTWATC


91G3 (1296) TATTGAGGGATGTGCATGCTGCCAAG-AGCAAGGGAGGCCATATGTAATC
1351 1400
83F5 (1345) AC-TTTTGTTGGTGTAAATGGAGTCGGRAAGTCTACCATTCTTGCTAAGA
83G7 (1348) AC-TTTTGTTGGGGTAAATGGAGTCGGAAAGTCTACCAATCTTGCTAAGA
91E9 (1342) ACATTTTGTTGGTGTAAATGGAGTCGGAAAGTCTACCAATCTTGCTAAGA
91G3 (1345) AC-TTTTGTTGGGGTAAATGGAGTCGGAAAGTCTACCAATCTTGCTAAGA
1401 1450
83F5 (1394) TAGCATACTGGCTTCAGCAGCACAATATCAGTGTTATGATGTCCGCTTGC
83G7 (1397) TAGCATACTGGCTTCAGCAGCACAATATCAGTGTTATGATGGCCGCTTGC
91E9 (1392) TAGCATACTGGCTTCAGCAGCACAATATCAGTGTTATGATGGCCGCTTGC
91G3 (1394) TAGCATACTGGCTTCAGCAGCACAATATCAGTGTTATGATGGCCGCTTGC
1451 1500
83F5 (1444) GATACATTCAGATCAGGGGCAGT-GAACAGCTTCGGACTCATGCTCGTAG
83G7 (1447) GATACATTCAGATCAGGGGCAGTTGAACAGCTTCGGACTCATGCTCGTAG
91E9 (1442) GATACATTCAGATCAGGGGCAGTTGAACAGCTTCGGACTCATGCTCGTAG
91G3 (1444) GATACATTCAGATCAGGGGCAGTTGAACAGCTTCGGACTCATGCTCGTAG
1501 1550
83F5 (1493) GCTCCAGATTCCTATATTTGAGAAAGGCTACGAAAAAGATCCT-GCTGTT
83G7 (1497) GCTCCAGATTCCTATATTTGAGAAAGGCTACGAAAAAGATCCT-GCTGTT
91E9 (1492) GCTCCAGATTCCTATATTTGAGAAAGGCTACGAAAAAGATCCTTGCTGTT
91G3 (1494) GCTCCAGATTCCTATATTTGAGAAAGGCTACGAAAAAGATCCT-GCTGTT
1551 1600
83F5 (1542) GTGGCAAAGGAAGCCATCCAAGAAGCAAGCCGAAACGGTTCTGACGTGGT
83G7 (1546) GTGGCAAAGGAAGCCATCCAAGAAGCMAGCCGAAACGGTTCTGACGTGGT
91E9 (1542) GTGGCAAAGGAAGCCATCCAAGAAGCAAGCCGAAACGGTTCTGACGTGGT
91G3 (1543) GTGGCAAAGGAAGCCATCCAAGAAGCAAGCCGAAACGGTTCTGACGTGGT
1601 1650
83F5 (1592) TCTTGTTGATACAGCTGGACGTATGCAGGACAACGAGCCATTGATGAGGG
83G7 (1596) TCTTGTTGATACAGCTGGACGTATGCAGGACAACGAGCCATTGATGAGGG
91E9 (1592) TCTTGTTGATACAGCTGGACGTATGCAGGACAACGAGCCATTGATGAGGG
91G3 (1593) TCTTGTTGATACAGCTGGACGTATGCAGGACAACGAGCCATTGATGAGGG
1651 1700
83F5 (1642) CCCTCTCAAAGCTCATATCCCTCAATAGTCCAGACCTAGTTCTGTTTGTT
83G7 (1646) CCCTCTCAAAGCTCATATCCCTCAATAGTCCAGACCTAGTTCTGTTTGTT
91E9 (1642) CCCTCTCAAAGCTCATATCCCTCAATAGTCCAGACCTAGTTCTGTTTGTT
91G3 (1643) CCCTCTCAAAGCTCATATCCCTCAATAGTCCAGACCTAGTTCTGTTTGTT
1701 1750
83F5 (1692) GGAGAAGCACTGGTTGGCAATGATGCTGTAGATCAACTATCGAAGTTCAA
83G7 (1696) GGAGAAGCCCTGGTTGGCAATGATGCTGTAGATCAACTATCGAAGTTCAA
91E9 (1692) GGAGAAGCACTGGTTGGCAATGATGCTGTAGATCAACTATCGAAGTTCAA
91G3 (1693) GGAGAAGCACTGGTTGGCAATGATGCTGTAGATCAACTATCGAAGTTCAA
1751 1800
83F5 (1742) TCAGAAATTGGCAGATCTCTCAAACTCGCCTAGTTCCAGGGTGATTGATG
83G7 (1746) TCAGAAATTGGCAGATCTCTCAAACTCGCCTAGTTCCAGGGTGATTGATG
91E9 (1742) TCAGAAATTGGCAGATCTCTCAAACTCGCCTAGTTCCAGGGTGATTGATG
91G3 (1743) TCAGAAATTGGCAGATCTCTCAAACTCGCCTAGTTCCAGGGTGATTGATG
1801 1850
83F5 (1792) GAATTGTGCTCACAAAGTTCGATTGTATTGATGATAAGGTTGGAGCTGCA
83G7 (1796) GAATTGTGCTCCCAAAGTTCGATTGTTTTGATGATAAGGTTGGGGCTGCC
91E9 (1792) GAATTGTGCTCACAAAGTTCGATTGTATTGATGATAAGGTTGGAGCTGCA
91G3 (1793) GAATTGTGCTCACAAAGTTCGATTGTATTGATGATAAGGTTGGAGCTGCA
1851 1900
83F5 (1842) CTTTCGATGGTGTATATATCTGGAGCACCAGTGATGTTTGTTGGTTGCGG
83G7 (1846) CTTTCGATGGTGTATATTTCTGGGGCCCCAGTGATGTTTGTTGGTTGCGG
91E9 (1842) CTTTCGATGGTGTATATATCTGGAGCACCAGTGATGTTTGTTGGTTGCGG
91G3 (1843) CTTTCGATGGTGTATATATCTGGAGCACCAGTGATGTTTGTTGGTTGCGG
1901 1950
83F5 (1892) ACAGTCCTATACAGACTTGAAAAAGCTGAATGTCAAAACCATTGTCAAGA
83G7 (1896) CCAGTCCTATCCCGCCTTGAAAAAGCTGAATGTCAAAACCCTTGTCCAGC
91E9 (1892) ACAGTCCTATACAGACTTGAAAAAGCTGAATGTCAAAACCATTGTCAAGA
91G3 (1893) ACAGTCCTATACAGACTTGAAAAAGCTGAATGTCAAAACCATTGTCAAGA
1951 2000
83F5 (1942) CTCTCCTCAAGTAGAAGAAACTCCCCTCGTGGAAAACGTTCTGTCTAGAT
83G7 (1946) CTCTCCTCAAGTAGAAGAAACTCCCCTCGTGGAAAACGTTCTGTCTAGAT
91E9 (1942) CTCTCCTCAAGTAGAAGAAACTCCCCTCGTGGAAAACGTTCTGTCTAGAT
91G3 (1943) CTCTCCTCAAGTAGAAGAAACTCCCCTCGTGGAAAACGTTCTGTCTAGAT
2001 2050
83F5 (1992) TACATGTTAAAGATTTGGCTAGTCAGGAGGAATAGCTATAGATTGGTTTC

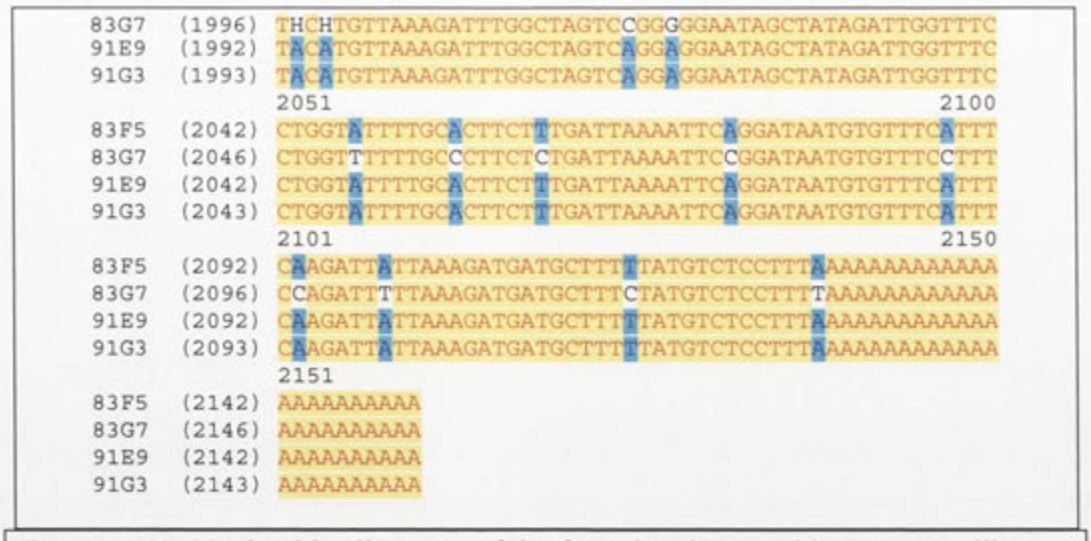
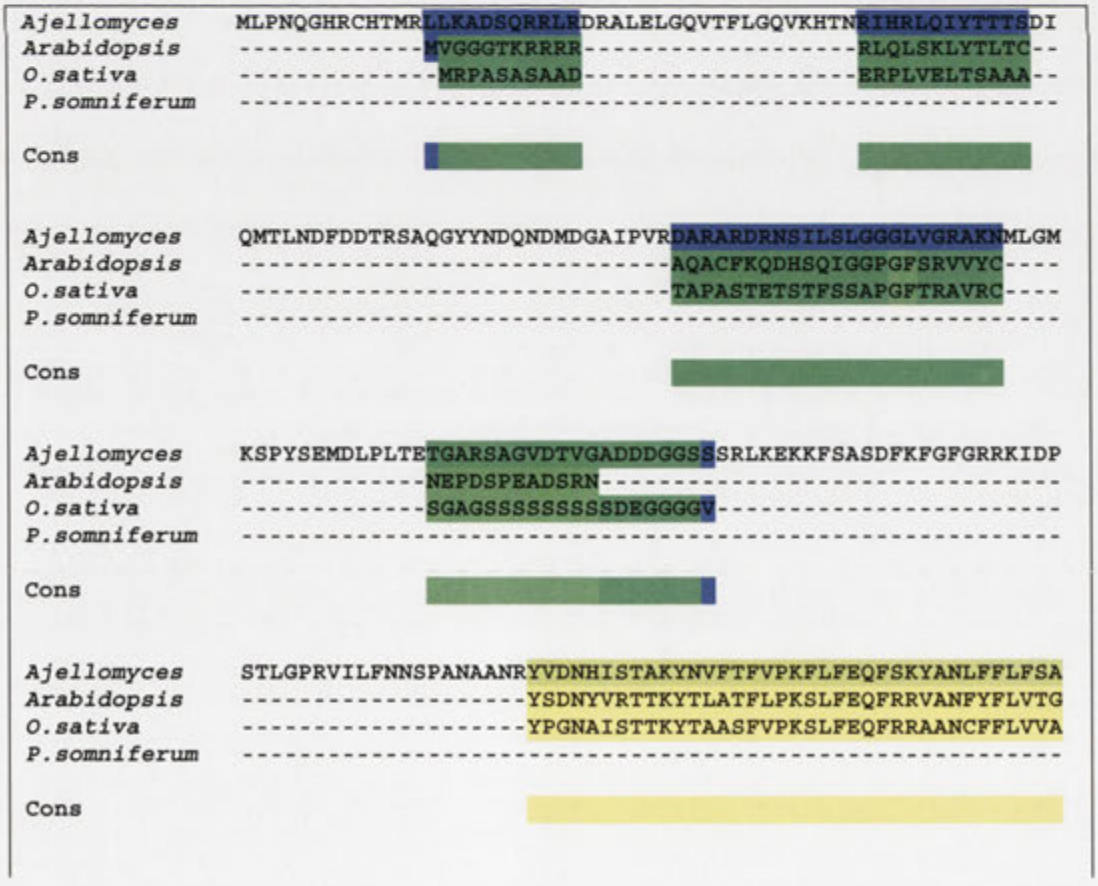
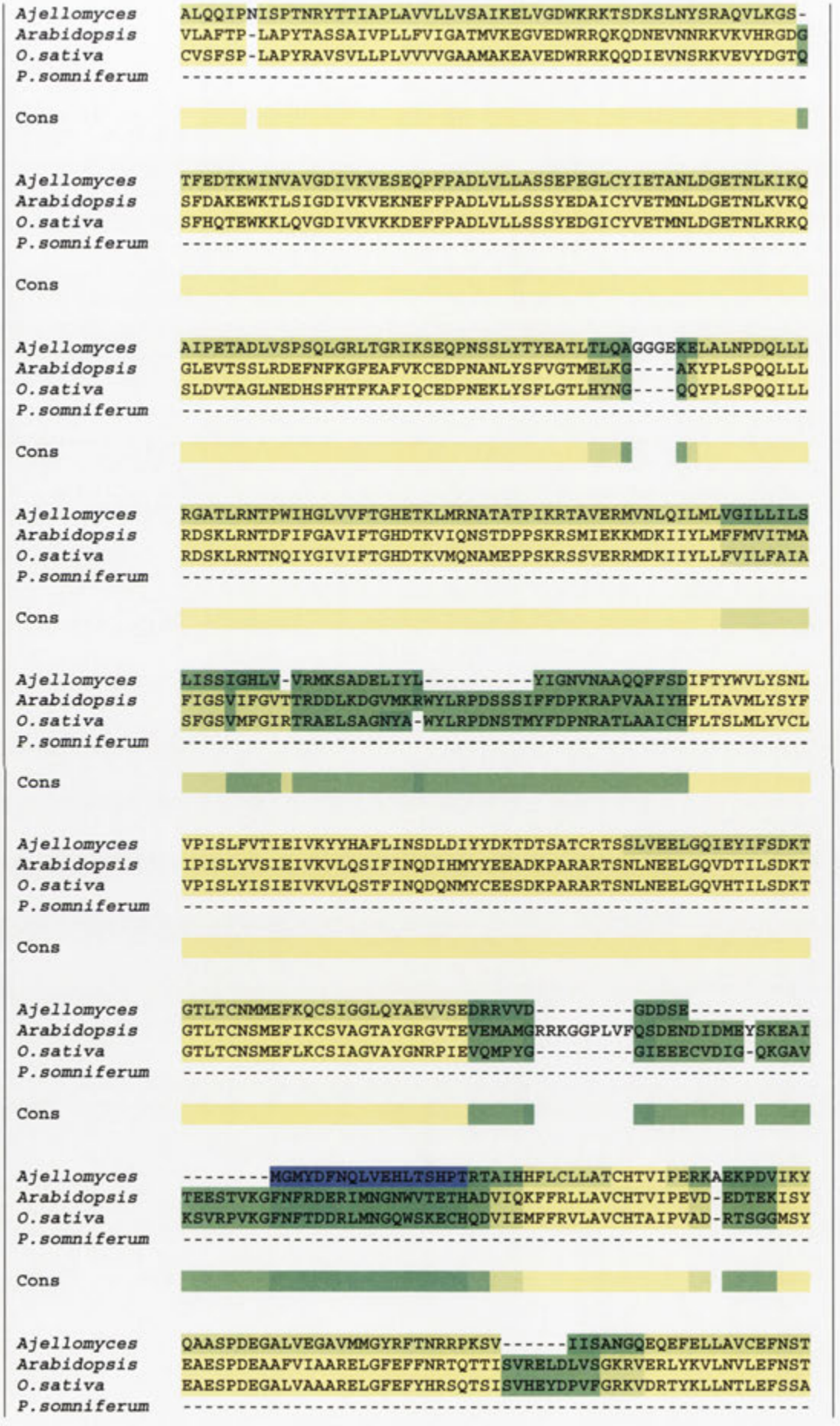
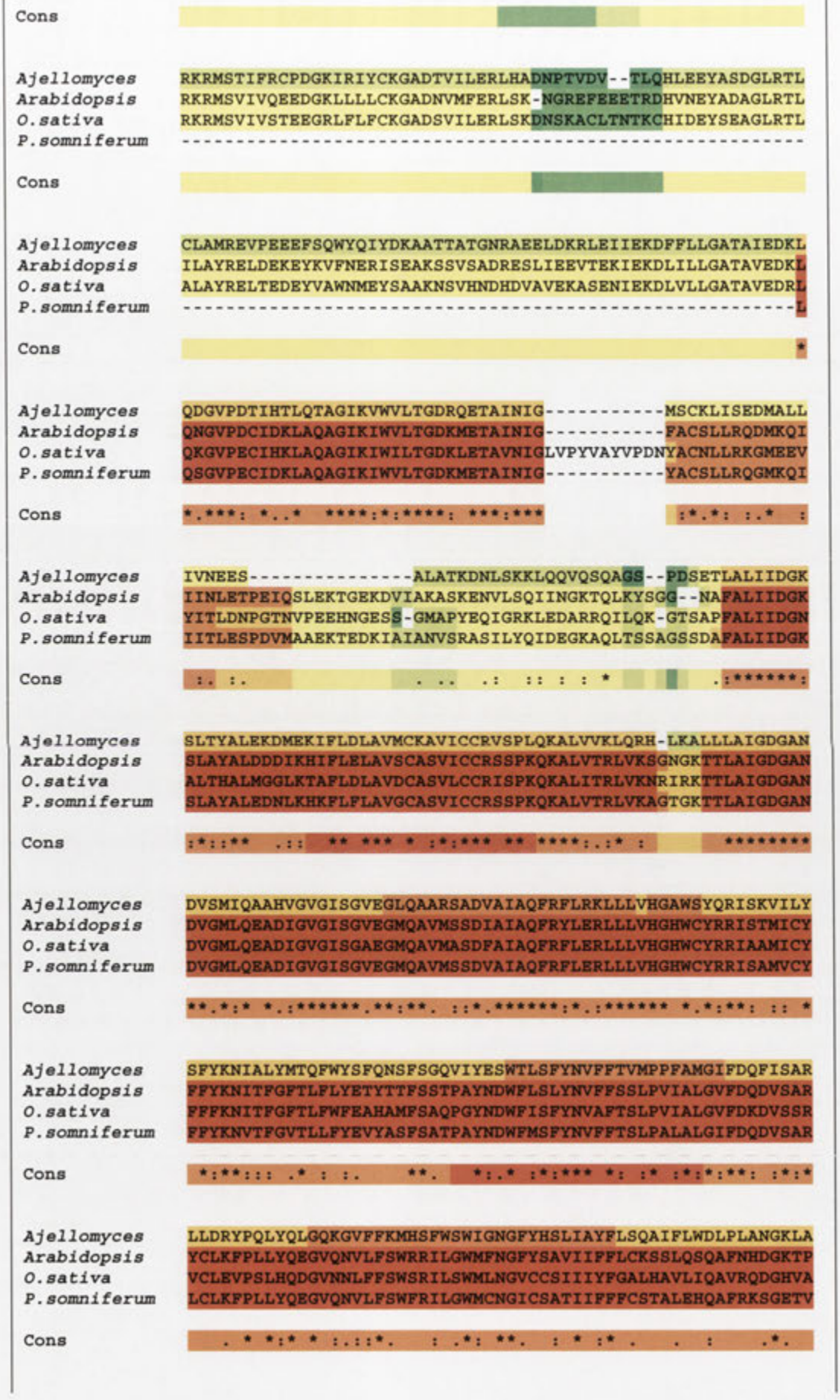


Figure 4.12: Nucleotide alignment of the four signal recognition receptor like proteins (83F5, 83G7, 91E9, 91G3) from the *top1* microarray experiments. The four contigs are homologous. The blue and white coloured nucleotides represent differences between the four sequences. These are possibly due to sequencing errors. Alignment undertaken using AlignX® A component of VNTI Suite 8 InforMax, Inc.




Ajellomyces ALQQIPNISPTNRYTTIAPLAVVLLVSAIKELVGDWKRKTSDKSLNYSRAQVLKGS-
Arabidopsis VLAFTP-LAPYTASSAIVPLLFVIGATMVKEGVEDWRRQKQDNEVNNRKVKVHRGDG
O.sativa CVSFSP-LAPYRAVSVLLPLVVVVGAAMAKEAVEDWRRKQQDIEVNSRKVEVYDGTQ
P.somniferum
Cons
Ajellomyces TFEDTKWINVAVGDIVKVESEQPFPADLVLLASSEPEGLCYIETANLDGETNLKIKQ
Arabidopsis SFDAKEWKTLSIGDIVKVEKNEFFPADLVLLSSSYEDAICYVETMNLDGETNLKVKQ
O.sativa SFHQTEWKKLQVGDIVKVKKDEFFPADLVLLSSSYEDGICYVETMNLDGETNLKRKQ
P.somniferum
Cons
Ajellomyces AIPETADLVSPSQLGRLTGRIKSEQPNSSLYTYEATLTLQAGGGEKELALNPDQLLL
Arabidopsis GLEVTSSLRDEFNFKGFEAFVKCEDPNANLYSFVGTMELKG----AKYPLSPQQLLL
O.sativa SLDVTAGLNEDHSFHTFKAFIQCEDPNEKLYSFLGTLHYNG----QQYPLSPQQILL
P.somniferum
Cons
Ajellomyces RGATLRNTPWIHGLVVFTGHETKLMRNATATPIKRTAVERMVNLQILMLVGILLILS
Arabidopsis RDSKLRNTDFIFGAVIFTGHDTKVIQNSTDPPSKRSMIEKKMDKIIYLMFFMVITMA
O.sativa RDSKLRNTNQIYGIVIFTGHDTKVMQNAMEPPSKRSSVERRMDKIIYLLFVILFAIA
P.somniferum
Cons
Ajellomyces LISSIGHLV-VRMKSADELIYL----------YIGNVNAAQQFFSDIFTYWVLYSNL
Arabidopsis FIGSVIFGVTTRDDLKDGVMKRWYLRPDSSSIFFDPKRAPVAAIYHFLTAVMLYSYF
O.sativa SFGSVMFGIRTRAELSAGNYA-WYLRPDNSTMYFDPNRATLAAICHFLTSLMLYVCL
P.somniferum
Cons
Ajellomyces VPISLFVTIEIVKYYHAFLINSDLDIYYDKTDTSATCRTSSLVEELGQIEYIFSDKT
Arabidopsis IPISLYVSIEIVKVLQSIFINQDIHMYYEEADKPARARTSNLNEELGQVDTILSDKT
O.sativa VPISLYISIEIVKVLQSTFINQDQNMYCEESDKPARARTSNLNEELGQVHTILSDKT
P.somniferum
Cons
Ajellomyces GTLTCNMMEFKQCSIGGLQYAEVVSEDRRVVD---------GDDSE-----------
Arabidopsis GTLTCNSMEFIKCSVAGTAYGRGVTEVEMAMGRRKGGPLVFQSDENDIDMEYSKEAI
O.sativa GTLTCNSMEFLKCSIAGVAYGNRPIEVQMPYG---------GIEEECVDIG-QKGAV
P.somniferum
Cons
Ajellomyces --------MGMYDFNQLVEHLTSHPTRTAIHHFLCLLATCHTVIPERKAEKPDVIKY
Arabidopsis TEESTVKGFNFRDERIMNGNWVTETHADVIQKFFRLLAVCHTVIPEVD-EDTEKISY
O.sativa KSVRPVKGFNFTDDRLMNGQWSKECHQDVIEMFFRVLAVCHTAIPVAD-RTSGGMSY
P.somniferum
Cons
Ajellomyces QAASPDEGALVEGAVMMGYRFTNRRPKSV------IISANGQEQEFELLAVCEFNST
Arabidopsis EAESPDEAAFVIAARELGFEFFNRTQTTISVRELDLVSGKRVERLYKVLNVLEFNST
O.sativa EAESPDEGALVAAARELGFEFYHRSQTSISVHEYDPVFGRKVDRTYKLLNTLEFSSA
P.somniferum


Cons
Ajellomyces RKRMSTIFRCPDGKIRIYCKGADTVILERLHADNPTVDV--TLQHLEEYASDGLRTL
Arabidopsis RKRMSVIVQEEDGKLLLLCKGADNVMFERLSK-NGREFEEETRDHVNEYADAGLRTL
O.sativa RKRMSVIVSTEEGRLFLFCKGADSVILERLSKDNSKACLTNTKCHIDEYSEAGLRTL
P.somniferum ----------
Cons
Ajellomyces CLAMREVPEEEFSQWYQIYDKAATTATGNRAEELDKRLEIIEKDFFLLGATAIEDKL
Arabidopsis ILAYRELDEKEYKVFNERISEAKSSVSADRESLIEEVTEKIEKDLILLGATAVEDKL
O.sativa ALAYRELTEDEYVAWNMEYSAAKNSVHNDHDVAVEKASENIEKDLVLLGATAVEDRL
P.somniferum ----------L
Cons *
Ajellomyces QDGVPDTIHTLQTAGIKVWVLTGDRQETAINIG----------MSCKLISEDMALL
Arabidopsis QNGVPDCIDKLAQAGIKIWVLTGDKMETAINIG----------FACSLLRQDMKQI
O.sativa QKGVPECIHKLAQAGIKIWILTGDKLETAVNIGLVPYVAYVPDNYACNLLRKGMEEV
P.somniferum QSGVPECIDKLAQAGIKIWVLTGDKMETAINIG----------YACSLLRQGMKQI
Cons
Ajellomyces IVNEES---------------ALATKDNLSKKLQQVQSQAGS--PDSETLALIIDGK
Arabidopsis IINLETPEIQSLEKTGEKDVIAKASKENVLSQIINGKTQLKYSGG--NAFALIIDGK
O.sativa YITLDNPGTNVPEEHNGESS-GMAPYEQIGRKLEDARRQILQK-GTSAPFALIIDGN
P.somniferum IITLESPDVMAAEKTEDKIAIANVSRASILYQIDEGKAQLTSSAGSSDAFALIIDGK
Cons
Ajellomyces SLTYALEKDMEKIFLDLAVMCKAVICCRVSPLQKALVVKLQRH-LKALLLAIGDGAN
Arabidopsis SLAYALDDDIKHIFLELAVSCASVICCRSSPKQKALVTRLVKSGNGKTTLAIGDGAN
O.sativa ALTHALMGGLKTAFLDLAVDCASVLCCRISPKQKALITRLVKNRIRKTTLAIGDGAN
P.somniferum SLAYALEDNLKHKFLFLAVGCASVICCRSSPKQKALVTRLVKAGTGKTTLAIGDGAN
Cons
Ajellomyces DVSMIQAAHVGVGISGVEGLQAARSADVAIAQFRFLRKLLLVHGAWSYQRISKVILY
Arabidopsis DVGMLQEADIGVGISGVEGMQAVMSSDIAIAQFRYLERLLLVHGHWCYRRISTMICY
O.sativa DVGMLQEADIGVGISGAEGMQAVMASDFAIAQFRFLERLLLVHGHWCYRRIAAMICY
P.somniferum DVGMLQEADIGVGISGVEGMQAVMSSDVAIAQFRFLERLLLVHGHWCYRRISAMVCY
Cons
Ajellomyces SFYKNIALYMTQFWYSFQNSFSGQVIYESWTLSFYNVFFTVMPPFAMGIFDQFISAR
Arabidopsis FFYKNITFGFTLFLYETYTTFSSTPAYNDWFLSLYNVFFSSLPVIALGVFDQDVSAR
O.sativa FFFKNITFGFTLFWFEAHAMFSAQPGYNDWFISFYNVAFTSLPVIALGVFDKDVSSR
P.somniferum FFYKNVTFGVTLLFYEVYASFSATPAYNDWFMSFYNVFFTSLPALALGIFDQDVSAR
Cons
Ajellomyces LLDRYPQLYQLGQKGVFFKMHSFWSWIGNGFYHSLIAYFLSQAIFLWDLPLANGKLA
Arabidopsis YCLKFPLLYQEGVQNVLFSWRRILGWMFNGFYSAVIIFFLCKSSLQSQAFNHDGKTP
O.sativa VCLEVPSLHQDGVNNLFFSWSRILSWMLNGVCCSIIIYFGALHAVLIQAVRQDGHVA
P.somniferum LCLKFPLLYQEGVQNVLFSWFRILGWMCNGICSATIIFFFCSTALEHQAFRKSGETV
Cons

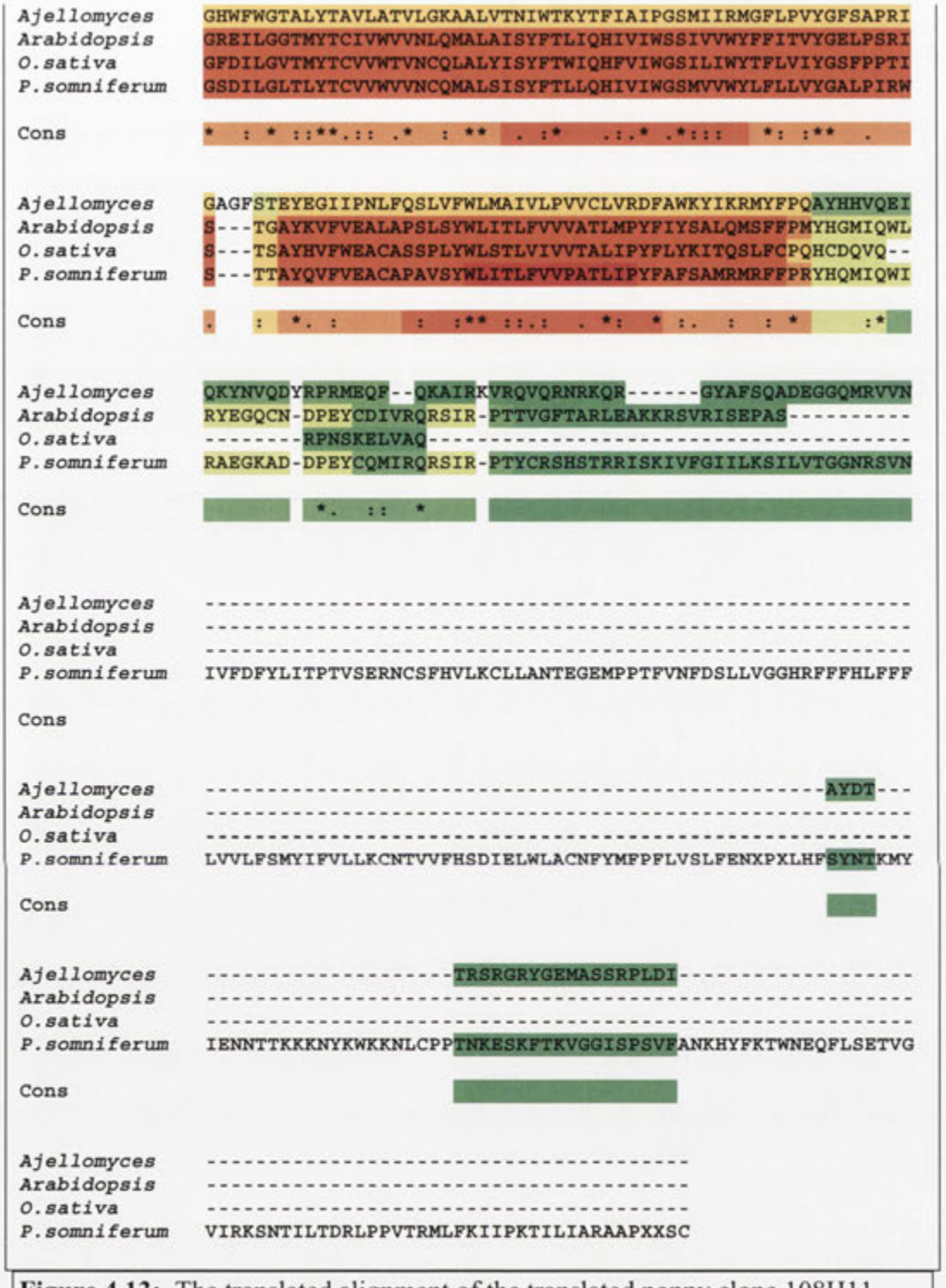


Figure 4.13: The translated alignment of the translated poppy clone 108H11 with three ATPase type transporter proteins from the NCBI genbank database.

Sequence homology is represented by the colour code-

BAD AVG GOOD

Ajellomyces: putative calcium transporting ATPase [*Ajellomyces capsulatus*].

Accession number: AAF90186.

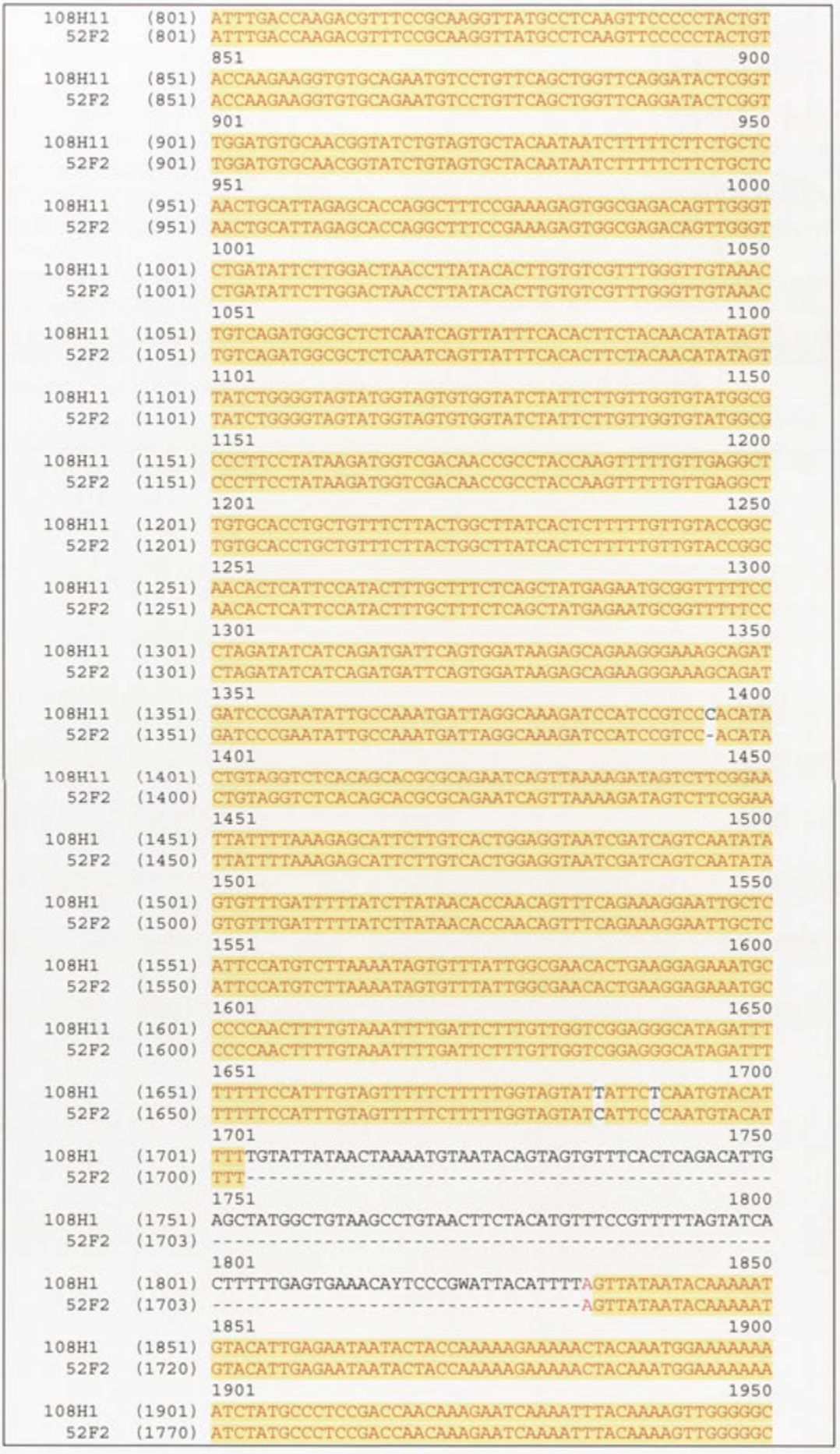
Arabidopsis: haloacid dehalogenase-like hydrolase family protein [*Arabidopsis thaliana*]. Accession number: NP_177038.

O. sativa: putative ATPase [*Oryza sativa* (japonica cultivar-group)]. Accession number: BAD05408.

P. somniferum: Clone 108H11 from the *top1* microarray results.

T-COFFEE, Version_1.41 was used for the amino acid alignments (Notredame *et al.*, 2000).

```
                    1                                                  50
108H11        (1)   CTCCAGTCTGGGGTACCAGAATGCATAGACAAACTCGCACAGGCTGGAAT
52F2          (1)   CTCCAGTCTGGGGTACCAGAATGCATAGACAAACTCGCACAGGCTGGAAT
                    51                                                 100
108H11       (51)   AAAGATATGGGTTTTGACTGGGGATAAGATGGAGACTGCCATCAATATTG
52F2         (51)   AAAGATATGGGTTTTGACTGGGGATAAGATGGAGACTGCCATCAATATTG
                    101                                                150
108H11      (101)   GGTATGCATGCAGTTTGCTTAGGCAAGGAATGAAACAAATTATTATCACC
52F2        (101)   GGTATGCATGCAGTTTGCTTAGGCAAGGAATGAAACAAATTATTATCACC
                    151                                                200
108H11      (151)   TTGGAGTCACCTGATGTTATGGCAGCTGAGAAAACGGAGGACAAAATTGC
52F2        (151)   TTGGAGTCACCTGATGTTATGGCAGCTGAGAAAACGGAGGACAAAATTGC
                    201                                                250
108H11      (201)   AATTGCAAATGTATCAAGGGCAAGCATACTTTATCAAATAGATGAAGGGA
52F2        (201)   AATTGCAAATGTATCAAGGGCAAGCATACTTTATCAAATAGATGAAGGGA
                    251                                                300
108H11      (251)   AGGCACAGTTAACTTCTTCAGCCGGAAGCTCAGATGCATTTGCGTTGATC
52F2        (251)   AGGCACAGTTAACTTCTTCAGCCGGAAGCTCAGATGCATTTGCGTTGATC
                    301                                                350
108H11      (301)   ATTGATGGGAAGTCACTCGCGTATGCCCTTGAAGACAATTTGAAGCACAA
52F2        (301)   ATTGATGGGAAGTCACTCGCGTATGCCCTTGAAGACAATTTGAAGCACAA
                    351                                                400
108H11      (351)   GTTTCTATTCCTTGCAGTTGGCTGTGCATCCGTTATTTGTTGCCGCTCAT
52F2        (351)   GTTTCTATTCCTTGCAGTTGGCTGTGCATCCGTTATTTGTTGCCGCTCAT
                    401                                                450
108H11      (401)   CACCTAAACAGAAAGCACTTGTAACACGTCTGGTTAAAGCTGGAACAGGT
52F2        (401)   CACCTAAACAGAAAGCACTTGTAACACGTCTGGTTAAAGCTGGAACAGGT
                    451                                                500
108H11      (451)   AAAACAACATTAGCAATTGGAGACGGGGCCAACGATGTGGGTATGCTTCA
52F2        (451)   AAAACAACATTAGCAATTGGAGACGGGGCCAACGATGTGGGTATGCTTCA
                    501                                                550
108H11      (501)   AGAAGCAGATATTGGGGTCGGAATTAGCGGAGTTGAAGGAATGCAGGCAG
52F2        (501)   AGAAGCAGATATTGGGGTCGGAATTAGCGGAGTTGAAGGAATGCAGGCAG
                    551                                                600
108H11      (551)   TCATGTCAAGTGATGTTGCAATTGCTCAGTTCCGATTTTTGGAGCGTCTG
52F2        (551)   TCATGTCAAGTGATGTTGCAATTGCTCAGTTCCGATTTTTGGAGCGTCTG
                    601                                                650
108H11      (601)   CTACTTGTACATGGTCATTGGTGTTACAGAAGAATATCAGCAATGGTATG
52F2        (601)   CTACTTGTACATGGTCATTGGTGTTACAGAAGAATATCAGCAATGGTATG
                    651                                                700
108H11      (651)   CTACTTCTTTTACAAGAATGTTACATTCGGTGTCACTCTGTTGTTCTACG
52F2        (651)   CTACTTCTTTTACAAGAATGTTACATTCGGTGTCACTCTGTTGTTCTACG
                    701                                                750
108H11      (701)   AGGTGTATGCATCATTCTCTGCCACACCTGCCTATAACGACTGGTTTATG
52F2        (701)   AGGTGTATGCATCATTCTCTGCCACACCTGCCTATAACGACTGGTTTATG
                    751                                                800
108H11      (751)   TCATTCTATAATGTTTTCTTCACATCATTACCTGCTCTTGCTCTTGGCAT
52F2        (751)   TCATTCTATAATGTTTTCTTCACATCATTACCTGCTCTTGCTCTTGGCAT
                    801                                                850
```


108H11 (801) ATTTGACCAAGACGTTTCCGCAAGGTTATGCCTCAAGTTCCCCCTACTGT
52F2 (801) ATTTGACCAAGACGTTTCCGCAAGGTTATGCCTCAAGTTCCCCCTACTGT
851 900
108H11 (851) ACCAAGAAGGTGTGCAGAATGTCCTGTTCAGCTGGTTCAGGATACTCGGT
52F2 (851) ACCAAGAAGGTGTGCAGAATGTCCTGTTCAGCTGGTTCAGGATACTCGGT
901 950
108H11 (901) TGGATGTGCAACGGTATCTGTAGTGCTACAATAATCTTTTTCTTCTGCTC
52F2 (901) TGGATGTGCAACGGTATCTGTAGTGCTACAATAATCTTTTTCTTCTGCTC
951 1000
108H11 (951) AACTGCATTAGAGCACCAGGCTTTCCGAAAGAGTGGCGAGACAGTTGGGT
52F2 (951) AACTGCATTAGAGCACCAGGCTTTCCGAAAGAGTGGCGAGACAGTTGGGT
1001 1050
108H11 (1001) CTGATATTCTTGGACTAACCTTATACACTTGTGTCGTTTGGGTTGTAAAC
52F2 (1001) CTGATATTCTTGGACTAACCTTATACACTTGTGTCGTTTGGGTTGTAAAC
1051 1100
108H11 (1051) TGTCAGATGGCGCTCTCAATCAGTTATTTCACACTTCTACAACATATAGT
52F2 (1051) TGTCAGATGGCGCTCTCAATCAGTTATTTCACACTTCTACAACATATAGT
1101 1150
108H11 (1101) TATCTGGGGTAGTATGGTAGTGTGGTATCTATTCTTGTTGGTGTATGGCG
52F2 (1101) TATCTGGGGTAGTATGGTAGTGTGGTATCTATTCTTGTTGGTGTATGGCG
1151 1200
108H11 (1151) CCCTTCCTATAAGATGGTCGACAACCGCCTACCAAGTTTTTGTTGAGGCT
52F2 (1151) CCCTTCCTATAAGATGGTCGACAACCGCCTACCAAGTTTTTGTTGAGGCT
1201 1250
108H11 (1201) TGTGCACCTGCTGTTTCTTACTGGCTTATCACTCTTTTTGTTGTACCGGC
52F2 (1201) TGTGCACCTGCTGTTTCTTACTGGCTTATCACTCTTTTTGTTGTACCGGC
1251 1300
108H11 (1251) AACACTCATTCCATACTTTGCTTTCTCAGCTATGAGAATGCGGTTTTTCC
52F2 (1251) AACACTCATTCCATACTTTGCTTTCTCAGCTATGAGAATGCGGTTTTTCC
1301 1350
108H11 (1301) CTAGATATCATCAGATGATTCAGTGGATAAGAGCAGAAGGGAAAGCAGAT
52F2 (1301) CTAGATATCATCAGATGATTCAGTGGATAAGAGCAGAAGGGAAAGCAGAT
1351 1400
108H11 (1351) GATCCCGAATATTGCCAAATGATTAGGCAAAGATCCATCCGTCCCACATA
52F2 (1351) GATCCCGAATATTGCCAAATGATTAGGCAAAGATCCATCCGTCC-ACATA
1401 1450
108H11 (1401) CTGTAGGTCTCACAGCACGCGCAGAATCAGTTAAAAGATAGTCTTCGGAA
52F2 (1400) CTGTAGGTCTCACAGCACGCGCAGAATCAGTTAAAAGATAGTCTTCGGAA
1451 1500
108H1 (1451) TTATTTTAAAGAGCATTCTTGTCACTGGAGGTAATCGATCAGTCAATATA
52F2 (1450) TTATTTTAAAGAGCATTCTTGTCACTGGAGGTAATCGATCAGTCAATATA
1501 1550
108H1 (1501) GTGTTTGATTTTTATCTTATAACACCAACAGTTTCAGAAAGGAATTGCTC
52F2 (1500) GTGTTTGATTTTTATCTTATAACACCAACAGTTTCAGAAAGGAATTGCTC
1551 1600
108H1 (1551) ATTCCATGTCTTAAAATAGTGTTTATTGGCGAACACTGAAGGAGAAATGC
52F2 (1550) ATTCCATGTCTTAAAATAGTGTTTATTGGCGAACACTGAAGGAGAAATGC
1601 1650
108H11 (1601) CCCCAACTTTTGTAAATTTTGATTCTTTGTTGGTCGGAGGGCATAGATTT
52F2 (1600) CCCCAACTTTTGTAAATTTTGATTCTTTGTTGGTCGGAGGGCATAGATTT
1651 1700
108H1 (1651) TTTTTCCATTTGTAGTTTTTCTTTTTGGTAGTATTATTCTCAATGTACAT
52F2 (1650) TTTTTCCATTTGTAGTTTTTCTTTTTGGTAGTATCATTCCCAATGTACAT
1701 1750
108H1 (1701) TTTTGTATTATAACTAAAATGTAATACAGTAGTGTTTCACTCAGACATTG
52F2 (1700) TTT----------
1751 1800
108H1 (1751) AGCTATGGCTGTAAGCCTGTAACTTCTACATGTTTCCGTTTTTAGTATCA
52F2 (1703) ----------
1801 1850
108H1 (1801) CTTTTTGAGTGAAACAYTCCCGWATTACATTTTAGTTATAATACAAAAAT
52F2 (1703) ----------AGTTATAATACAAAAAT
1851 1900
108H1 (1851) GTACATTGAGAATAATACTACCAAAAAGAAAAACTACAAATGGAAAAAAA
52F2 (1720) GTACATTGAGAATAATACTACCAAAAAGAAAAACTACAAATGGAAAAAAA
1901 1950
108H1 (1901) ATCTATGCCCTCCGACCAACAAAGAATCAAAATTTACAAAAGTTGGGGGC
52F2 (1770) ATCTATGCCCTCCGACCAACAAAGAATCAAAATTTACAAAAGTTGGGGGC

```
                1951                                                 2000
108H1   (1951) ATTTCTCCTTCAGTGTTCGCCAATAAACACTATTTTAAGACATGGAATGA
 52F2   (1820) ATTTCTCCTTCAGTGTTCGCCAATAAACACTATTTTAAGACATGGAATGA
                2001                                                 2050
108H1   (2001) GCAATTCCTTTCTGAAACTGTTGGTGTTATAAGATAAAAATCAAACACTA
 52F2   (1870) GCAATTCCTTTCTGAAACTGTTGGTGTTATAAGATAAAAATCAAACACTA
                2051                                                 2100
108H1   (2051) TATTGACTGATCGATTACCTCCAGTGACAAGAATGCTCTTTAAAATAATT
 52F2   (1920) TATTGACTGATCGATTACCTCCAGTGACAAGAATGCTCTTTAAAATAATT
                2101                                             2148
108H1   (2101) CCGAAGACTATCTTAATAGCTCGAGCTGCTCC--CWGTARATCATGCC
 52F2   (1970) CCGAAGACTATCTTT-TAACTGATTCTGCGCGTGCTKGAGACCAATC-
```

Figure 4.14: The nucleotide alignment of the two cDNAs 52F2 and 108H11. The two clones are almost identical (Identity 92.9%) except the cDNA clone 108H11 contains 130 extra nucleotides in the 3' untranslated region (highlighted in green). This indicates that two this gene is encoded by at least two different genes. The translated sequence alignment with flippase genes from other species shown in the previous Figure 4.17. Alignment obtained using AlignX® A component of VNTI Suite 8 InforMax, Inc.

4.5 Germplasm with Altered Morphinan Alkaloid Yields

4.5.1 *P. somniferum* cv. Marianne

A total of five slides were used for the *P. somniferum* cv C052 versus *P. somniferum* cv. Marianne experiments. The microarray experiments included 3 biological replicates and a technical replicate, dye swap, for two of the biological replicates. The analysis revealed a total of 460 clones, which were differentially expressed between the two cultivars (Figure 4.15).

Of the 460 differentially expressed genes 269 were more highly expressed in Marianne than in C052 using a cut-off expression ratio value of 2. The other 191 genes were more highly expressed in the high morphine producing C052 (Appendix 3). None of the control sequences printed on the array were differentially expressed between the two cultivars. Due to the large number of differentially expressed genes sequencing was not undertaken on the whole set. Rather the results were compared to the sets of differentially expressed genes with the other genotypes and mutants (see section 4.6).

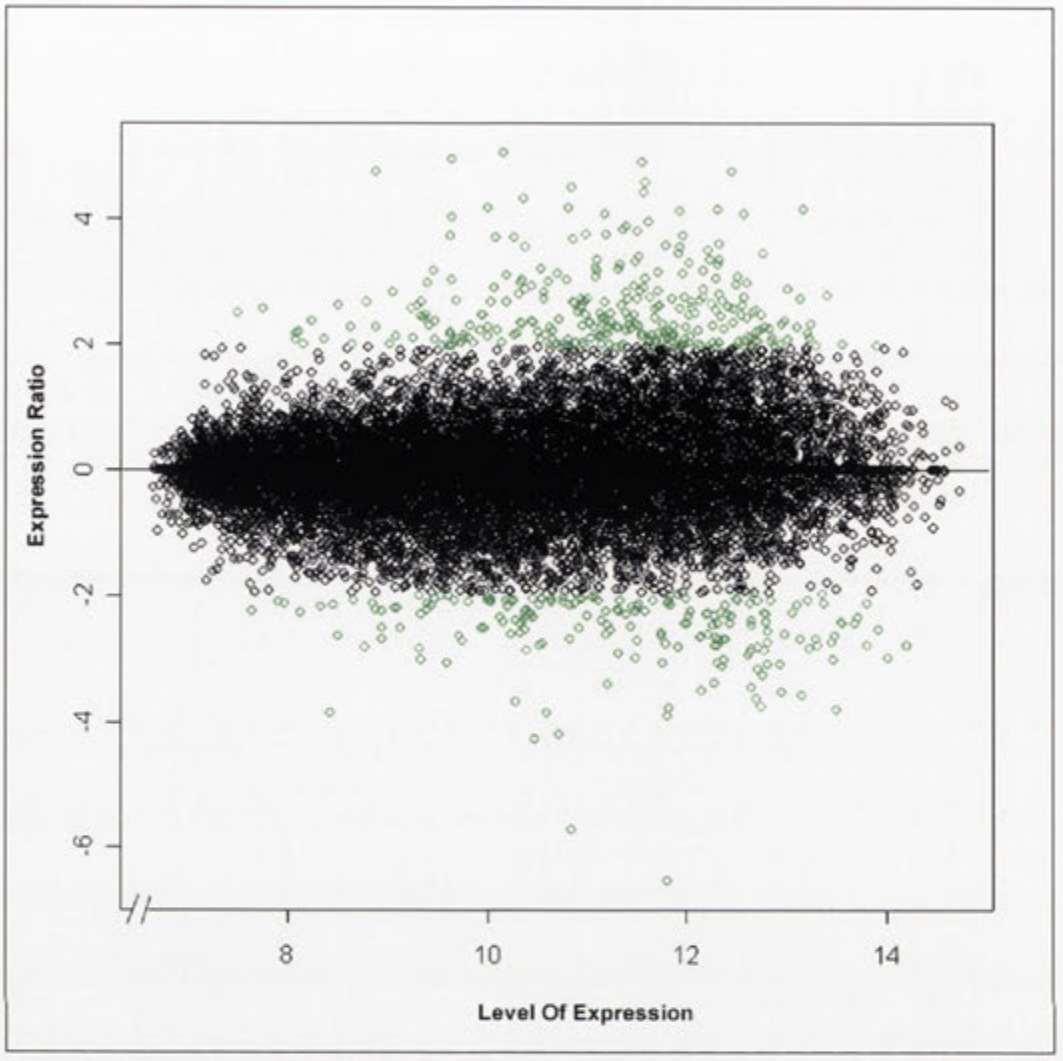


Figure 4.15: Graphical representation of the differentially expressed genes between *P. somniferum* cv. C052 and *P. somniferum* cv. Marianne. Using the differential ratio cut-off of 2 revealed a total of 460 genes [the green circles] were differentially expressed. 269 of these genes were more highly expressed in the Marianne cultivar than in the high morphine line C052. The remaining 191 genes were more highly expressed in the high morphine producing line C052. No control sequences were differentially expressed between the two cultivars.

4.5.2 *P. somniferum* Low Morphine Accumulating Mutant D205

Microarray experiments were undertaken using the procedures detailed in Chapter 2. Only two slides were used for experiments with mutant D205 and these were dye swaps of the one RNA preparation. To reduce the effect of biological differences the single biological sample for each mutant was prepared

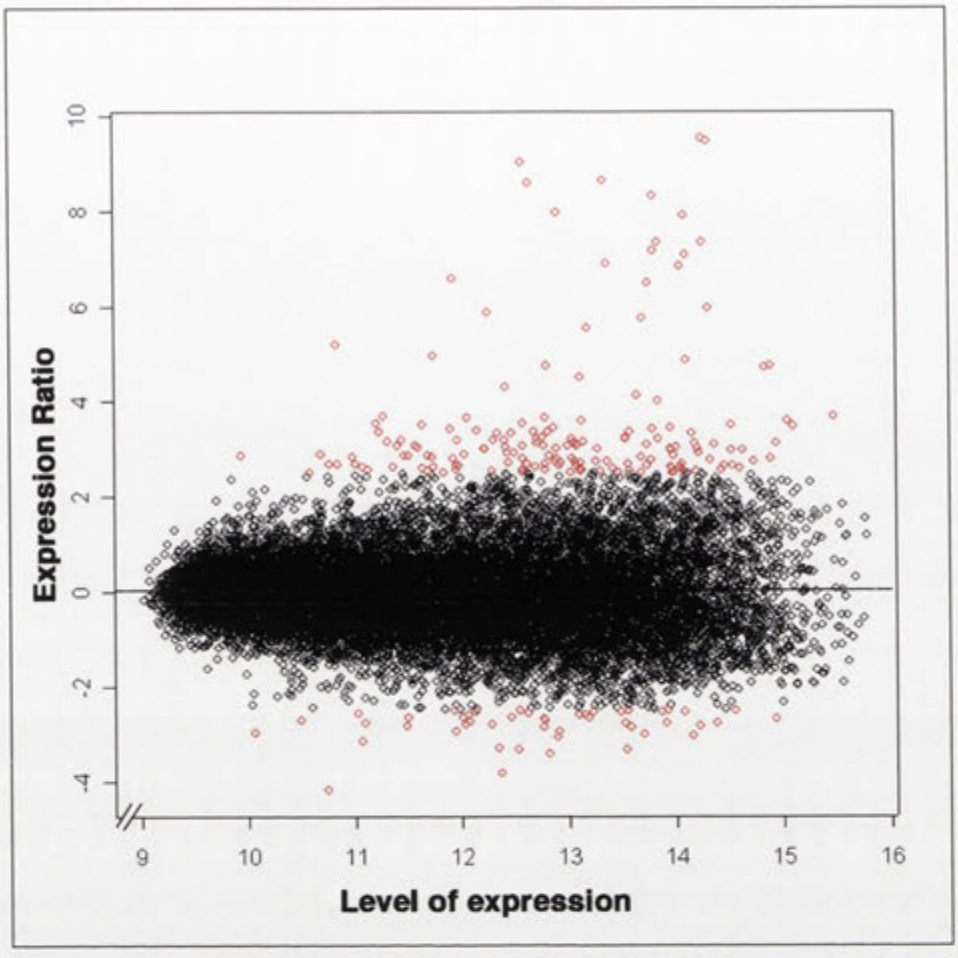


Figure 4.16: Graphical representation of the differentially expressed genes (red circles) between the low morphine-accumulating mutant, D205 and the parental control C052.

by pooling six independent mutant plants. A total of eighty-two cDNAs were differentially under-expressed in the mutant and only eight were over-expressed in the mutant compared to the control (Figure 4.16). For the full list of differentially expressed genes see Appendix 4. Due to the large number of clones found to be differential, sequencing of the entire set was not undertaken.

4.5.3 *P. somniferum* High Total Alkaloid Accumulating Mutant D242

Microarray experiments were undertaken using the procedures detailed in Chapter 2. Only two slides were used for experiments with mutant D242 and these were dye swaps of the one RNA preparation. To reduce the effect of

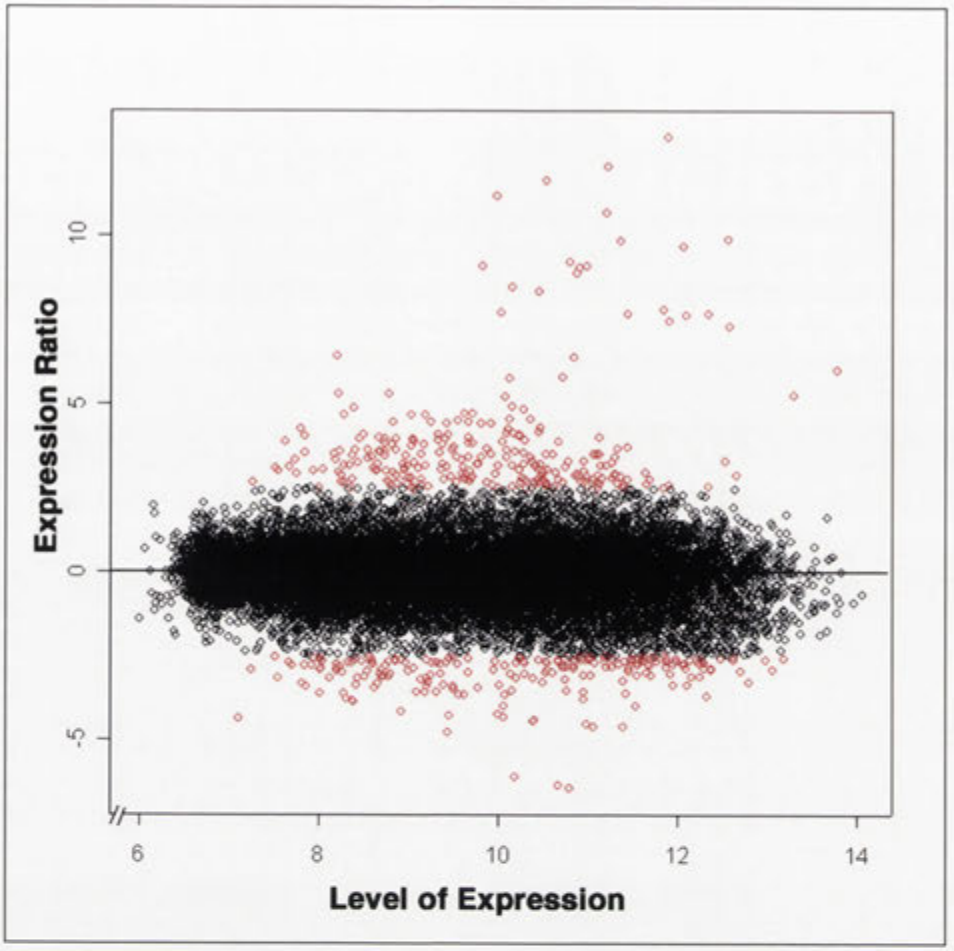


Figure 4.17: Graphical representation of the differentially expressed genes (red circles) of the high thebaine and oripavine mutant, D242 compared to the parental control C052.

biological differences the single biological sample for each mutant was prepared by pooling six independent mutant plants. A total of eighty-six cDNAs were differentially under-expressed in the mutant and only fourteen were over-expressed in the mutant compared to the control (Figure 4.17). For the full list of differentially expressed genes see Appendix 5. Due the large number of clones found to be differential, sequencing of the entire set was not undertaken.

4.5.4 *P. somniferum* Low Morphinan Alkaloid Mutant High Papaverine Accumulating Mutant D120

Microarray experiments were undertaken using the procedures detailed in Chapter 2. Only two slides were used for experiments with mutant D120 and

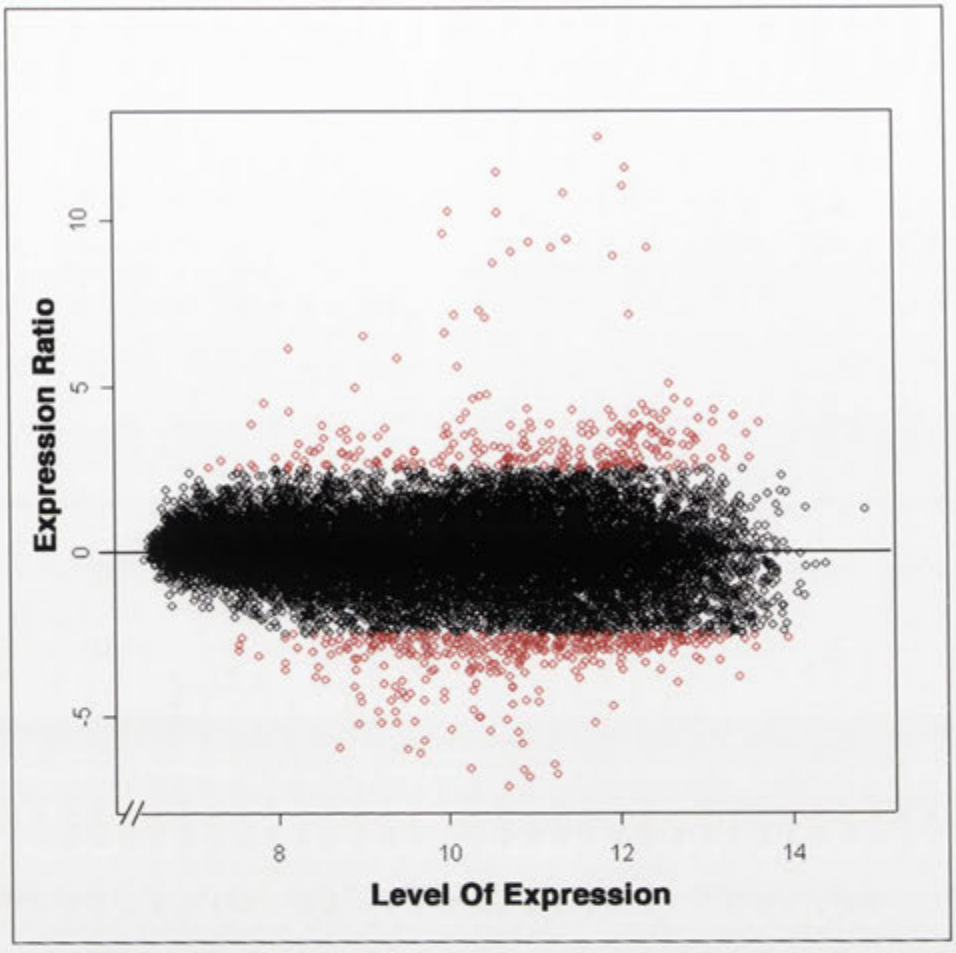


Figure 4.18: Graphical representation of the differentially expressed genes (red circles) of the high papaverine-accumulating mutant, D120 compared to the parental control C052.

these were dye swaps of the one RNA preparation. To reduce the effect of biological differences the single biological sample for each mutant was prepared by pooling six independent mutant plants.

The microarray analysis revealed one hundred and four cDNAs to be under-expressed and ninety-six were over-expressed in the mutant compared to the control parental phenotype C052 and the papaverine-accumulating mutant D120 (Figure 4.18). For a full list of the differentially expressed genes see Appendix 6. Further analyses of this mutant may provide useful leads to the nature of regulatory control at the point of (*S*)-coclaurine in the pathway.

Further microarray experiments would obviously reduce the number of candidate genes, however resources were diverted to the next two subchapters rather than concentrating on the germplasm in this subchapter. While sequencing was not undertaken of these complete sets of differential genes, these uncharacterised differential sets were compared to each other and the differentials of the other experiments (section 4.6) in deciding on a set for sequencing and further analysis.

4.6 Comparisons Across The Experiments

As a further focus to the *top1* experiments the ten genes that were differentially expressed were then analysed for their differential expression across the other germplasm that were tested in this thesis (Table 4.7). This was undertaken to establish which genes of the set of ten were implicated as either affected by the morphinan pathway or involved in the pathway. The entire set of ten genes were also differentially expressed in: the other thebaine only accumulating mutant (F075); the low morphine and high papaverine accumulating mutant D120; the low morphine mutant D205; the high total morphinan accumulating mutant D242; and the *E. californica* germplasm. The only lines that these genes were not differentially expressed were in: the reticuline accumulating germplasm of the mutant E40/41; the Ir-cor reticuline accumulating transgenic and the low morphinan alkaloid accumulating *P. somniferum* cultivar Marianne.

Table 4.7: The list of ten differential genes from the *top1* microarray and their respective differential expression values, if any, in the other microarray experiments undertaken in this thesis. The codeinone reductase RNAi experiment Ir-cor, Reticuline accumulating mutant (E40/41) and the low morphine *P. somniferum* cv Marianne are listed in the same column, as none of the genes from the *top1* experiment were differential.

Name	F075 (Thebaine)	E40/41 (Reticuline), Ir-cor (Reticuline), *P.somniferum* cv Marianne	D120 (Low M, high Pap)	D242 (High M+C+T+O)	D205 (Low M)	*E. californica*
116D7	5.1	nd	6.3	4.3	6.5	6.3
91G3	9.0	nd	14.1	8.7	5.9	6.3
27D12	10.2	nd	11.6	10.0	7.5	7.7
157C4	6.8	nd	6.8	5.4	5.5	9.1
91B6	6.2	nd	10.3	6.2	7.2	2.7
221C9	6.6	nd	7.6	5.3	3.6	4.6
54F9	5.8	nd	8.6	4.9	2.8	4.8
52F2	6.3	nd	7.2	6.8	4.8	6.3
151H8	8.8	nd	7.3	7.6	3.9	5.3
111C7	7.7	nd	8.8	7.7	6.4	7.1

D242 is high oripavine and thebaine. High total M (morphine)+C (codeine)+O (oripavine)+T (thebaine).

D120 is the high (Pap) papaverine low morphine mutant.

D205 is the low morphine mutant.

Reticuline is the reticuline-accumulating mutant E40/41.

F075 is thebaine only accumulating mutant.

All values are indicative of reduced expression in the test sample compared to the test sample (the wildtype control *P. somniferum*).

nd, represents the gene was not differential in this experiment.

4.7 Summary

The use of microarrays has given more information regarding the expression profile and significance of the set of ten genes that were shown through an exhaustive set of microarray experiments to be differentially expressed in the *top1* mutant. Table 4.7 summarises the results of the expression ratios across each of the different mutants and species comparison experiments, excluding the codeinone reductase RNAi transgenic and *P. somniferum* cv. Marianne experiments, in which none of the ten genes were differential.

The experiments described in Chapter 4 have implicated a small number of genes to be associated with the pathway. The set of ten was arrived at after highly repeated experiments with *top1* mutation; furthermore their expression is consistently affected by other perturbations in the alkaloid pathway (Table 4.7). It was not necessarily expected that the microarray analyses would reveal the mutated gene in the various studied mutation lines, and none of the differential genes can be considered a strong candidate as the causal gene for any of the mutations. However the consistently differential genes are implicated to be closely associated with the pathway, directly or indirectly. The strongest evidence of association is for the set of ten genes from the twelve *top1*-microarray experiments. In addition several other genes were differential with other mutants or transgenics that are possible candidates for alkaloid transport and control and warrant further consideration and experimentation.

The codeinone reductase RNAi transgenics have an almost identical alkaloid phenotype to the reticuline-accumulating mutant E40/41. The microarray results for these two genotypes however are quite different. None of the ten genes differentially expressed in *top1* are differential in the codeinone reductase RNAi transgenics.

Likewise none of the ten were significantly differentially expressed in the EMS generated reticuline-accumulating mutant, E40/41. However the two sets of differential genes for these two reticuline accumulating genotypes, a mutant and a knockout transgenic, are also quite different to each other.

The blast results from the ten differentially expressed genes from the *top1* microarray experiments, does not suggest any as candidates for the demethylation of thebaine to neopinone. If the gene responsible for the demethylation reaction is a P450, as suggested by homology to the codeine and thebaine detoxifying protein in rat microsomes (Kodaira and Spector, 1994), it is possible that the P450 is one of the unknown genes from the array. P450 enzymes vary remarkably at both the nucleotide and amino acid level (reviewed in Chapple 1998). If however none of the differential genes is the implicated demethylase, some of the possible explanations for failing to detect a convincing candidate for the demethylase reaction are:

- The cDNA responsible for the *top1* phenotype is of low abundance and was not printed on the array.
- An homologous gene (or genes) is expressed to such an extent that it masks the difference in expression of the mutated demethylase gene.
- The leaf material used for expression analysis is the wrong material and the gene is not expressed sufficiently in that tissue. Counter to this argument is the fact that morphine and codeine can be detected in leaf extracts. It is still unknown if the morphinan alkaloids are made in the leaf parenchyma cells and transported into the latex, however several genes have been shown to be expressed in leaf tissue (Kutchan and Dittrich 1995; Fachinni and De Luca 1994 and 1995b; Rosco *et al.*, 1997; Pauli and Kutchn 1998; Unterlinner *et al.*, 1999; Facchini and Park 2003; Ounaroon *et al.*, 2003; Weid *et al.*, 2004).
- The transcript levels of the demethylase gene(s) may remain unaffected in the mutant even though the mutation affects the catalytic activity, transport and/or translational modification process of the enzyme. Further to this post transcriptional control may be taking place affecting the level of translation and subsequently the level of enzyme activity. The transcript of the demethylase may still be produced at the same levels but the lifetime of the RNA may be altered again affecting the level of translated product but not necessarily the transcript levels seen on the

array. Similarly alternative splicing of the RNA could affect translation levels but not message levels.

Of course it also conceivable that *top1* does not directly affect the demethylation reaction, rather it is a structural change, which prevents the thebaine being delivered to the compartment where the demethylase is active. For example thebaine-accumulating vacuoles in the latex appears to be distinct from the morphine accumulating vacuoles in size and density (Pham and Roberts, 1991).

It is also possible that one or more of the *top1* ten differential genes is physically linked to the mutation responsible for the *top1* phenotype and its transcription is affected by the mutation. An argument against this hypothesis is that the same ten genes are also differentially expressed in all but one of the morphinan alkaloid mutants studied in this thesis and the low morphine *P. somniferum* cv Marianne. This suggests that these ten genes are not specifically causative of any of the mutations but are affected by the different perturbations to the alkaloid pathway. This chapter will investigate, where possible, each of the set of ten genes to establish a possible link with the alkaloid pathway. The analysis of these genes will include transcript analysis to further authenticate the set of ten genes as differentially expressed.

Some of the candidate genes have strong matches in the databases, which will allow some speculation of how they are associated with morphine biosynthesis and will guide future experimentation to confirm their roles. The initial analysis might involve comparing the genomic sequences of any of these differentially expressed genes from the mutant with the control to see if there are sequence variations in the structural enzyme. This initial comparison would not include possible differences in the promoter regions, which would require isolation of promoter sequences. A second stage of analysis of differential genes could be to knock out their activity in transgenic control plants; any effects on alkaloids would confirm the association with the pathway; a phenotype mimicking the original mutant would implicate the gene as the point of lesion in the original

mutant. The ultimate confirmation that the gene was the point of lesion would be to "rescue" the wild type phenotype by expression of the gene in the mutant.

The ten differentially expressed *top1* genes will be briefly reviewed in the next chapter, and where possible within the time constraints of this PhD study, further experiments described to advance understanding of the potential association with the morphine pathway.

Chapter 5

Further Characterisation Of The Genes Specific To The *top1* Mutant

5.1 Introduction

The large quantity of secondary metabolites, their structural complexity and their use of nitrogen, an essential component of primary metabolism, are highley suggestive of an important role for secondary metabolites in the plant's fitness and survival. The morphinan pathway is a case in point and recent evidence suggests that the morphinans are under intricate control and regulation (Bird *et al.*, 2003; Weid *et al.*, 2004). Plant breeding programs have been responsible for large increases in morphinan alkaloid yields over recent decades, however even wild *Papaver* species have these complex products accumulating to high concentrations, suggesting they are contributing to fitness.

The *P. somniferum* morphinan alkaloid pathway is important not only because it produces such diverse and complex compounds more efficiently and in larger quantities than can be achieved via *in vitro* synthesis, but because of the insights it offers into how enzymes can perform such complicated reactions. The chemical reactions required to make these compounds, for example the reaction catalysed by the BBE enzyme, are so complex they defy *in vitro* mimicry.

Progress in unravelling the morphinan alkaloid pathway has been constrained by its restriction to a non-model species and by the limited availability of mutants in the pathway. Professor Zenk and co-workers have elucidated many of the enzymatic steps of the pathway from cell suspensions (Rueffer *et al.*, 1981; 1983; Rueffer and Zenk, 1987; Stadler *et al.*, 1988, Frenzel and Zenk, 1990; De-Eknamkul and Zenk,

1990; De-Eknamkul and Zenk, 1992; Gerardy and Zenk, 1993a, b; Lenz and Zenk, 1994; 1995a,c), and the laboratories of Professor Kutchan and Professor Facchini (Facchini and De Luca 1994; Pauli and Kutchan, 1998; Unterlinner *et al.*, 1999; Ounaroon, *et al.*, 2003) have isolated several structural genes. Understanding of the steps and structural genes of the pathway is slowly emerging, however almost nothing is known of its regulation and the transport of the various intermediates and the end products.

Early work has shown that the specialised vacuoles of the laticifers of the opium poppy are responsible for the storage of the various products and intermediates. However surprisingly there is some immunolocalisation results suggesting the enzymatic steps responsible for these products do not occur in these compartments (Bird *et al.*, 2003; Weid *et al.*, 2004). This would suggest that there are several levels of intermediate shuttling occurring via transport vesicles in and out of these vacuoles. Recently proteomics work has been published on the latex of *P. somniferum* revealing a snapshot of the proteins found in this specialised system (Decker *et al.*, 2002). This work is substantial and offers a general insight into some of the proteins found in the latex. However many of the protein spots were not sequenced and for those that were, only short lengths of N-terminal sequencing were obtained. Furthermore there is a lack of annotated *P. somniferum* genes present in the database; consequently Decker et al., 2000 failed to give any functional assignment to many of the proteins.

Several mutants have been isolated at Tasmanian Alkaloids Pty Ltd. and one such mutant *top1* has been characterised using biochemical methods of analysis. Further characterisation of the transcriptome has been undertaken in this thesis using cDNA microarrays to reveal a set of ten differentially expressed genes that are all under-expressed not only in the *top1* mutant but across several independent mutants compared to the control C052.

This chapter will investigate, where possible, the roles of the seven genes with associated function and show further evidence to their involvement in the mutant *top1* phenotype. This will include: the further authentification of the microarray data using quantitative real time PCR analysis; analysis of the ten genes' expression across root, stem and leaf tissue using reverse transcriptase PCR techniques; several of the genes with substantial associated function though blast searches will be further analysed; and finally an F3 population segregating for the *top1* phenotype will be analysed for the ten genes' traits. This work will then be dicscussed as to its relevance and future directions proposed.

5.2 Quantitative Real Time PCR Analysis

Microarrays are beneficial in that they enable a large number of genes to be screened quickly. However a major drawback of cDNA spotted arrays is the chance of cross-hybridisation of families of homologous expressed genes; since the cDNAs are long sequences there are more opportunities for segments of homology to be shared with various gene families. Cross-hybridisation can lead to the dilution of the real value for differential expression of a gene or can sometimes lead to the wrong cDNA being calculated as differentially expressed when in fact it is another differentially expressed gene that has reasonable sequence homology binding to the cDNA on the array. To confirm if the genes that are differentially expressed in the arrays are in fact the right genes semi-quantitative real time-PCR (qRT-PCR) analysis was undertaken. QRT-PCR enables specific primers to be used enabling the determination of expression of single genes if nucleotide differences can be obtained. This method is also more sensitive than PCR analysis and can allow the semi-quantification of the expression of each of the cDNAs to authenticate the microarray results.

Quantitative real time PCR analysis was undertaken on the same tissue type as the microarray experiments using new biological material. The qRT-PCR data

authenticated the microarray data (Table 5.1) showing the genes that could be tested were down regulated in the mutant *top1* compared to the control C052. Data could not be obtained for the gene 52F2. Several primer sets were designed and tested on RNA extracted from both the control and the mutant, but it was not possible to get a single band for qRT-PCR quantification. The presence of several bands suggests that there is possibly a gene family representing this clone with multiple isoforms in the poppy genome, several transcript variants or a combination of both situations making it difficult to get a single band using PCR primers until more sequence data is known of this gene or gene family.

The expression ratios for several of the genes however, differed between the microarray and the qRT-PC calculations. The specificity of the qRT-PCR primers enables targeting of a particular cDNA sequence whereas the cDNAs on the microarray are subject to hybridisation of a number of homologous sequences that may not be differential, reducing the differential expression ratio.

Five of the genes' (116D7, 91G3, 157C4, 91B6 and 151H8) qRT-PCR differential ratios were lower than the corresponding microarray ratios, and two of these five genes, 91B6 and 151H8, had differential ratios less than the arbitrary cut-off of two (Table 5.1). The possible reason for this is that the cDNA on the microarray is binding other homologous cDNAs, all of which are differentially expressed; the qRT-PCR primers may be more specific and detect only some of these differential homologues. Four of the genes, 27D12; 54F9; 221C9 and 111C7, had larger differential expression ratios (27D12: 7.5, 54F9: 6.7, 221C9: 69, and 111C7: 3.2) than the microarray ratios (27D12: 4.7, 54F9: 3.9, 221C9: 3.4, and 111C7: 2.9) (Table 5.1). The possible reason for this is that the qRT-PCR primers were designed to the specific cDNA which is affected by the *top1* mutation and that on the microarray the homologous transcripts are masking the differential effect. A study using a microarray containing synthetic gene fragments with decreasing sequence identities ranging from 100% to 60% revealed a cross hybridisation result between

80-90% (Girke *et al.*, 2000). One gene 221C9 had a 69-fold decrease in expression in the mutant *top1* compared to the wildtype C052.

The qRT-PCR analysis has strengthened the conclusion that at least seven of the ten genes are reproducibly differentially expressed in *top1*. This in turn has strengthened the impetus to find the link between these genes and the morphine pathway.

5.3 Reverse Transcriptase PCR

To further our understanding of the set of ten genes it was decided to analyse their spatial expression and determine if the genes are expressed where it is thought the alkaloids are predominantly synthesised.

To study spatial expression, reverse transcriptase PCR was used to compare the gene's expression between the *top1* mutant and the control in leaf, root and stem tissue. This method was used in preference to qRT-PCR, as it was not possible to find a control gene for use in standardisation that remained unchanged across all three tissues types used. Primers designed to several *P. somniferum* genes were tried as qRT-PCR controls such as codeinone reductase, cyp80, BBE, TYDC, and the 3 different Major Latex Proteins all of which were deemed to variable across the three tissue types and between the two genotypes to be used as successful controls. The study was undertaken on root, stem and leaf mRNA as this material is easily isolated and is thought to be the main areas of synthesis of alkaloids. The expression of these genes in latex would also be important information in determining the relationship of the expression of these genes and the accumulation of the morphinan alkaloids in the latex, however RNA extractions from the latex was difficult and only yielded degraded RNA.

The analysis of message levels was undertaken in leaf, root and stem tissue taken from the mutant *top1* and the control. RNA was extracted from new plants grown

under the same conditions as the plants used for the microarray analysis (see sub Chapters 2.2 and 2.3) with the exception being material collected for the stem which was taken 4 days post petal drop (see Figure 1.7). Primer sets from the qRT-PCR were used and these worked for all of the ten genes except for two (52F2 and 91B6) that always gave multiple bands despite trying four different sets of primers for each gene and high annealing temperatures (62-65°C) for increased stringency. This suggests these two genes may belong to multi-gene families with different sized mRNA populations, or different RNA products, but with very similar sequence homology to the cDNA that was isolated from the microarray analysis.

The results of these experiments verified the qRT-PCR results for the leaf samples, however they also showed some differences between the expressions of these genes in the mutant *top1* compared to the control over the other tissue types (Figure 5.1). Not surprisingly the primers for 116D7 could not detect any message in the roots of either the mutant or the wild-type plant as this encodes a gene with a substantial hit to a chlorophyll a/b-binding protein. Expression of the other seven genes could be detected in stem and root tissue types.

The Rt-PCR experiments also give some information as to the level of expression of these genes across the various tissue types. The genomic controls yielded larger PCR products using the same primer combinations for several of the genes (91G3, 27D12, 157C4 and 111C7) suggesting there is an intron or introns present between the sequence/s that the specific primer combinations yielded. The size difference between the RNA amplified products and the genomic material gives confidence in there being no contamination of genomic DNA in the RNA extractions.

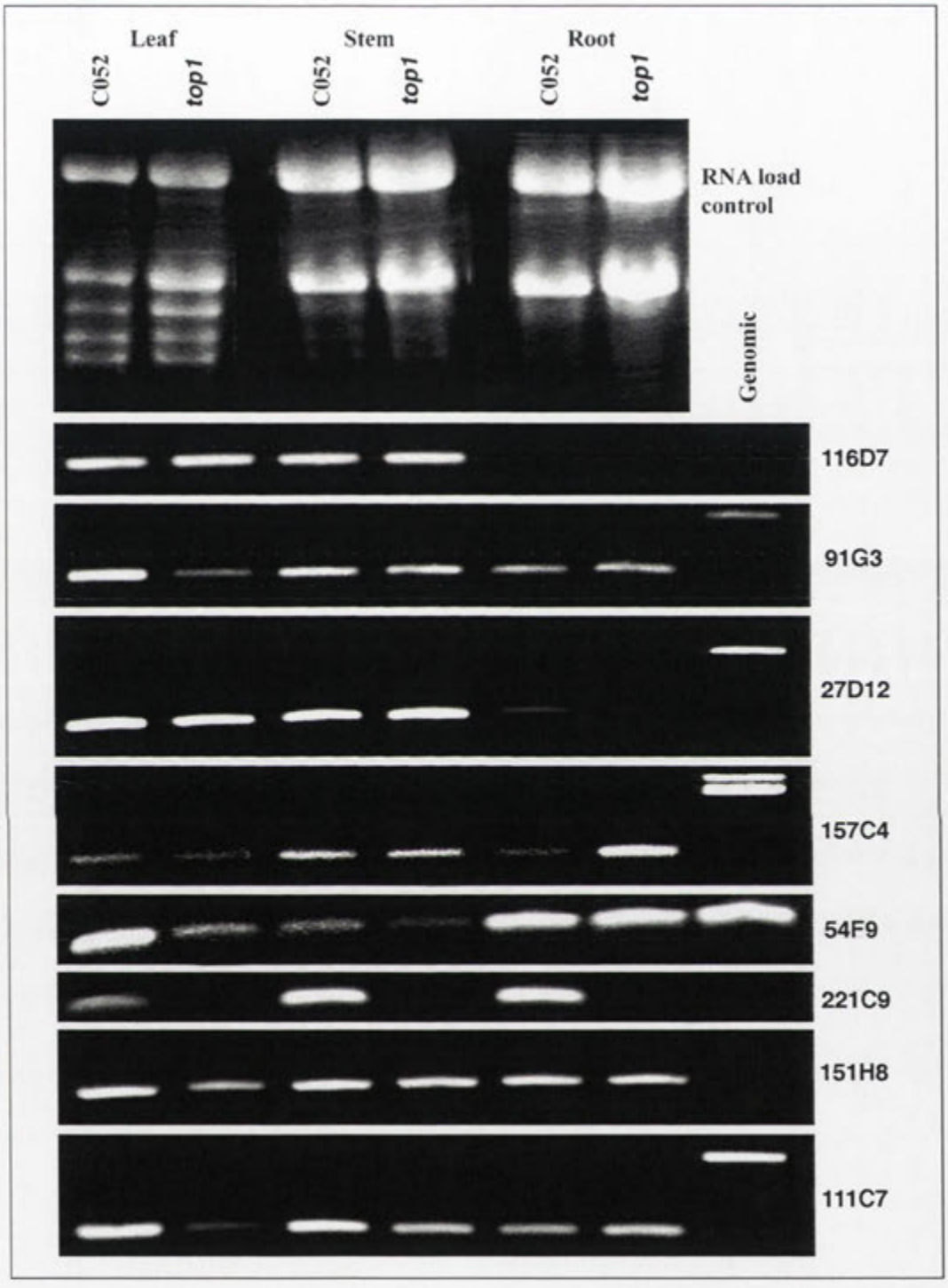


Figure 5.1: Reverse transcriptase PCR results of the eight differentially regulated genes from the *top1* mutant across leaf, stem and root transcript. The genomic control is also shown on the far right. This was obtained using C052 genomic DNA and the same primers used for the qRT-PCR reaction.

Table 5.1: Relative expression of clones found to be down regulated between C052 and *top1*.

Clone	Description	E-value	Array Ratio	qRT PCR Ratio[a]	RT-PCR Ratios					
					Leaf		Stem		Root	
					C052	*top1*	C052	*top1*	C052	*top1*
116D7	*Arabidopsis thaliana* psi type ii chlorophyll a/b-binding protein, putative.	6E-95	7.9	2.9	+++	+++	+++	+++	-	-
91G3*	*Arabidopsis thaliana* signal recognition particle receptor-like protein (at4g30600).	1E-153	5.1	2.4	+++	+	+++	++	+	+
27D12	No Significant Hits	-	4.7	7.5	+++	+++	+++	+++	+	-
157C4	*Arabidopsis thaliana* phosphoenolpyruvate carboxykinase [atp] (ec 4.1.1.49)	1E-164	4.4	3.6	++	+	++	++	+	+++
91B6	No Significant Hits	-	4.4	1.5			na			
54F9	No Significant Hits	-	3.9	6.7	+++	++	++	+	+++	+++
52F2*	*Arabidopsis thaliana* potential phospholipid-transporting ATPase 9 (ec 3.6.3.1)	0	3.6	na			na			

221C9	*Securigera parviflora* aspartate aminotransferase.	2E-54	3.4	69	+	-	+++	-	+++	-
151H8	No Significant Hits	-	2.9	1.7	+++	+	+++	++	++	+
111C7	*Arabidopsis thaliana* apospory-associated protein c; aldose 1-epimerase.	4E-52	2.7	3.2	+++	+	+++	+	+	++

From previous page Table 5.1: The complete set of non-redundant differentially expressed genes in the *top1* mutant including identified genes and unidentified genes, the latter based on E values greater than E^{-50}. The table lists: the clone identification (Clone Id); number of redundant genes (Number of Copies) and the correlating expression ratio as determined by qRT-PCR is also given as the control expression over the mutant genes expression. All but one of the genes could be checked by qRT-PCR. Only two genes showed a relative expression ratio less than the cut-off value of 2. **a**. The calculated CT ratio of the expression of the gene in the control over the mutant *top1* adjusted for RNA loading by comparison to the control expression.

* 91G3 represents 4 identical clones; 52F2 represents 2 similar clones.

na, results were unobtainable due to excess copies of genes..

RT-PCR ratios are given as + or – values depending upon the amount of product present see Figure 5.1. The more pluses the more product that could be visualised.

Interestingly 3 genes, 54F9, 151H8 and 221C9 were differentially under expressed across all of the tissue types tested in the mutant *top1* compared to the control tissue samples. 221C9, which has homology to an aspartate aminotransferase, also had a large large decrease in expression in the mutant further authenticating the qRT-PCR and microarray data showing. The other two genes 54F9 and151H8 have no substantial homology to any sequences in the NCBI database. 151H8 had a lower qRT-PCR value than the cut-off of 2 suggesting that the level of expression does not change substantially between the mutant *top1* and the control C052.

Two genes, 157C4 and 111C7, showed a large increase in expression in the *top1* mutant root over the control root expression (Figure 5.1 and Table 5.1) which has much lower levels of expression of either gene in the stem or leaf material. It is difficult to explain this situation except for the possibility of a feedback system. This could be explained by the mutation occurring in a positive regulator of expression of either or both the genes which is specifically expressed in the upper part of the plant (stem and leaf) and the plant is compensating by increasing transcript production in the roots with a different regulator which has been unaffected by the mutation.

Each of the genes, from the set of ten, with homology to known genes in the database will now be discussed as to their possible involvement in alkaloid biosynthesis.

5.4 Chlorophyll a/b Binding Protein 116D7

The cDNA (116D7) that has homology to a chlorophyll a/b binding protein seems to be differentially expressed in a number of mutations (Table 4.7) including the main focus of this thesis, the *top1* mutant. However this gene is not differentially under-expressed in the reticuline-accumulating mutant. These results suggest that the altered expression of this gene is a pleiotropic affect when the alkaloid pathway is perturbed in a number of ways. It is also conceivable that this chlorophyll a/b gene is affected due to possible pleiotropic changes in energy supply or that this gene is in

competition with another gene/s associated with the pathway. The gene 116D7 is only expressed in green tissue as confirmed by reverse transcriptase PCR expression data (Fig. 5.1 and Table 5.1) showed that it is only expressed in green tissue consistent with its presumed association with the photosynthetic machinery and its being under similar expression controls.

The chlorophyll a/b gene that 116D7 is closest to is that of the light-harvesting chlorophyll a/b binding protein Lhca2.1. The Lhca2 gene, to which Lhca2.1 is most similar, encodes one of the polypeptides of the light-harvesting complex of photosystem I. Lhca2.1 is thought to be too lowly expressed in *Arabidopsis* to have a role in the light-harvesting antennae and its function is still to be determined (Jansson 1999). Figure 5.2 illustrates the relatedness of the translated *P. somniferum* cDNA 116D7 to the *Arabidopsis* Lhca2.1 protein and to the rest of the *Arabidopsis* Lhca gene family. Another translated *P. somniferum* cDNA, 209F1, differentially expressed from the reticuline-accumulating mutants (E40/41) microarray, is also compared to the *Arabidopsis* Lhca proteins to show how divergent the *P. somniferum* proteins are form the *Arabidopsis* Lhca proteins.

A search of the literature did not reveal any direct links between photosynthesis and alkaloid production. The only possible link that can be attributed at this time would be the correlation that alkaloid genes and hence alkaloid synthesis (Facchini and Park, 2003) is increased due to stresses on the plant such as fungal attack and that photosynthesis is known to be down regulated in similar situations (Jung *et al.*, 2003). While both the alkaloid and photosynthetic pathways require nitrogen it is possible that the increase in alkaloids and nitrogen is offset by a reduction in photosynthesis. However the level of photosynthetic activity in molar terms far exceeds the level of alkaloid production in the plant making this simplistic link an unlikely explanation.

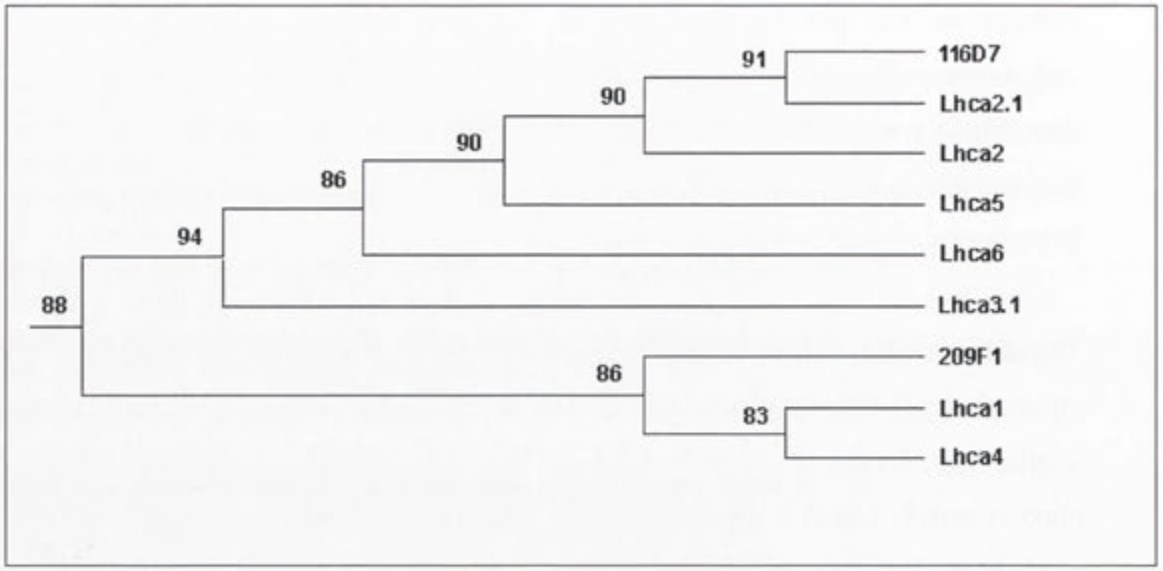


Figure 5.2: Tree illustrating the relatedness of seven *Arabidopsis thaliana* chlorophyll light harvesting complex I (Lhca) protein sequences (Lhca1, Lhca2, Lhca2i, Lhca3i, Lhca4, Lhca5 and Lhca6). The sequence of 116D7 showed closest homology to the Lhca2i gene. The translated sequence of the 209F1 clone, shown here to reveal homology between the *Papaver* and *Arabidopsis* Lhca genes from the reticuline accumulating experiment (subchapter 4.4.2) was closest to the Lhca1 gene. (Adapted from Jansson 1999) Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 3.0 (Kumar *et al.*, 2004). Numbers on branches are ClustalW scores for the two closest aligned protein sequences from each branch.

Arabidopsis Lhca1- NCBI Accesion Number M85150.

Arabidopsis Lhca2- NCBI Accesion Number AT3G61470.1.

Arabidopsis Lhca2.1- NCBI Accesion Number AT1G19150.1.

Arabidopsis Lhca3.1- NCBI Accesion Number U01103.

Arabidopsis Lhca4- NCBI Accesion Number X71878.

Arabidopsis Lhca5- NCBI Accesion Number AF134121.

Arabidopsis Lhca6- NCBI Accesion Number AF134130.

Papaver somniferum- 209F1 and 116D7 from this thesis.

5.5 Aminotransferase like cDNA 221C9

Some attention has been given to 221C9, the amino transferase-like gene, because of the large difference in expression levels in the qRT-PCR results and all three tissue types in the RT-PCR data. The blast results revealed high homology to a gene with a known and characterised aminotransferase activity. The sequence has no similarity to cytochrome P450 enzymes and it seems very unlikely an aminotransferase could have been modified to catalyse the demethylation reaction of thebaine to neopinone.

GGTTTGCAAATCATCGGATGTGGCTAGCAGAGTTGAAAGTCAGCTGAAACTTGTTATCAG
GCCTATGTATTCAAATCCTCCCGTGCATGGCCCATCTATTGTGGCTACGGTCCTGAAAGA
TAGAGACTTGTACAATGAGTGGACTCTAGAGCTGAAAGCTATGGCTGACAGAATAATAA
GTATGCGTCAGCAACTATTTGATGCATTACGTGCAAGAGGTACACCAGGTGATTGGAGTC
ACATTATTAAACAAATTGGGATGTTTACTTTCACTGGTTTGAATAAAGAGCAAGTTGCTT
TCATGACTAGAGAATATCACATCTACATGACTTCTGATGGG>taaatgtctctatgcttacttgattacaattgc
aaggtcaatgctccttgattacaattgcaaggtcaatgctcaggatcaatatgggctattgttcagaatcccatgtcattttcatgggatatcattataagt
ctcctgtgctgcttgcatttcgtgtagatcctggaggaaacctatccaatattacagacaccggtttcctttaatttttttgttccttttctgaaacatgaattt
gtaaaggaattctcttgtcatgcaggagaattagcatggcgggttta<AGTTCAAGGACGGTGCCTCACCTTGCGGATG
CTATTAATGCTGCTGTGACCCGTATGGGTTGAGCAAAACTTGGTATCAAATTGTGTAGGA
TGCCTTTCGGAGGTTTGCTTTTTAATTTTATTTTTATTCTCTGTTGCAATAAAAAACAGCTC
AACAATTAGTGATGTAAGCTGCAGTTCCTAATTCGATTGAATCTCACGTGAGATATTTGC
CCCATCTTTTATAGAAATGAAACTTACTATGTGAGTTTTGTTTTTCCTTTTCTCTGGAAGTC
GAGGATTTTAATAGAGTGACTAACAAATGGACTCTTGTTGGAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Figure 5.3: The nucleotide sequence for the *P. somniferum* cDNA 221C9. The blue sequence, enclosed by the symbols < >, is the region of the cDNA that does not align with any of the ESTs from the NCBI database.

Poppy cDNA 221C9 was found to be differentially under-expressed in microarray experiments comparing a number of other mutants to the controls; this was also true for the entire set of 10 *top1* differentially expressed genes (Table 4.7). However, the set of ten genes from the *top1* microarray experiments were not differentially expressed in either the Reticuline mutant or in the hairpin RNAi codeinone reductase

GGTTTGCAAATCATCGGATGTGGCTAGCAGAGTTGAAAGTCAGCTGAAACTTGTTATCAG
GCCTATGTATTCAAATCCTCCCGTGCATGGCCCATCTATTGTGGCTACGGTCCTGAAAGA
TAGAGACTTGTACAATGAGTGGACTCTAGAGCTGAAAGCTATGGCTGACAGAATAATAA
GTATGCGTCAGCAACTATTTGATGCATTACGTGCAAGAGGTACACCAGGTGATTGGAGTC
ACATTATTAAACAAATTGGGATGTTTACTTTCACTGGTTTGAATAAAGAGCAAGTTGCTT
TCATGACTAGAGAATATCACATCTACATGACTTCTGATGGGAGTTCAAGGACGGTGCCTC
ACCTTGCGGATGCTATTAATGCTGCTGTGACCCGTATGGGTTGAGCAAAACTTGGTATCA
AATTGTGTAGGATGCCTTTCGGAGGTTTGCTTTTTAATTTTATTTTTATTCTCTGTTGCAAT
AAAAAACAGCTCAACAATTAGTGATGTAAGCTGCAGTTCCTAATTCGATTGAATCTCACG
TGAGATATTTGCCCCATCTTTTATAGAAATGAAACTTACTATGTGAGTTTTGTTTTTCCTTT
TCTCTGGAAGTCGAGGATTTTAATAGAGTGACTAACAAATGGACTCTTGTTGGAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Figure 5.4: The nucleotide sequence for Del_221C9. This is the nucleotide sequence from Figure 5.1 without the nucleotide sequence in blue that does not align to any of the ESTs from the NCBI database. The stop codon TGA is represented in red.

microarray experiments (Table 4.7). This suggests that the 221C9 cDNA (and the other nine *top1* differential genes) is associated with perturbations to the pathway in the later morphinan branch but not to perturbations resulting in accumulation of the earlier reticuline intermediate.

Blasting the poppy cDNA 221C9 against the NCBI EST database shows that the 221C9 cDNA has an insertion that is not present in other ESTs that align to this gene. It seems likely that it is an unspliced intron present in cDNA 221C9. This insertion is shown in blue in the sequence in the Figure 5.3. Using this information a sequence was constructed without the novel inserted sequence (blue insertion) from Figure 5.3 and called Del_221C9 (Figure 5.4). The new nucleotide sequence translates into a longer protein by 20 amino acids (Figure 5.5), as the stop codon that occurred in the intron is no longer present to terminate translation.

a). Del_221C9 Translated - Longest ORF [Frame 2] 112 amino acids

MYSNPPVHGPSIVATVLKDRDLYNEWTLELKAMADRIISMRQQLFDALRARGTPGDWSHIIK
QIGMFTFTGLNKEQVAFMTREYHIYMTSDGSSRTVPHLADAINAAVTRMG

b). Poppy 221C9 Translated - Longest ORF [Frame 2] 92 amino acids

MYSNPPVHGPSIVATVLKDRDLYNEWTLELKAMADRIISMRQQLFDALRARGTPGDWSHIIK
QIGMFTFTGLNKEQVAFMTREYHIYMTSDG

Figure 5.5: The translated open reading frames of the cDNA 221C9. a). Is the longest translated open reading frame of the Del_221C9 nucleotide sequence which shows the extra 20 amino acids in red that are not present in the translated 221C9 sequence (b) that still contains the extra nucleotide sequence not found in the aligning ESTs in the database.

Alignment of the poppy cDNA 221C9 with homologues in the database revealed that the poppy clone contained an unspliced intron at the 3' end (Figure 5.6). New primers (Figure 5.6) were designed to various regions of the cDNA for RT-PCR to check the qRT-PCR data. The results showed that the transcript intron seems to splice in the mutant but not in the control.

The 5' forward primer of the original primer set (AGM180) used for the qRT-PCR was designed to the unspliced region in the 221C9 cDNA. The original cDNA clone 221C9 was taken from the cDNA library generated from the control C052 plant and contains the unspliced intron. The mutant contains spliced versions of the transcript and it is the difference in the splicing of the intron that has caused this gene to appear to be differentially under-expressed; perhaps the presence of the intron in the transcript renders it more stable to degradation and therefore the microarrays will register the difference in the size of the transcript pool. It is possible that 221C9 transcript in the control translates to a protein with no function compared to the

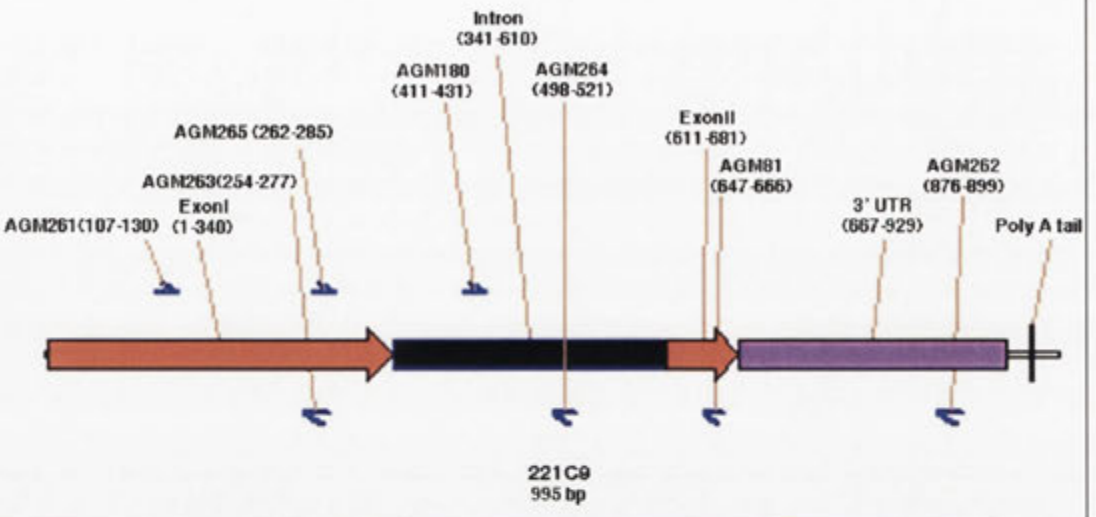


Figure 5.6: A diagrammatic representation of the cDNA 221C9, which shows the unspliced intron found in the control phenotype. The diagram also shows the 3' untranslated region (3' UTR), the two coding regions in red (Exon) and the Poly A tail of the cDNA. The blue arrows illustrate where each of the primers were designed to in the nucleotide sequence with the arrow showing the direction in which that primer amplified in for PCR. The primers AGM180 (5' forward) and AGM 181 (3' reverse) were the original primers used for the qRT-PCR experiments that gave a 69-fold difference in expression levels between the mutant *top1* and the control C052. AGM 180 is to the intron itself. Figure 5.7 shows the results obtained from the different combinations of each of the primers. The blue arrow representing each primer represents the DNA strand to which it binds and the arrow signifies the direction of amplification.

spliced version in *top1* plants. The non-spliced transcript may also translate to a protein with altered function. The altered splicing may not just be specific to this gene it may be occurring in other genes, one of which could be responsible for the alkaloid defect. This could also explain why so many protein changes are observed but not transcriptional changes.

Sequencing comparisons of the gene in genomic DNA of the mutant and control was undertaken, using primers to the 5' UTR region of the poppy aminotransferase

221C9, to ascertain if there was an internal sequence mutation causing the splice-variant of 221C9 in the mutant. This experiment was unsuccessful; sequence from both genomic samples was of good quality for a few hundred base pairs from the 5' end but then suddenly became undecipherable, suggesting that there are several polymorphic copies of the gene. Only by cloning each of the genes could this be resolved further.

More primers were designed to try to ascertain if a mutation lies between the mutant and control genes of 221C9. RT-PCR was conducted with various combinations of primers on RNA isolated from both the mutant *top1* and control to see if there were differences in the size of the products (Table 5.2) resulting from EMS-induced deletion. The resulting RT-PCR reactions using these primer combinations revealed some interesting results (Figure 5.7). Equal concentrations of RNA were used for the RT-PCR reactions and each reaction was run for 30 cycles to determine a difference in transcript levels. The first primer combination (Figure 5.7 number 1) of AGM180 and AGM181 was used as a control as these were the primers used for the verification of the of the microarray data. This primer combination amplifies part of the intron and the second exon (Labelled ExonII in the Figure 5.6). The expected amplified bands are of 255 bp if the intron is present and no band is expected in the RT-PCR reaction of the RNA if the intron has been spliced. The expected non-excised band of 255 bp was seen in both the control and the *top1* mutant but to a much lesser extent. The genomic DNA showed that the 255 bp DNA fragment is present in both the mutant and control genomes. Another control reaction was with primers designed to the first open reading frame only (primers AGM261 and AGM263). As expected both the control and *top1* RT-PCR reactions produced a 170 bp fragment. The genomic PCR revealed that there was more than just the 170 bp fragment present in both the control and mutant genotypes. It is possible that there is more than one gene coding for the aminotransferase 221C9 with different sized introns that are excised and not seen in the RT-PCR reactions. To test this theory the fragments would need to be cloned and sequenced and this sequence data compared

Table 5.2: The primer combinations used to test for intron splicing of 221C9.

#	Primer 5'	Primer 3'	Region Amplified	Size expected + Intron (bp)	Size Expected - Intron (bp)
1	AGM180	AGM181	Intron to ExonII	255	0
2	AGM180	AGM262	Intron to 3' UTR	488	0
3	AGM180	AGM264	Internal to Intron	110	0
4	AGM261	AGM181	5'ExonI to ExonII	559	290
5	AGM261	AGM262	5' ExonI to 3' UTR	792	523
6	AGM261	AGM263	5' ExonI	170	170
7	AGM261	AGM264	5' ExonI to Intron	414	0
8	AGM265	AGM181	5' ExonI to ExonII	404	135

Primer combinations used to test for differences between the *top1* and control aminotransferase 221C9 gene. Each primer pair was tested on RNA and genomic DNA extracted from the *top1* mutant and control plant. See Figure 5.6 for the position of each of the primers. The number (#) of the primer combination is given for reference for the resulting RT-PCR reactions (Figure 5.6). For primer sequences see Table 2.1.

to the aminotransferase 221C9 sequence. The second reaction undertaken (Figure 5.7- 2) of the primers AGM180 and AGM262 were internal to the intron but this time to the 3' untranslated region to detect if there are some different translational modifications. As can be seen in figure 5.7 there are three different bands in the control RT-PCR lane ranging in size from the expected 488 bp to about 250 bp and this range is also seen in the genomic samples for both the control and mutant phenotype suggesting that there are several copies of the same gene with different length regions between this primer pair or different splice sites 3', especially as it means that the intron seq needs to be conserved. The *top1* RNA showed a reduced level of amplification of the same size band suggesting that there is a very low level of non-splicing of the intron taking place or another copy of the gene with the intron present and not spliced giving a lower level of transcript from which to amplify the product.

The third reaction was undertaken using primers AGM180 and AGM262, which are both internal to the intron. The resulting 110 bp fragment could be seen in the control and again to a lesser extent in the mutant. Another band at around 700 bp could be seen amplifying in the control RT-PCR and in both the control and mutant genomic DNA PCR reactions. It is possible this is a primer artefact, however it is unusual that the band could not be seen in the *top1* mutant RT-PCR reaction suggesting that there may be another amplification product specific to this primer pair that is also downregulated in the mutant. Further work needs to be done on this however by sequencing the product and comparing the sequence to the 221C9 sequence.

The next two sets of primers dealt with the 5' end of the 221C9 sequence. Firstly two primers were designed to the 5' end of the first open reading frame (ExonI in Figure 5.6). The first of the primers AGM261 was to a region of high homology across the aminotransferases in the database and the second 5' primer AGM262 was to a less conserved region of the first open reading frame of the compared sequences of the aminotransferases in the database. Using the first primer to the more conserved region AGM261 no result could be obtained for reaction four with the 3' primer AGM181, which was internal to the second open reading frame (ExonII in Figure 5.6). The size fragment expected from this reaction was 559 bp if the intron was not excised or 290 bp if it was. Both of these bands could only be seen in the *top1* genomic sample but not in the control or any of the RT-PCR reactions. Several other distinct bands could be seen in the *top1* genomic sample which again were not reciprocated in the control genomic or either of the RT-PCR reactions. It is possible that this was a bad primer combination. This reaction was repeated in reaction 8 using the primer AGM262 to the less conserved region of the first open reading frame (ExonI in Figure 5.6) and the same 3' primer as in reaction four. This time the expected result for the non-splicing event (404 bp fragment) could be seen in the control along with some splicing taking place as well, the 135 bp fragment. The *top1* mutant RT-PCR produced very little product of the unspliced form (404 bp in size)

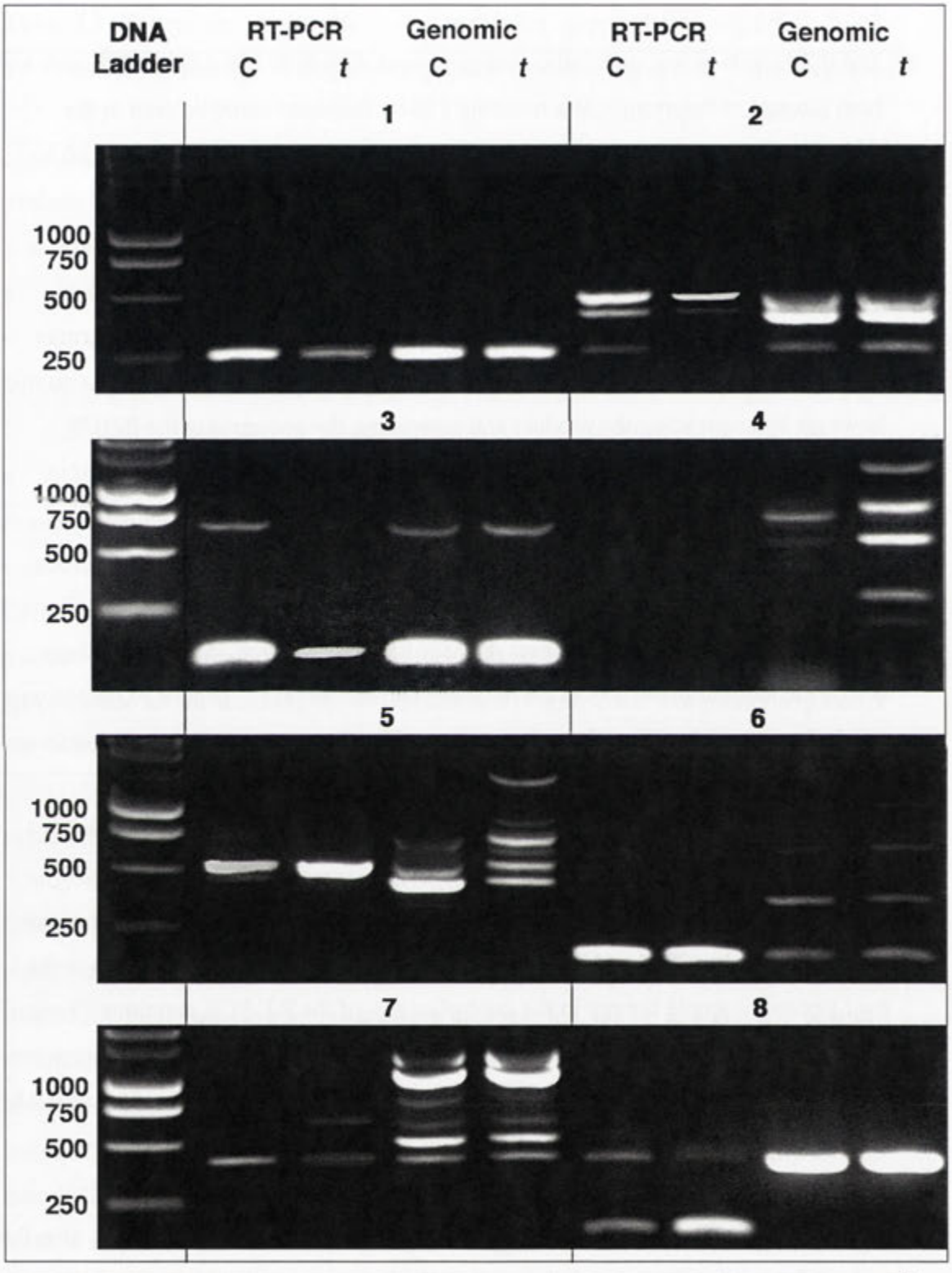


Figure 5.7: RT-PCR results of the 8 primer combinations used to test for any deletion mutations between the *top1* splicing-variant of the poppy aminotransferase 221C9 and the non-splicing control version. The *top1* genomic DNA or RNA

sample used for the RT-PCR or PCR respectively is represented by (*t*) and the control sample is represented by (C). RNA was used for the RT-PCR and genomic DNA for a PCR of the respective control or *top1* samples. The numbers 1-8 represent the primer combinations used in each of the RT-PCR reactions. See Table 5.2 for the list of primer combinations and the expected RT-PCR product sizes. The Promega 1 Kb marker is run on the left as a DNA ladder for approximate size determination.

and the major product was of the spliced version (135 bp). The genomic DNA amplified a single band of 404 bp showing that both genomes contain the unspliced intron.

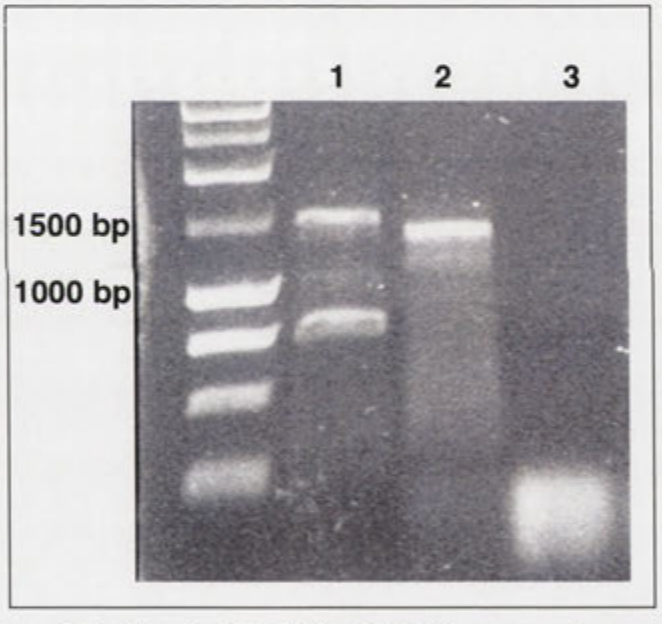


Figure 5.8: Cloning of the 5'end of 221C9 cDNA (Poppy aminotransferase) Using the 5' primer SP6 to the vector containing the cDNA and 3' primer (1) AGM181 (647-666 bp) Extra 900 bp (not cloned), (2) AGM263 (254-277 bp) Extra 1200 bp (cloned) and (3) AGM264 (no product). See Figure 5.5 for primer positions. The sizes are listed on the gel for the 1000 bp and 1500 bp fragments of the DNA ladder (Promega 1 Kb DNA ladder) (left lane).

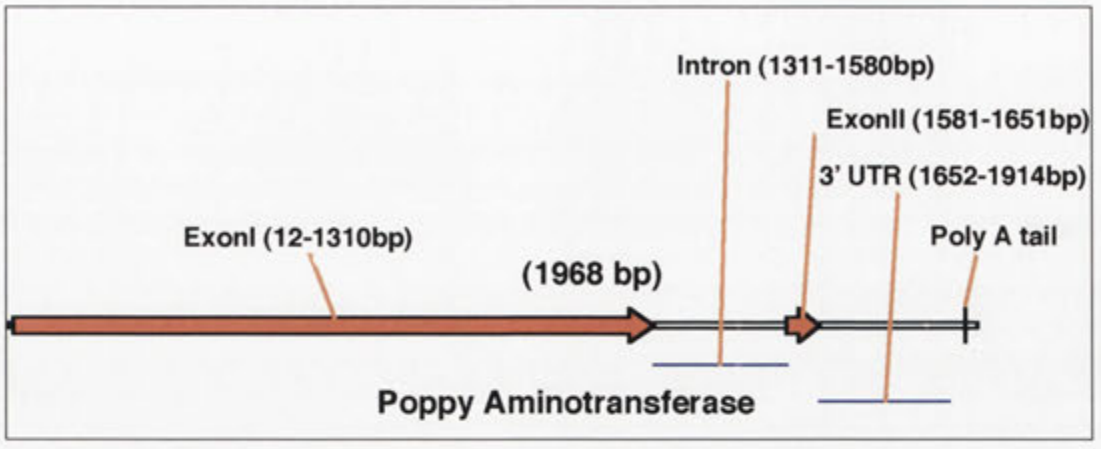


Figure 5.9: The poppy aminotransferase gene. The diagram represents the cDNA 221C9 from the *top1* microarray experiments along with an extra 1003 bp at the 5' end of the gene found using the SP6 vector primer at the 5' end of the inserted cDNA sequence and the internal 3' reverse primer AGM263.

The last comparison that was made between the *top1* mutant and control 221C9 transcripts was with a 5' primer (AGM261) to the first open reading frame (ExonI in Figure 5.6) and to either a 3' primer (AGM264) internal to the intron (Reaction 7 in Figure 5.7) or a 3' primer (AGM262) to the 3' untranslated region intron (Reaction 5 in Figure 5.7). The result for the primer set amplifying from within the intron (Reaction 7 Figure 5.7) showed a band running at a size of approximately 400 bp in both genotypes, which is the expected size for the product if the intron is not excised. The *top1* mutant also gave a band of the same size but the band was not as intense suggesting that there was less non-excised intron 221C9 transcript available for the RT-PCR reaction. The genomic PCR of both samples gave several bands including an approximate 400 bp fragment. Again this could be due to several copies of the aminotransferase with different sized introns being present or the aminotransferase genes could be close together on the same chromosome/s. The reaction that encompasses a fragment from the first open reading frame across the intron to the 3' untranslated region of 221C9 (Reaction 5 in Figure 5.7) gave some interesting results. The control RT-PCR reaction gave a band the same size (500 bp) as expected for the *top1* mutant without the intron present. If the intron was not excised then there should be a 792 bp fragment, which isn't seen. It is possible that the

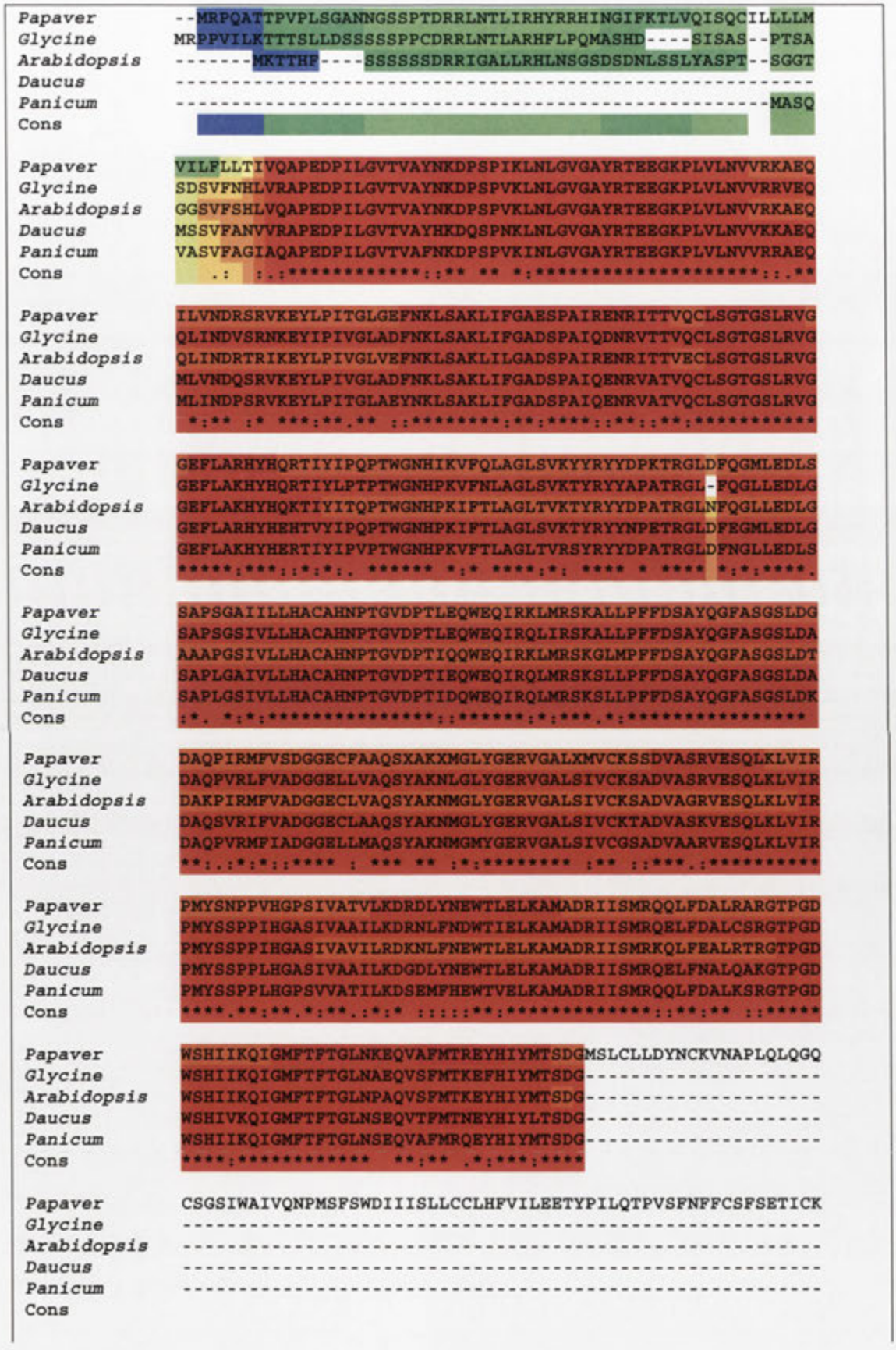

Papaver --MRPQATTPVPLSGANNGSSPTDRRLNTLIRHYRRHINGIFKTLVQISQCILLLLM
Glycine MRPPVILKTTTSLLDSSSSSPPCDRRLNTLARHFLPQMASHD----SISAS--PTSA
Arabidopsis -------MKTTHP----SSSSSSDRRIGALLRHLNSGSDSDNLSSLYASPT--SGGT
Daucus ----------
Panicum ----------MASQ
Cons
Papaver VILFLLTIVQAPEDPILGVTVAYNKDPSPIKLNLGVGAYRTEEGKPLVLNVVRKAEQ
Glycine SDSVFNHLVRAPEDPILGVTVAYNKDPSPVKLNLGVGAYRTEEGKPLVLNVVRRVEQ
Arabidopsis GGSVFSHLVQAPEDPILGVTVAYNKDPSPVKLNLGVGAYRTEEGKPLVLNVVRKAEQ
Daucus MSSVFANVVRAPEDPILGVTVAYHKDQSPNKLNLGVGAYRTEEGKPLVLNVVKKAEQ
Panicum VASVFAGIAQAPEDPILGVTVAFNKDPSPVKINLGVGAYRTEEGKPLVLNVVRRAEQ
Cons
Papaver ILVNDRSRVKEYLPITGLGEFNKLSAKLIFGAESPAIRENRITTVQCLSGTGSLRVG
Glycine QLINDVSRNKEYIPIVGLADFNKLSAKLIFGADSPAIQDNRVTTVQCLSGTGSLRVG
Arabidopsis QLINDRTRIKEYLPIVGLVEFNKLSAKLILGADSPAIRENRITTVECLSGTGSLRVG
Daucus MLVNDQSRVKEYLPIVGLADFNKLSAKLIFGADSPAIQENRVATVQCLSGTGSLRVG
Panicum MLINDPSRVKEYLPITGLAEYNKLSAKLIFGADSPAIQENRVATVQCLSGTGSLRVG
Cons
Papaver GEFLARHYHQRTIYIPQPTWGNHIKVPQLAGLSVKYYRYYDPKTRGLDFQGMLEDLS
Glycine GEFLAKHYHQRTIYLPTPTWGNHPKVFNLAGLSVKTYRYYAPATRGL-FQGLLEDLG
Arabidopsis GEFLAKHYHQKTIYITQPTWGNHPKIFTLAGLTVKTYRYYDPATRGLNFQGLLEDLG
Daucus GEFLARHYHEHTVYIPQPTWGNHPKIFTLAGLSVKTYRYYNPETRGLDFEGMLEDLG
Panicum GEFLAKHYHERTIYIPVPTWGNHPKVFTLAGLTVRSYRYYDPATRGLDFNGLLEDLS
Cons
Papaver SAPSGAIILLHACAHNPTGVDPTLEQWEQIRKLMRSKALLPFFDSAYQGFASGSLDG
Glycine SAPSGSIVLLHACAHNPTGVDPTLEQWEQIRQLIRSKALLPFFDSAYQGFASGSLDA
Arabidopsis AAAPGSIVLLHACAHNPTGVDPTIQQWEQIRKLMRSKGLMPFFDSAYQGFASGSLDT
Daucus SAPLGAIVLLHACAHNPTGVDPTIEQWEQIRQLMRSKSLLPFFDSAYQGFASGSLDA
Panicum SAPLGSIVLLHACAHNPTGVDPTIDQWEQIRQLMRSKSLLPFFDSAYQGFASGSLDK
Cons
Papaver DAQPIRMFVSDGGECFAAQSXAKXMGLYGERVGALXMVCKSSDVASRVESQLKLVIR
Glycine DAQPVRLFVADGGELLVAQSYAKNLGLYGERVGALSIVCKSADVASRVESQLKLVIR
Arabidopsis DAKPIRMFVADGGECLVAQSYAKNMGLYGERVGALSIVCKSADVAGRVESQLKLVIR
Daucus DAQSVRIFVADGGECLAAQSYAKNMGLYGERVGALSIVCKTADVASKVESQLKLVIR
Panicum DAQPVRMFIADGGELLMAQSYAKNMGMYGERVGALSIVCGSADVAARVESQLKLVIR
Cons
Papaver PMYSNPPVHGPSIVATVLKDRDLYNEWTLELKAMADRIISMRQQLFDALRARGTPGD
Glycine PMYSSPPIHGASIVAAILKDRNLFNDWTIELKAMADRIISMRQELFDALCSRGTPGD
Arabidopsis PMYSSPPIHGASIVAVILRDKNLFNEWTLELKAMADRIISMRKQLFEALRTRGTPGD
Daucus PMYSSPPLHGASIVAAILKDGDLYNEWTLELKAMADRIISMRQELFNALQAKGTPGD
Panicum PMYSSPPLHGPSVVATILKDSEMFHEWTVELKAMADRIISMRQQLFDALKSRGTPGD
Cons
Papaver WSHIIKQIGMFTFTGLNKEQVAFMTREYHIYMTSDGMSLCLLDYNCKVNAPLQLQGQ
Glycine WSHIIKQIGMFTFTGLNAEQVSFMTKEFHIYMTSDG----------
Arabidopsis WSHIIKQIGMFTFTGLNPAQVSFMTKEYHIYMTSDG----------
Daucus WSHIVKQIGMFTFTGLNSEQVTFMTNEYHIYLTSDG----------
Panicum WSHIIKQIGMFTFTGLNSEQVAFMRQEYHIYMTSDG----------
Cons
Papaver CSGSIWAIVQNPMSFSWDIIISLLCCLHFVILEETYPILQTPVSFNFFCSFSETICK
Glycine ----------
Arabidopsis ----------
Daucus ----------
Panicum ----------
Cons

Figure 5.10: Alignment of the Poppy aminotransferase gene 221C9, including the cloned 5' coding region, with other aminotransferases from the database. The colour code represents the quality of the alignment. BAD AVG GOOD

Panicum miliaceum (proso millet) aspartate transaminase (EC 2.6.1.1) AAT2 – proso millet. NCBI Accession number S53303.

Daucus carota (carrot) aspartate aminotransferase, cytoplasmic (Transaminase A). NCBI Accession number P28734.

Arabidopsis thaliana aspartate aminotransferase, chloroplast / transaminase A (ASP3) (YLS4). NCBI Accession number NP_196713.

Glycine max (soybean) aspartate transaminase (EC 2.6.1.1) AAT1 peroxisomal/ glyoxysomal precursor - soybean. NCBI Accession number T06136.

T-COFFEE, Version_1.41 was used for the amino acid alignments (Notredame *et al.*, 2000).

excised version of the cDNA was amplified in the PCR reaction preferentially over the larger fragment, or that there is another aminotransferase present that this primer combination can amplify but that previous combinations could not. The genomic PCR results using the same combination of primers gave different banding patterns between the control and the mutant. This again suggests that there are multiple copies of the gene with different sized introns or tandem repeats of the aminotransferase gene.

Using the reverse primers AGM181, AGM263 and AGM264 (See Figure 5.6) and the vector primer to the 5' end of the cloned cDNA insert it was possible to amplify the 5'end of the aminotransferase gene from the control C052 cDNA library (Figures 5.8 and 5.9). One fragment was cloned from the primer combination 5' SP6 and 3' AGM263 which was approximately 1400bp in length. Sequencing and BLAST x analysis showed that this fragment was longer than the coding region of

aminotransferases in the database and therefore is giving some of the 5' UTR region. Alignment of this sequence with the poppy cDNA 221C9 showed that the two sequences overlapped and the subsequent BLAST x analysis of the full contig revealed that the gene was in fact an aminotransferase. It aligned best to the *Daucus* (carrot) aspartate aminotransferase (Figure 5.10). There are 4 distinct aminotransferases in *Arabidopsis* with 3 different localisations; AAT2 is cytosolic, AAT1 is mitochondrial and AAT3 is chloroplastic. The Poppy aminotransferase falls between the plastid, mitochondrial and cytosolic AAT types and not to any particular one type. Without further analysis such as localisation studies it is difficult to classify the cDNA 221C9.

Several bands could be seen in these reactions that were in the control but not in the *top1* mutant (for example the 750 bp product seen in reaction 3) suggesting that there may be homologous regions outside the aminotransferase coding region and that the primer pairs are amplifying a larger construct. The first obvious step would be to undertake Southern analysis of the aminotransferase gene in the poppy to see if there are more copies than in *A. thaliana* and any possible gene duplication events.

The benzylisoquinoline alkaloid pathway requires this type of enzyme, still yet to be cloned, to perform the transaminase reaction on tyrosine to form 4–hydroxyphenylpyruvate. However 221C9 has highest homology to an aspartate aminotransferase rather than a tyrosine aminotransferase. As explained earlier, it is not feasible that an aminotransferase type enzyme could be directly involved in a demethylation reaction. Because of the splicing difference that has been uncovered in 221C9, the most likely link between 221C9 and the *top1* mutation is that they are caused by two unrelated mutations in the *top1* germplasm. The F3 segregation data in chapter 5.11 reveals that this gene does segregate with the *top1* phenotype adding further evidence to the possibility of the mutation in the 221C9 gene being at least linked with the *top1* phenotype.

5.6 Potential phospholipid-transporting ATPase cDNA 52F2

The transporter-like proteins represented by cDNA 52F2 are differentially under-expressed in a number of the mutants of morphinan alkaloid biosynthesis examined in this thesis.

Morphinan alkaloids accumulate in latex vesicles within poppy laticifers and it is now generally accepted in the literature that laticifer ultrastructure and differentiation is related to morphinan accumulation. Several sequential enzymatic steps in alkaloid biosynthesis have been shown to occur in distinct cell types indicating that intercellular transport of intermediates occurs. Most of the information that is known comes from expression patterns and localisation studies performed on several of the alkaloids. Measurements of the alkaloids; morphine, codeine, thebaine, papaverine, and noscapine in latex showed that more than 90% of the alkaloid content was found in two different types of vacuoles (Fairbairn *et al.*, 1974; Fairbairn and Steele, 1981; Homeyer and Roberts, 1984; Roberts *et al.*, 1983; Roberts, 1987). One vacuole type was designated 900x g and the other 1100x g due to their sedimentation properties. Of the two types of vacuoles the 900x g fraction of vacuoles seemed to contain the higher proportion of alkaloids and seemed to accumulate all of the aforementioned alkaloids measured except thebaine in the three weeks after petal opening (Pham and Roberts, 1991).

Two models for alkaloid transport in *P. somniferum* have been proposed:

1. The first, diffusion and ion trap mechanism, assumes lipophilic alkaloids can diffuse across the membranes and are trapped in the acidic vacuolar compartment by protonation and salt formation (Neumann *et al.*, 1983; Matile 1984; reviewed in Roberts *et al.*, 1991). It has been shown in *Chelidonium majus* that chelidonic acid readily complexes alkaloids (Hauser and Wink, 1990) and possibly explains alkaloid accumulation against a concentration gradient as observed for sanguinarine and various other

lipophilic alkaloids into latex vacuoles (Matile, 1976; Hauser and Wink, 1990). In *P. somniferum* meconic acid is the major complexing agent found in the latex in high concentrations (Fairbairn and Williamson, 1978). High levels of sulphate and chloride were also found consistent with some form of ion trap mechanism for the major alkaloids of opium poppies (Pham and Roberts, 1991). There is however no direct evidence for an "ion trap" being responsible for the accumulation of morphinan alkaloids in latex vacuoles. Indeed morphine accumulates in opium poppy preferentially over several more basic alkaloids; this is inconsistent with ion trapping. An addition of ATP caused morphine uptake into the vacuoles to increase, however ATP inhibitors did not inhibit uptake (Roberts *et al.*, 1991). These authors examined the uptake of various alkaloids and found no clear correlation with pK or lipophilicity. They proposed some sort of channel uptake mechanism with specificities to particular alkaloid molecules.

It is possible that the vacuoles in the latex may undertake this role. It is thought that the vacuoles of *P. somniferum* latex may be specialist vacuoles due to their size, resistance to external measures designed to reduce the Δ pH and capacity to sequester millimolar quantities of alkaloids (Roberts *et al.*, 1991). Further to this the rapid and specific uptake of the major opium alkaloids has no absolute requirement for MgATP, is independent of temperature and does not show saturation kinetics (Homeyer and Roberts, 1984).

2. The second model is based upon specific membrane bound alkaloid transporters. Other alkaloids have been shown to pass the tonoplast via a carrier mechanism. This is evident for *S*-reticuline, a charged species under cytosolic pH, and this mechanism appears to be stereo-selective for the *S*- rather than the *R*- enantiomer (Deus-Neumann and Zenk, 1986). This discrimination for the naturally occurring *S*-enantiomers has also been shown for the alkaloid *S*-scoulerine (Deus-Neumann and Zenk, 1986). Uptake is

thought to occur by a H^+ antiport mechanism such as for lupanine (Mende and Wink, 1987), *S*-reticuline, *S*-scoulerine (Deus-Neumann and Zenk, 1986), and 1-(malonyl amino)-cyclopropane-1-carboxylic acid (Bouzayen *et al.*, 1989). H^+- ATPase is a major vacuolar protein (Leigh and Walker, 1980; Sze, 1985; Bremberger *et al.*, 1988) and along with pyrophosphatase (Leigh and Walker, 1980; Rea and Sanders, 1987; Bremberger *et al.*, 1988) is responsible for tonoplast H^+ concentration being greater than the cytoplasm (Thom and Komor, 1984; Sze, 1985; Rea and Sanders, 1987; Heidrich *et al.*, 1989; Taiz, 1992). This proton gradient has been shown to be utilized for a substrate- proton antiport transport mechanism (Lüttge *et al.*, 1981; Hager and Hermsdorf, 1981; Thom and Komor, 1984, 1985; Briskin *et al.*, 1985; Blumwald and Poole, 1985 a,b; Blumwald, 1987; Blackford *et al.*, 1990; Getz, 1991).

The phospholipid-transporting ATPases or flippases are known to transport lipids across membranes. They function by the substrate first partitioning into the cytoplasmic leaflet of the lipid bilayer from which it gains access to the transporter. The substrate is then 'flipped' by the transporter in an ATP-dependent manner to the opposite leaflet of the membrane and thence to the adjacent aqueous medium. MDR (*multidrug resistance*) transporters are also known to act in this way (van Helvoort *et al.*, 1996) and the human MDR1 P-glycoprotein is a flippase of broad specificity that catalyses the transfer of phospholipids across the plasma membrane. Many molecules have been shown to be transported by flippases including the alkaloid vinblastine (Tamai and Safa 1991). While some flippases have been shown to be highly specific (van Helvoort *et al.*, 1996), further detailed analysis of the flippases found in this study will be required to determine their specificity. These cDNAs could be representative of the ATP requiring transporters needed to shuttle alkaloids across the various compartments where they accumulate.

In order to test if these flippases are involved in alkaloid shuttling or some control process in morphinan synthesis it would be beneficial to use either antisense or

RNAi gene silencing to respectively partially or totally knockout this gene. Antisense technology may be more useful if this transporter is not specific in the transport of morphinans or associated compounds, as it would enable crucial cellular processes to continue. Overexpression studies of this gene would also yield invaluable insights into its function with the creation of still further phenotypes to study. Once it is known what compounds this flippase is specifically transporting studies could then be undertaken to ascertain how it performs this.

5.7 Signal recognition particle receptor-like protein cDNA 91G3

The signal recognition particle receptor-like proteins were also differentially under expressed in the mutant in each of the alkaloid mutants in comparison to their respective control phenotypes.

Signal recognition particle receptor (SRPR) proteins are necessary for efficient export of extra-cytoplasmic ER-targeted proteins. They are known to bind to the signal sequence when it emerges from the ribosomes; the signal recognition particle, which it binds to, consists of a 4.5S RNA molecule and protein Ffh. The protein has a two-domain structure: the G-domain binds GTP; the m-domain binds the RNA and also binds the peptide signal sequence. The *Arabidopsis* srpr-like proteins that align with the SRPR-like proteins found in these experiments have only been given function by homology. However these proteins are highly conserved across all classes of phyla due to their important roles in the transfer of the newly synthesised protein to its targeted membrane or compartment within the cell.

The SRPR-like proteins found to be highly differentially downregulated in the mutant *top1* could be responsible for:

- the transfer of the CytP450 enzymes assumed to be responsible for the demethylation of thebaine, codeine and oripavine, to the endoplasmic

reticulum and subsequently to the latex membranes or vacuoles where they act to synthesise codeine and morphine from thebaine; or

- the targeting to the ER of a membrane transporter required to relocate alkaloids between cells or sub-cellular compartments.

The high conservation of SRPR-like proteins suggests RNAi induced silencing may have too many effects to be viable. Transgenic overexpression may prove the more revealing approach.

5.8 Phosphoenolpyruvate carboxykinase cDNA 157C4

PEPCK catalyses the decarboxylation of oxaloacetate to phosphoenolpyruvate (Figure 5.11). In C4 plants it is involved in CO_2 concentration, while in CAM plants it is involved in C4 acid decarboxylation. The exact role of PEPCK is unclear in C3 plants though it appears to be involved in gluconeogenesis (the synthesis of glucose from non-carbohydrates) and may also be involved in amino acid synthesis. In plants that use the C4 photosynthetic pathway, PEPCK is characteristically expressed with several other photosynthetic enzymes to high levels in bundle sheath cells allowing concentration of CO_2 around Rubisco, and thus reduced photorespiration. C3 plants accumulate PEPCK enzyme in cells around the vascular system of petioles and stems where they are expressed to comparable levels to C4 plants (Hibberd & Quick, 2002).

PEPCK can be found in other tissues where its function is unclear. Examples include: in resin ducts, where it may be involved in plant defence; in developing seeds, with a possible role in metabolism of nitrogenous assimilates; and in cucumber phloem where it may be involved in amino acid metabolism (Chen *et al.*, 2004). So there is a potential role for PEPCK in the metabolism and/or conversion of amino acids in plants. It is possible that an affect on the morphinan alkaloid pathway, which is derived from the amino acid tyrosine, could cause a corresponding

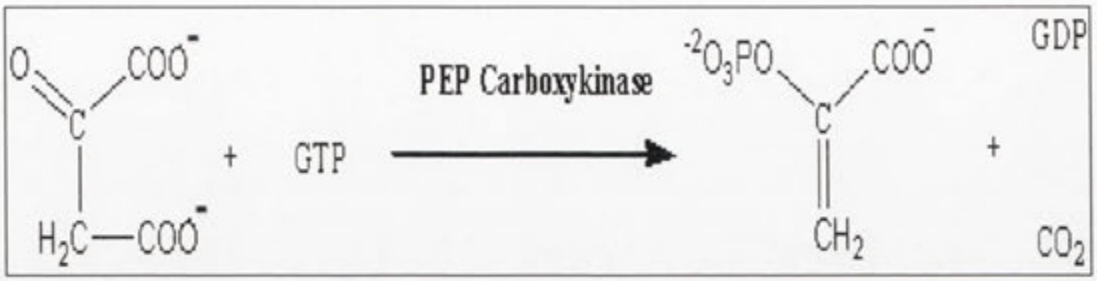


Figure 5.11: Illustration of the catalysis of oxaloacetate to phosphoenolpyruvate by phosphoenolpyruvate carboxykinase.

feedback on PEPCK if it is part of tyrosine metabolism. However this is considered a very unlikely association with *top1*.

5.9 Aldose 1- epimerase-like cDNA 111C7

The aldose 1-epimerase-like cDNA 111C7 has homology to a family of galactose mutarotase-like proteins. These proteins have been well characterised in bacteria, fungi and animals. They are thought to have a carbohydrate binding domain and act in the epimerisation of sugars, such as converting β-D-galactose to α-D-galactose (Thoden *et al.*, 2002). Expressed genes with homology to these mutarotase genes have been found in plants and have been shown to be glycoproteins localised to the nuclear rim (Heese-Peck, A., and Raikhel 1998). Enzymatic studies have been undertaken on purified extracts of plant material and have shown that the enzyme can catalyse a number of interconversions of sugars, particularly pentoses and hexoses related to D-glucose (Bailey, *et al.*, 1967). The poppy cDNA 111C7 has homology to the sugar binding mutarotase domain of *Arabidopsis* genes in the database (Figure 5.12).

The domain of highest homology between the aldose epimerises is the sequence: EQGDAITFEGEVDKVYVSTPTKIAILDHEKKRTFVLRKDGLPDAVVWNPWD KKAKAMADFGDDEYKHMLCVEAAVVEKPITLKPGEEWKGRQELS. This is

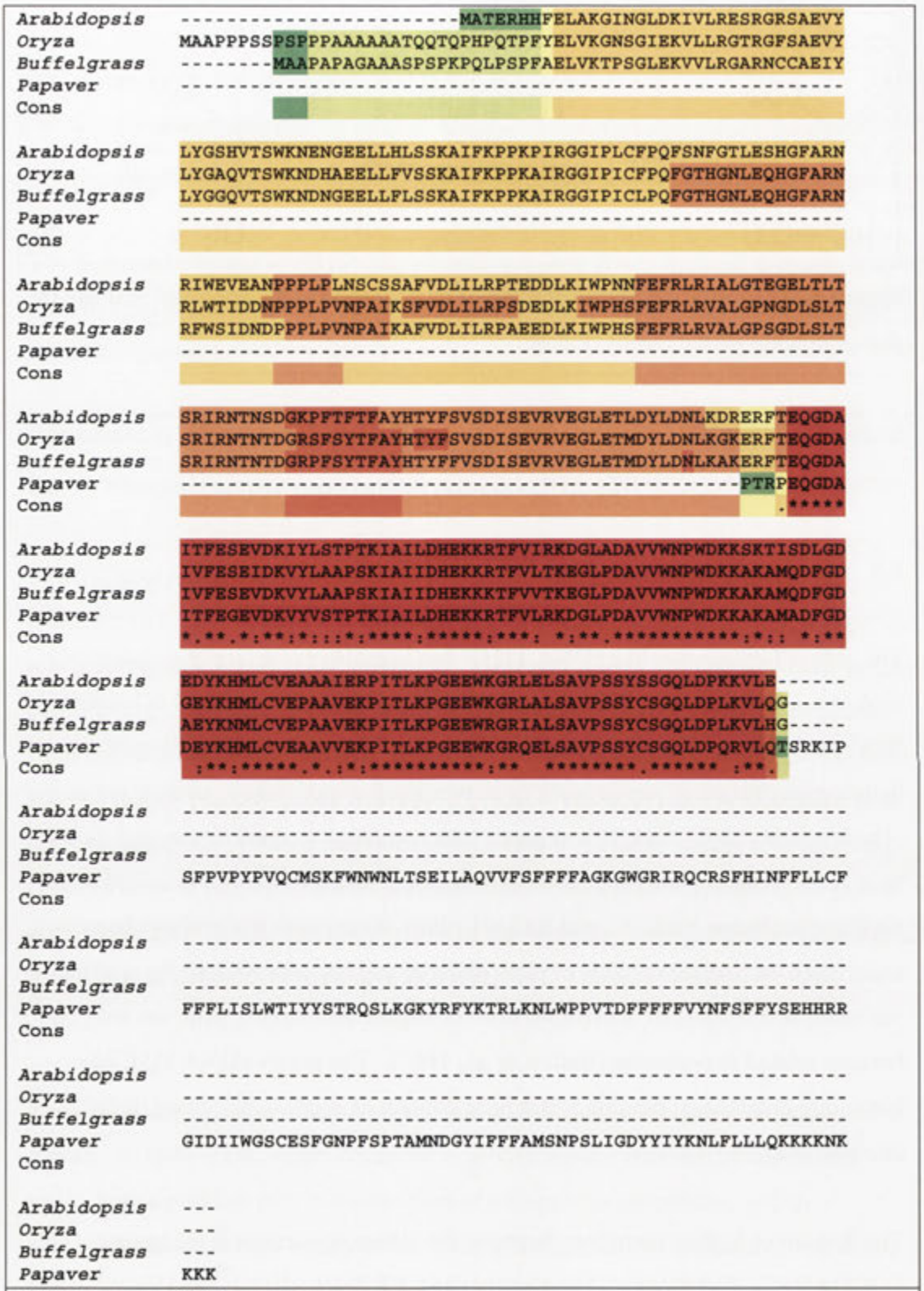


Figure 5.12: Alignment of the *P. somniferum* from this thesis 111C7 with several aldose 1-epimerases from the database. The label Cons is the conserved sequence

between the translated alignments. The colour code BAD AVG GOOD represents the quality of the alignment.

Arabidopsis thaliana aldose 1-epimerase family protein. NCBI Accession number NP_200543.

Oryza sativa (japonica cultivar-group) putative Aldose 1-epimerase. NCBI Accession number BAD05401.

Pennisetum ciliare (*Buffelgrass*). Possible apospory-associated protein c. NCBI Accession number Q40784.

T-COFFEE, Version_1.41 was used for the amino acid alignments (Notredame *et al.*, 2000).

the mutarotase domain from the pFam database (Schultz *et al.*, 1998). The *P. somniferum* cDNA 111C7 extends much further than any of the plant aldose epimerises in the database. It is possible that this is 3' UTR or that the gene could have a different function.

5.10 cDNAs with no homology to genes in the database (27D12, 91B6, 54F9, and 151H8)

Several cDNAs revealed no homology to any sequences when blasted against the NCBI database. These three genes are also differentially under-expressed in several other alkaloid mutations, suggesting a general pleiotropic effect of various perturbations of the pathway. Further work testing the association of the expression of these genes with the inheritance of the *top1* phenotype is required.

Two of the genes, 91B6 and 151H8, without any homology to genes in the NCBI database, were under-expressed in the *top1* phenotype by the microarray analysis, however the qRT-PCR data showed that there was less than a two-fold difference in expression of both genes. It is possible that a mutated gene could have lost all function and yet its transcription is nearly unaffected. It is also possible that a very

subtle drop in transcript level of a regulator could have much greater effects on the enzymes and pathway.

A further gene, 54F9, with no significant alignments in the database, showed an increased difference of expression between the *top1* mutant and the *P. somniferum* control in the qRTPCR data compared to the microarray data. This can be explained by the presence of a gene family or more than one gene with homologous regions that are binding to the cDNA spot on the array during the hybridisation process. If primers are designed to a region that is specific to the cDNA of interest (in this case 54F9) it is possible to eliminate the other expressed genes with homologous regions.

With so few likely candidates it is definitely worth considering transgenic experiments to test 54F9, 91B6 and151H8 in either or both of the following ways:

- Hairpin RNAi silencing of 54F9, 91B6 and 151H8 to see if the *top1* phenotype is recreated or morphine synthesis otherwise perturbed.
- Sense expression of 54F9, 91B6 and 151H8 in the *top1* mutant to see if the mutation is complemented.

5.11 F3 Segregating Population from *top1* x wild type cross

An F3 segregating population from a cross between *top1* and a wild type poppy was available from Tasmanian Alkaloids. It was considered likely that any causal association between the *top1* mutation and the set of ten differentially expressed genes would be evident as a co-segregation in the F3 population.
The material used for the *top1* experiments was derived from commercial lines that have undergone some breeding so as to eliminate any non *top1* phenotypic mutations, however the breeding was not extensive and there was still a possibility that there were secondary unrelated mutations in the *top1* germplasm. It was hoped that microarray experiments with the F3 segregating population would further shorten the list of genes of interest.

RNA was extracted from eight F3 segregants that were shown to have the *top1* phenotype (thebaine and oripavine) and from four F3 segregants shown to have the morphine phenotype. The bulked segregant RNA was used for a microarray experiment. Plant phenotypes were determined by a combination of HPLC and TLC analysis as described in Chapter 2. 25 µg of RNA from each of the four plants of each phenotype were pooled, giving the 100 µg of RNA required for the microarray labelling reaction, for microarray analysis. A total of four slides were used for the F3 bulked segregant experiments, including two biological replicates and a technical replicate of each. Biological replicates consisted of pooling RNA from a different set of four F3 plants of either the thebaine or morphine phenotype (Figure 5.13). By pooling the RNA samples of the F3 progeny of the same phenotype it was expected to eliminate the variability in expression between individual plants and highlight those genes that were only directly related to the *top1* phenotype. Microarray results showed several genes that were differentially up or down regulated in the F3 *top1* phenotype plants compared with the F3 morphine phenotype plants (Figure 5.14). Of the genes differentially expressed in the F3 bulked segregant microarray analysis only three (116D7, 221C9, and 151H8) of the set of ten genes from the earlier *top1* experiments were found to be differentially expressed and these three genes were below the differential cut-off value of 2 which has been applied in the earlier *top1* experiments (Table 5.3). 116D7 has its closest BLAST x hit to a putative chlorophyll a-b binding protein from *A. thaliana*. 221C9 has its closest match to an aspartate aminotransferase. Neither of these genes are likely candidates for the enzyme responsible for the demethylation of thebaine and oripavine. The third cDNA, 151H8, has no substantial hits in the database. Further work is needed to determine the nature of this gene and its interaction with morphine biosynthesis.

Given that the F3 bulked segregant microarray analysis was undertaken with limited replication, it is difficult to judge how much weight to give its conclusions relative to the highly replicated earlier *top1* analysis. Certainly it strengthens the argument for further attention to be paid to the overlapping three genes mentioned above. None of

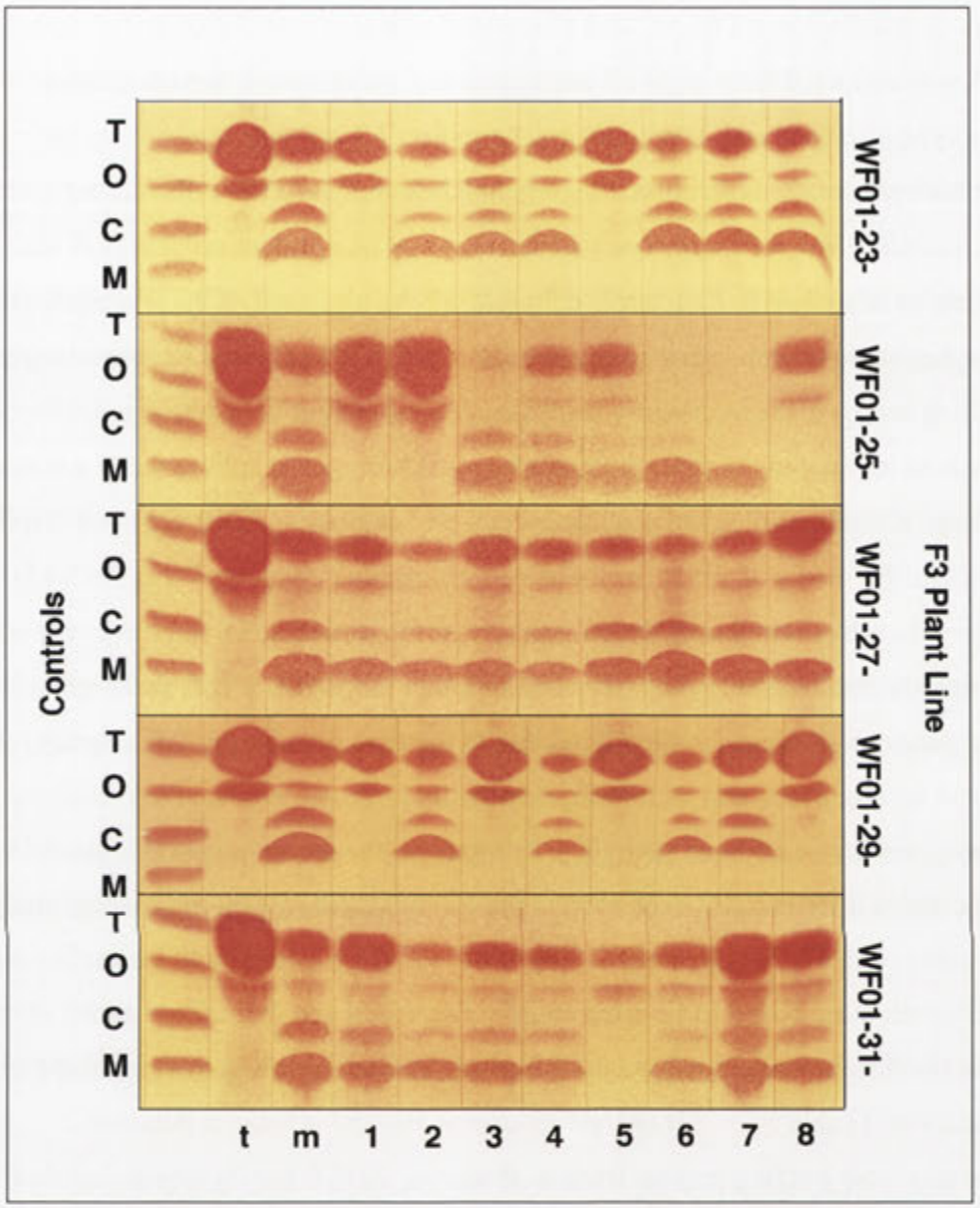


Figure 5.13: TLCs of the five segregating F3 populations (WF01-31-, WF01-29-, WF01-27-, WF01-25-, WF01-23-) tested for the morphine or thebaine phenotype. Four thebaine and oripavine accumulating mutants were selected for the test sample on the microarray (WF01-31-5, WF01-29-1, WF01-25-8, WF01-23-5). Four of the F3 progeny segregating for the morphine control phenotype were selected for the reference sample (WF01-31-2, WF01-29-6, WF01-25-3, WF01-23-6). The alkaloid controls are run in the far left lane with morphine (M), codeine (C), oripavine (O) and thebaine (T). The *top1* parental phenotype is the second lane from the left (t)

and the third lane is the morphine control parental phenotype (m). Lanes 1-8 represent the different F3 individuals of each segregating F3 population.

---

the three genes has a hit to a gene resembling a gene with demethylase activity or a P450. It is thought that the gene responsible for this enzymatic demethylation of thebaine and oripavine is a P450 in animal tissues (Kodeira and Spector 1988). However the plant may have evolved an alternate *O*-demethylase that is responsible for the reaction and this may not be a cytochrome P450.

It is interesting to observe that one of the differential genes (101D1) in the F3 bulked segregant microarray experiment was a codeinone reductase cDNA which was more highly expressed in the mutant F3 lines than the control F3 lines (Table 5.3). This codeinone reductase from the cDNA library is very similar to the NCBI entries. It is known that codeinone reductase is a member of a large gene family with as many as twelve copies thought to be present in *P. somniferum* (Unterlinner *et al.*, 1999) and 101D1 may be a different codeinone reductase than has yet been described. It is surprising that only one codeinone reductase clone appeared as differential, given the high homology between members of this multigene family.

Another interesting gene that was differentially expressed in this experiment was cDNA 2A2 (more highly expressed in the mutant), which has homology to a serine/threonine- protein kinase involved in brassinosteroid signal transduction. It is thought to be involved in the perception of systemin, a peptide hormone responsible for the systemic activation of defence genes in the leaves of wounded plants. It is also thought to regulate a signaling cascade involved in plant development (reviewed in Luan 2003). However, neither the codeinone reductase cDNA nor cDNA 2A2 were differentially expressed in the original *top1* microarray experiments, suggesting that these genes are being differentially expressed as a result of factors other than the mutation responsible for the *top1* phenotype. These factors could be as a result of differences in the growing conditions, small differences in light due to placement of

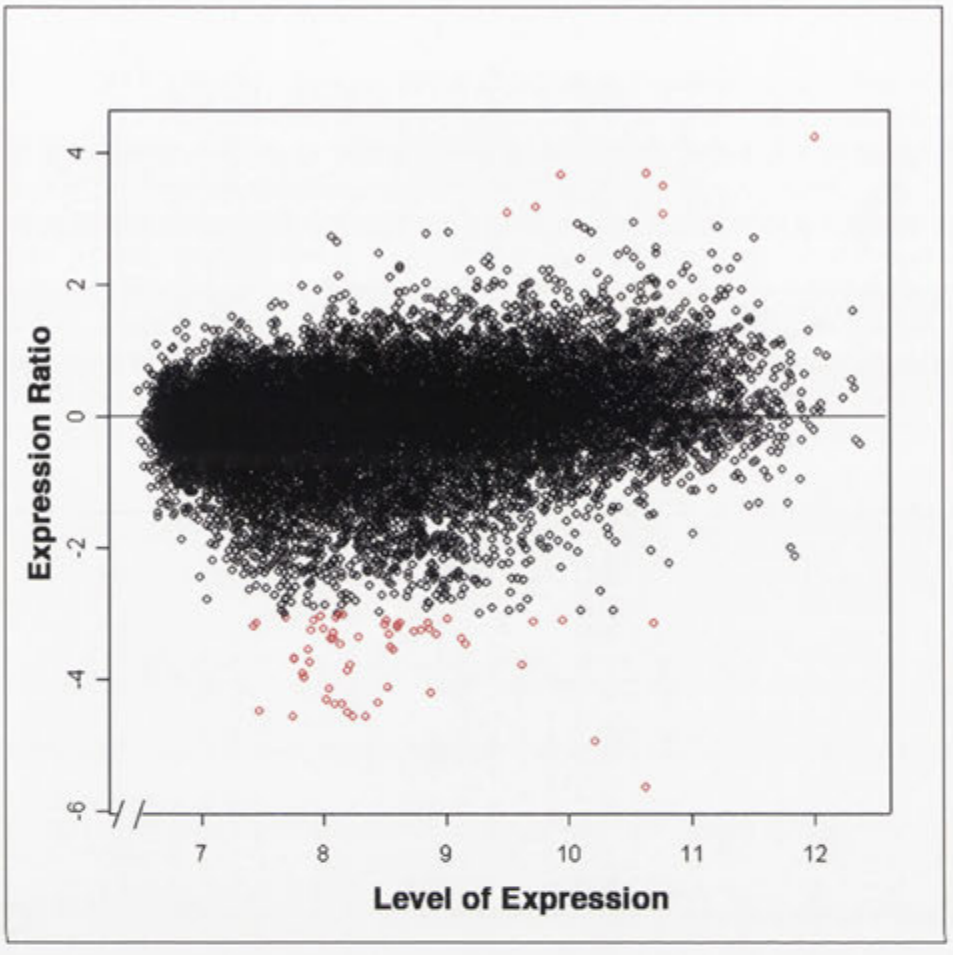


Figure 5.14: Graphical representation of the differentially expressed genes in the F3 population segregating for the *top1* phenotype compared to the F3 population segregating for the morphine wildtype phenotype.

---

the pots in the glasshouse or even possible insect or microbe attack. C13 is a genomic DNA sequence from *Arabidopsis,* which is a non-expressed genomic DNA sequence used as a negative control and printed several times on the microarray. Only one of these replicate control spots was differential in the F3 experiment (Table 5.3). There was apparently an expressed poppy sequence having enough homology to cross hybridize. The failure of the replicate spots to be differential suggests the appearance of C13 is an undefined technical error. Several other differential genes in Table 5.3 did not have high enough homology to any sequences in the NCBI database and so remain without any known function. Also several cDNAs with good homology to photosynthetic enzymes appear as differential in the F3 experiment

Table 5.3: Differentially expressed genes in the F3 segregating population for the *top1* phenotype.

Name	Expression Ratio	Accession Number	Description	E-value
148A9	+2.6	-	No Significant Hits	-
84C9	+2.2	O64452	*N. paniculata*. npca1 Carbonic anhydrase, chloroplast precursor (Carbonate dehydratase) (ec 4.2.1.1).	4E-99
33H5	+1.8	AF389262	*Dicentra eximia* 26S ribosomal RNA gene, partial sequence.	1E-130^
51C1	+1.8	CB24_TOBAC	*N. tabacum* chlorophyll a-b binding protein 40, chloroplast precursor	1E-135
116D7†	+1.7	ATU03395	*A. thaliana* psi type ii chlorophyll a/b-binding protein, putative.	6E-95
221C9†	+1.7	AF419301	*Securigera parviflora*. Aspartate aminotransferase.	2E-54
151H8†	+1.6	-	No Significant Hits	-
91F8*	-2.1	-	No Significant Hits	-
C13	-2.1	-	Genomic repeat	-
29A10	-2.1	-	No Significant Hits	-
101D1	-2.1	Q9SQ69	*P. somniferum* NADPH-dependent codeinone reductase (ec 1.1.1.247)	1E-124
52H3	-2.1	AAD27877	*Vigna radiata* LHCII type III chlorophyll a/b binding protein	1E-123
133D7	-2.0	ATPA_TOBAC	*N. tabacum* ATP synthase alpha chain (ec 3.6.3.14).	1E-169
103D8*	-2.0	-	No Significant Hits	-
30F5	-2.0	-	No Significant Hits	-
2A2	-2.0	Q8LDB7	*A. thaliana* somatic embryogenesis receptor-like kinase, putative.	3E-55

*Only 50% of the spots analyzed for the cDNAs 91F8 and 103D8 were reliable for statistical analysis. A + value represents the gene is downregulated in the F3 plants with the *top1* phenotype. A – value represents the gene is upregulated in the F3 plants with the *top1* phenotype. † represents the three cDNAs that were significantly differentially under-expressed in the *top1* mutant microarray analysis that were also downregulated in the F3 experiment.
^ All E-values are the best BLAST x hits except for 33H5 which is the best BLAST n value.

that were not differential in the original *top1* experiments, and are unlikely to have any direct functional link to morphine biosynthesis.

5.12 Discussion & Conclusions

In summary this research has opened up a previously difficult area to study. Plant secondary metabolism is now recognised as a unique cross meshing of enzymes and products to give many useful products both to the plant (anthocyanins, plant defence compounds) and to man (morphine, codeine, caffeine, vinblastine). The *Papaveraceae* family is unique in its ability to accumulate morphine and *P. somniferum* is one of only a few members of this family to accumulate morphine to appreciable levels. The *top1* mutant has shown that intermediate morphinan alkaloids can accumulate in the plant to levels similar to that of the parent phenotype without any adverse affects on growth or maturity. The blockage in the pathway in *top1* seems to be at the enzymatic reaction following the synthesis of thebaine and oripavine in the bifurcated morphinan pathway. The microarray analysis, the first of its kind in any *Papaver* species, has revealed ten genes whose expression is affected in the *top1* mutant, and that may be involved in morphine biosynthesis indirectly; a number of those genes are novel, unknown in the GenBank database. The case for their interaction with the morphine pathway is strengthened by the fact that they are also differentially expressed in a number of other perturbations of the pathway. Their further study may add to our understanding of the regulation, enzymology and transport of the morphinan alkaloids.

This thesis has contributed to the understanding of morphinan alkaloid synthesis and accumulation in the 'opium poppy' *Papaver somniferum*. It describes the first molecular genetic analysis of the first stable mutations reported of morphine biosynthesis. What is believed to be a relatively simple mutation in a secondary metabolic pathway can have several other affects in what would have been thought to be unrelated metabolic areas. It is possible that several of these changes are attributable to indirect pleiotropic affects since they appear in several different alkaloid mutants. Further work is needed to address the nature of these associations.

Specifically this thesis has succeeded in each of its aims: firstly a *P. somniferum* cDNA library has been produced and used for printing the first published poppy microarray; secondly, several new *P. somniferum* mutants have been characterised; thirdly these mutants (and other pathway-variant germplasm) have been further analysed for transcriptional changes using the poppy cDNA microarray; fourthly, the transcriptional changes have been analysed through sequencing and Blast x analysis with public domain databases; lastly, cross comparisons of experiments and further molecular analyses have foccused attention on a set of ten genes and particularly a set of three genes. The aims of this thesis have been met and directions for further research have been set.

5.13 Future Work

Since the morphinan alkaloid pathway is not present in any model genetic system, resources such as defined insertional mutations are not available. Fortunately a reasonably efficient transformation system has been developed and the best method of further study of candidate genes will be to produce transgenic *P. somniferum* plants. Two types of transgenics could be informative: those containing hairpin RNAi silencing constructs of each individual gene; and those with sense over-expression constructs into various backgrounds including the various mutants. The possibility exists that RNAi transgenic plants may fail to give a phenotype because the genes may belong to functionally redundant families and where only some members are silenced by the chosen hairpin sequences. On the other hand the candidate gene may be functional in several other metabolic pathways, including primary metabolism, such that silencing could be lethal; in this case the only plants that might be regenerated would not be effectively silenced and show no phenotype. If a multigene family is suspected, such as the transporter-like proteins to which cDNA 52F2 belongs, it would be possible to isolate 5' and or 3' UTRs of several members of the family and create knockouts of each individual of the gene family

separately and determine which member of the family is responsible or associated with the morphine pathway.

The set of three genes could possibly be homologues of other genes that are actually involved in the pathway. Microarray studies utilise the principal of nucleotide hybridisation and it is possible several different types of genes are cross hybridising to the cDNAs printed on the array. The hybridisation could involve gene families or classes of genes with specific regions of homology. The hybridisation procedure used was relatively stringent but it will not necessarily stop all permutations of cross hybridisation (Girke *et al.*, 2000). This could explain why several of the qRT-PCR results were not consistent with the microarray results. This could also explain why the clone 116D7 was significant, even though it has a region of high homology to a chlorophyll a/b binding protein. There may be a sufficient region of homology between 116D7's transcript and another gene's transcript that is associated with the *top1* mutant phenotype.

There are several possible explanations for why the gene responsible for the mutation was not found using the microarray analyses.

1. The gene responsible for *top1* is not on the array. The cDNA library employed for the microarray showed very low redundancy, but there was still some redundancy. Estimates for the number of genes in the *Arabidopsis* genome range from 20 000 to 60 000 (Bevan *et al.*, 1998). Therefore the *P. somniferum* 17 000 cDNA array used in this study may have represented around 50-75% of the expressed genome. It is therefore possible that the gene responsible for *top1* is not represented on the array.
2. The *top1* mutation does not affect transcription. It is possible that the cDNA was present on the array, however the *top1* mutation does not affect transcription. Function may be lost due to a critical change in the coding region but the transcript level is unaltered. The mutant peptide may be altered in folding, active enzyme site or post-translational modification. A mutant enzyme may have altered signal peptides affecting localisation or anchoring which prevent access to

its normal substrates. If the mutated *top1* gene encodes a morphinan regulator protein, it may be altered in DNA binding while leaving the transcript levels unchanged.

Using the arrays however has implicated several genes as having some role or association in alkaloid biosynthesis either indirectly or directly. A number of these warrant further study, such as the cDNAs associated with *top1* with no significant database homology. Further characterisation of these genes would involve isolating full-length clones and sequencing to firstly determine if there are homologous clones in the database to give some clues as to their function. Secondly if no obvious function could be ascertained, it would then be important to create transgenic knockouts, either using antisense for a limited response or inverted repeat technology for a more severe affect. A recreation of a *top1* phenotype in a 'knockout' would confirm the causal relationship to the mutation. Thirdly, the *top1* phenotype could be complemented (repaired) with a full-length sense version of the gene or cDNA.

Further analysis using the microarrays could be undertaken using RNA isolated from the latex and the roots specifically since it has been shown that the ten genes do have different levels of expression in the roots and stem of the plants.

While the gene or genes responsible for the *top1* phenotype have not yet been identified from these experiments, the set of ten genes have yielded several important leads. In addition the three genes that were differentially down regulated, even in the F3 segregating population for the *top1* phenotype, 116D7, 221C9, 151H8, could be valuable as molecular markers for this economically important trait. Further work is required to determine if one or more of these genes is sufficiently physically linked to the mutation and if a useful polymorphism can be found to facillitate marker-assisted breeding. A molecular marker map of *P. somniferum* is being constructed by Tasmanian Alkaloids; eventually it may be possible to clone the *top1* gene by map-based walking beginning with these three apparently linked genes.

The genes isolated using the current microarray study are the first step in characterising the structural enzymes and illustrating the varying level of controls of the morphinan pathway. Other experiments to undertake would include further analysis of the differential proteins in the latex. By looking at 2D protein gels of segregating populations for the *top1* phenotype it would be possible to isolate proteins that are differential in the highly restricted morphine-accumulating tissue sampled in the latex. Analysis of the latex proteins assumes that the genes responsible for the phenotype are found here and that it is not a trafficking of intermediates and end products in and out of the laticifer system itself. The advantage of using this technique is that the latex fraction will be enriched for non-photosynthetic proteins, which are unlikely to be involved in the direct synthesis of morphinans. By sequencing the differential proteins from the 2D gels the short protein sequences can be compared to the translations of the differentially expressed genes identified from the microarray experiments.

The morphine biosynthetic mutants and their global transcript analysis described in this thesis have advanced knowledge of the morphinan alkaloid pathway in *P. somniferum*. Much remains to be done before we comprehensively understand the regulatory controls, the nature of uncharacterised enzymatic steps, the missing gene sequences of structural enzymes, the various locations of biosynthesis, the trafficking of enzymes and intermediates, and the constraints on product accumulation.

Bibliography

Abrahams, S., Tanner, G.J., Larkin, P.J. and Ashton, A.R. "Identification and biochemical characterization of mutants in the proanthocyanidin pathway in *Arabidopsis*." Plant Physiology 130: 561-576. 2003a

Abrahams, S., Lee, E., Walker, A.R., Tanner, G.J., Larkin, P.J. and Ashton, A.R. "The *Arabidopsis* TDS4 gene encodes leucoanthocyanidin dioxygenase (LDOX) and is essential for proanthocyanidin synthesis and vacuole development." The Plant Journal 35 (5): 624-636. 2003b

Aerts, R.J., Gisi, D., De Carolis, E., De Luca, V. and Baumann, T.W. "Methyl jasmonate vapour increases the developmentally controlled synthesis of alkaloids in *Catharanthus* and *Cichona* seedlings." The Plant Journal 5: 635-643. 1994

Akil, H., Watson, S.J. and Walker, J.M. "Endogenous opioids: biology and function." Annual Review Neurosciences 7: 223-255. 1984

Allen, R., Millgate, A., Chitty, J., Thistleton, J., Miller, J., Fist, A., Gerlach, W. and Larkin, P.J. "Replacing morphine with the non-narcotic reticuline in opium poppy by RNAi gene silencing." Nature Biotechnology 22: 1559-1566. 2004

Amann, T., Roos, P.H., Huh, H. and Zenk, M.H. "Purification and characterisation of a cytochrome P450 enzyme from pig liver, catalysing the phenol oxidative coupling of (*R*)-reticuline to salutaridine, the critical step in morphine biosynthesis." Heterocycles 40: 425-440. 1995

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic Acids Research 25: 3389-3402. 1997

Ayyangar, N.R. and Bhide, S.R. " Seperation of eight alkaloids and meconic acid and quantitation of five principal alkaloids in gum opium by gradient reverse-phase high-performance liquid chromatography." Journal of Chromatography 436: 455-465. 1988

Bailey, J.M., Fishman, P.H. and Pentchev, P.G. "Studies on mutarotases. I. Purification and properties of a mutarotase from higher plants." The Journal of Biological Chemistry 242 (18): 4262-4269. 1967

Balbi, G. "Poppy seed oil in manufacture of paints and varnishes." Olearia 14: 97-104. 1960

Baldwin, B.G. "Phylogenetic utility of the internal transcribed spacers of nuclear

ribosomal DNA in plants: An example from the *Compositae*." Molecular Phylogenetics and Evolution 1: 3-16. 1992

Baldwin, I.T. and Ohnmeiss, T.E. "Alkaloid responses to damage in *Nicotiana* native to North America." Journal of Chemical Ecology 19: 1143-1153. 1993

Bare, C.E., Coffman, C.B. and Gentner, W.A. "Growth, development and thebaine content of *Papaver bracteatum* Lindl. in relation to spacing." Agronomy Journal 79 (5): 935-937. 1987

Beare, R.J.L., Gray, L., Nera, E.A. and Levin, O.L. "Nutritional properties of poppy seed oil relative to some other oils." Nutrition and Metabolism 23 (4): 335-346. 1979

Bentley, K.W. and Dyke, S.F. "The structure of pseudomorphine." Journal Chemical Society: 2574-2577. 1955

Bernáth, J. "VIII Utilization of poppy seed" In: Poppy: The Genus *Papaver*. Pg 347-342. Ed: Jeno Bernath. Amsterdam Harwood Academic, 1998

Bernáth, J., Németh, É. and Petheõ, F. "Alkaloid accumulation in capsules of the selfed and cross-pollinated poppy." Plant Breeding 122: 263-267. 2003

Bevan , M. "Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*." Nature 391: 485-488. 1998

Bird D.A., Franceschi, V.R. and Facchini, P.J. "A tale of three cell types: Alkaloid biosynthesis is localized to sieve elements in opium poppy." The Plant Cell 15: 2626-2635. 2003

Blackford, S., Rea, P.A. and Sanders, D. "Voltage sensitivity of H^+/Ca^{2+} antiport in higher plant tonoplast suggests a role in vacuolar calcium accumulation."The Journal Biological Chemistry 265: 9617-9620. 1990

Blumwald, E. and Poole, R. J. "Na/H-antiport in isolated tonoplast vesicles from storage tissue of *Beta vulgaris*." Plant Physiology 78: 163-167. 1985a

Blumwald, E. and Poole, R. "Nitrate storage and retrieval in *Beta vulgaris*- Effects of nitrate and chloride on proton gradients in tonoplast vesicles." Proceedings of the National Academy of Sciences USA 82: 3683-3687. 1985b

Blumwald, E. "Tonoplast vesicles as a tool in the of ion transport at the plant vacuole." Physiologia Plantarum 69: 731-734. 1987

Bock, A., Wanner, G. and Zenk, M.H. "Immunocytological localization of two enzymes involved in berberine biosynthesis." Planta 216: 57-63. 2002

Bouzayen, M., Ltche, A., Pech, P-C. and Marigo, G. "Carrier-mediated uptake of 1-(malonylamino)-cyclopropane-1-carboxylic acid in vacuoles isolated from *Catharanthus roseus* cells." Plant Physiology 91: 1317-1322. 1989

Bremberger, C., Haschke, H-P. and Lüttge, U. "Separation and purification of the tonoplast ATPase and pyrophosphatase from plants with constitutive and inducible crassulacean acid metabolism." Planta 175: 465-470. 1988

Breton, G., Danyluck, J., Frenette, J-B. and Sahan, F. "Expression profiling and bioinformatic analyses of a novel stress-regulated multispanning transmembrane family from cereals and *Arabidopsis*." Plant Physiology 132: 64-74. 2003

Briskin, D.P., Thornley, W.R. and Wyse, R.E. "Membrane transport in isolated vesicles from sugarbeet taproot. II. Evidence for a sucrose/K^+-antiport." Plant Physiology 78: 871-875. 1985

Brown, P.O. and Botstein, D. "Exploring the new world of microarrays." Nature Genetics Supplement 21: 33-37. 1999

Brochmann-Hanssen, E. "A second pathway for the terminal steps in the biosynthesis of morphine." Planta Medica 50: 343-345. 1984

Brochmann-Hanssen, E., Hirai, K., Nielsen, S., Pfeifer, S., Mann, I. and Kühn, L. "Isolation of *N*-methyl-14-*O*-desmethylepiporphyroxine." Journal of Pharmaceutical Sciences 57: 30-35. 1968

Brownstein, M.J. "A brief history of opiates, opioid peptides, opioid receptors." Proceedings of the National Academy of Sciences USA 90: 5391-5393. 1993

Carporale, L.H. "Chemical ecology: a view from the pharmaceutical industry." Proceedings of the National Academy of Sciences USA 92: 72-82. 1995

Chapple, C. "Molecular-genetic analysis of plant cytochrome P450-dependent monooxygenases." Annual Reviews of Plant Physiology and Plant Molecular Biology 49: 311-343. 1998

Chauhan, S.P. and Patra, N.K. "Mutagenic effects of combined and single doses and EMS in opium poppy." Plant Breeding 110 (4): 342-345. 1993

Chen, Li Mei; Li, Kun Zhi; Miwa, T. and Izui, K."Overexpression of a cyanobacterial phosphoenolpyruvate carboxylase with diminished sensitivity to feedback inhibition in *Arabidopsis* changes amino acid metabolism." Planta 219 (3): 440-449. 2004

Chitty, J. A., Allen, R.S., Fist, A.J. and Larkin, P.J. "Genetic transformation in commercial Tasmanian cultivars of opium poppy, *Papaver somniferum*, and movement of transgenic pollen in the field." Functional Plant Biology, 30: 1045-1058. 2003

Choi, K., Morishige, T. and Sato, F. "Purification and characterization of coclaurine *N*-methyltransferase from cultured *Coptis japonica* cells." Phytochemistry 56: 649-655. 2001

Choi, K., Morishige, T., Shitan, N., Yazaki, K. and Fumihiko, S. "Molecular cloning and characterization of coclaurine *N*-methyltransferase from cultured cells of *Coptis japonica*." The Journal of Biological Chemistry 277 (1): 830-835. 2002

Chou, W-M. and Kutchan, T. "Enzymatic oxidations in the biosynthesis of complex alkaloids." The Plant Journal 15 (3): 289-300. 1998.

Coelho, P.S.R., Bryan, A.C., Kumar, A., Shadel, G.S. and Michael Snyder. "A novel mitochondrial protein, Tar1p, is encoded on the antisense strand of the nuclear 25S rDNA." Genes and Development 16: 2755-2760. 2002.

Corrigan, D. and Martyn, E.M. "The thebaine content of ornamental poppies belonging to the *Papaver* section *Oxytona*." Planta Medica 42 (1): 45-49. 1981

Cox, B.M., Gentleman, S. and Su, T.P. "Further characterization of morphine-like peptides (endorphins) from pituitary." Brain Research 115: 285-296. 1976

Cunningham, F.X., Jr., Pogson, B., Sun, Z., McDonald, K.A., Della-Penna, D. and Gantt, E. "Functional analysis of the β and ε lycopene cyclase enzymes of *Arabidopsis* reveals a mechanism for control of cyclic carotenoid formation." The Plant Cell 8: 1613-1626. 1996

Darwin, W.D. and Cone, E.J. "Fluorescence properties of pseudomorphine and congeners: structure-activity relationships." Journal of Pharmacological Sciences 69, 253-7. 1980

Dawson, R.F. "The localization of the nicotine synthetic mechanism in the tobacco plant." Science 94 (2443): 396-397. 1941

Dawson, R.F. " Accumulation of nicotine in reciprocal grafts of tomato and tobacco." American Journal of Botany 29: 66-71. 1942

Debeaujon, I., Peeters, A.J.M., Leon-Kloosterziel, K.M. and Koorneef, M. "The *TRANSPARENT TESTA12* gene of *Arabidopsis* encodes a multidrug secondary transporter-like protein required for flavonoid sequestration in vacuoles of the seed coat endothelium." The Plant Cell 13: 853-871. 2002

Decker, G., Wanner, G., Zenk, M.H. and Lottspeich, F."Characterization of proteins in latex of the opium poppy (*Papaver somniferum*) using two-dimensional gel electrophoresis and microsequencing." Electrophoresis 21 (16): 3500-3516. 2000

De-Eknamkul, W. and Zenk, M.H. "Enzymatic formation of (*R*)-reticuline from 1,2-dehydroreticuline in the opium poppy plant." Tetrahedron Letters 31(34): 4855-4858. 1990

De-Eknamkul, W. and Zenk, M.H. "Purification and properties of 1,2-dehydroreticuline reductase from *Papaver somniferum* seedlings." Phytochemistry, 31: 813-821. 1992

De Luca, V., Marineau, C. and Brisson, N. "Molecular cloning and analysis of a cDNA encoding a plant tryptophan decarboxylase: comparison with animal DOPA decarboxylases." Proceedings of the National Academy of Sciences of the USA 86: 2582-2586. 1989

De Luca, V. and St Pierre, B. "The cell and developmental biology of alkaloid biosynthesis." Trends in Plant Science 5 (4): 168-173. 2000

De Luca, V. and Laflamme, P. "The expanding universe of alkaloid biosynthesis." Current Opinion in Plant Biology, 4: 225-233. 2001

Deus-Neumann, B. and Zenk, M.H. "A highly selective alkaloid uptake system in vacuoles of higher plants." Planta 162: 250-260. 1986

De Smet, P.A.G.M. "The role of plant derived-drugs and herbal medicines in healthcare." Drugs 54: 801-840. 1997

Dewick, P.M. "Alkaloids 6". In: Medicinal Natural Products: A Biosynthetic Approach. John Wiley & Sons, Ltd. Brisbane. 1998. Pages 315-337.

Diachenko, L.B., Ledesma, J., Chenchik, A.A. and Siebert, P.D. "Combining the technique of RNA fingerprinting and differential display to obtain differentially expressed mRNA." Biochemistry Biophysical Research Communications 219: 824-828. 1996

Dittrich, H. and Kutchan, T.M. "Molecular cloning, expression, and induction of berberine bridge enzyme, an enzyme essential to the formation of benzophenanthridine alkaloids in the response of plants to pathogenic attack." Proceedings of the National Academy of Sciences USA 88: 9969-9973. 1991

Dong, X., Braun, E.L. and Grotewald, E. "Functional conservation of plant secondary metabolic enzymes revealed by complementation of *Arabidopsis* flavonoid mutants with maize genes." Plant Physiology 127: 46-57. 2001

Donnerer, J., Oka, K., Brossi, A., Rice, K.C. and Spector, S. "Presence and formation of codeine and morphine in the rat." Proceedings of the National Academy of Sciences USA 83: 4566-4567. 1986

Dos Santos, R.I., Schripsema, J. and Verpoorte, R. "Ajmalicine metabolism in *Catharanthus roseus* cell cultures." Phytochemistry 35: 677-681. 1994.

Dostal, J. and Slavik, J. "Recent knowledge on sanguinarine and related alkaloids." Chemicke Listy 94 (1): 15-20. 2000

Dragar, C. and Bick, I.R.C. "Somniferine, a novel dimeric opium alkaloid." Tetrahedron letters 29: 3115-3116. 1988

D'Souza, I. and Schellenberg, G.D. "Tau Exon 10 expression involves a bipartite intron 10 regulatory sequence and weak 5'-3' splice sites."The Journal of Biological Chemistry 277: 26587 - 26599: 2002

El-Ahmady, S.H. and Nessler, C.L. "Cellular localization of tyrosine decarboxylase expression in transgenic opium poppy and tobacco." Plant Cell Reports 20: 313-317. 2001

Ellner, S.P. "Review of R, version 1.1.1." Bulletin of the Ecological Society of America 82; 127-128. 2001

Evdokimova, L. I. and Rat'kin, A. V. "Biochemical nature of flower color mutants of poppy *Papaver somniferum* L." Biology Bulletin of the Russian Academy of Sciences 26 (5): 430-435. 1999

Ensminger, P.A. "Control of developments in plants and fungi by far-UV radiation." Physiologia Plantarum 88, 3: 501-508. 1993

Ewing, B., Hillier, L., Wendl, M.C. and Green, P. "Base-calling of automated sequencer traces using phred. I. Accuracy assessment." Genome Research 8 (3): 175-185. 1998

Ewing, B. and Green, P. "Base-calling of automated sequencer traces using phred. II. Error probabilities." Genome Research 8 (3): 186-194. 1998

Facchini, P.J. and De Luca, V. "Differential and tissue-specific expression of a gene family for tyrosine/dopa decarboxylases in opium poppy." Journal of Biological Chemistry 269: 26684-26690. 1994

Facchini, P.J. and De Luca, V. "Expression in *Escherichia coli* and partial characterisation of two tyrosine/DOPA decarboxylases from opium poppy." Phytochemistry 38: 1119-1126. 1995a

Facchini, P.J. and De Luca, V. "Phloem-specific expression of tyrosine/dopa decarboxylase genes and the biosynthesis of isoquinoline alkaloids in opium poppy." The Plant Cell 7, 1811-1821. 1995b

Facchini, P. J., Penzes, C., Johnson, A.G. and Bull, D. "Molecular characterization of berberine bridge enzyme genes from opium poppy." Plant Physiology 112: 1669-1667. 1996

Facchini, P.J., Johnson, A.G. and Bird, D.A. "Regulation of alkaloid biosynthesis in opium poppy and related species: prospects for pathway engineering." AgBiotech News and Information 9 (6): 131N-138N. 1997

Facchini, P.J., Penzes-Yost, C., Samanani, N. and Kowalchuck, B. "Expression patterns conferred by tyrosine/dihydroxyphenylalinedecarboxylase promoters from opium poppy are conserved in transgenic tobacco." Plant Physiology 118: 69-81. 1998

Facchini, P., J. "Alkaloid biosynthesis in plants: biochemistry, cell biology, molecular regulation, and metabolic engineering applications." Annual Reviews in Plant Physiology and Plant Molecular Biology 52: 29-66. 2001

Facchini, P.J. and Park, S-U. "Developmental and inducible accumulation of gene transcripts involved in alkaloid biosynthesis in opium poppy." Phytochemistry 64: 177-186. 2003

Facchini P.J., Bird D.A., MacLeod B.P., Park S-U. and Samanani N. "Multiple levels of control in the regulation of alkaloid biosynthesis." In: Romeo JT, (Ed.) *Integrative Phytochemistry: From Ethnobotany to Molecular Ecology.* Amsterdam: Pergamon. 2003. Pages 143-180.

Fairbairn, J.W., Hakim, F. and Kheir, Y. E. "Alkaloidal storage, metabolism and translocation in the vesicles of *Papaver somniferum* latex." Phytochemistry 13: 1133-1139. 1974

Fairbairn, J.W. and Steele, M.J. "Biosynthetic and metabolic activities of some organelles in *Papaver somniferum* latex." Phytochemistry 20 (5): 1031-1036. 1981

Fairbairn, J.W. and Williamson, E.M. "Meconic acid as a chemotaxonomic marker in the *Papaveraceae*." Phytochemistry 17: 2087-2089. 1978

Farnsworth, N.R. and Bingel, A.S. In: New Natural Products And Plant Drugs With Pharmacological, Biological Or Therapeutical Activity. Springer: New York. 1997. Pages 61-73.

Fiehn, O. and Weckwerth, W. "Minireview: Deciphering metabolic networks".

European Journal of Biochemistry 270: 579-588. 2003

Fist, A.J., Byrne, C.J. and Gerlach, W.L. "*Papaver somniferum* strain with high concentration of thebaine and oripavine." United States Patent 6,067,749. 2000

Franke, R.; Hemm, M. R.; Denault, J. W.; Ruegger, M. O.; Humphreys, J. M. and Chapple, C. "Changes in secondary metabolism and deposition of an unusual lignin in the *Ref8* Mutant of *Arabidopsis*."The Plant Journal 30 (1): 47-59. 2002

Frenzel, T. and Zenk, M., H. "Purification and characterization of three isoforms of *S*-adenosyl-L-methionine: (*R*, *S*)-tetrahydrobenzylisoquinoline-N-methyltransferase from *Berberis koetineana* cell cultures." Phytochemistry 29: 3491-3497. 1990

Frick, S. and Kutchan, T., M. "Molecular cloning and functional expression of *O*-methyltransferases common to isoquinoline alkaloid and phenylpropanoid biosynthesis." Plant Journal 17: 329-339. 1999

Frick, S., Ounaroon, A. and Kutchan, T., M. "Combinatorial biochemistry in plants: the case of *O*-methyltransferases." Phytochemistry 56: 1-4. 2001

Frick, S., Chitty, J. A., Kramell, R., Scmidt, J., Larkin, P. J. and Kutchan, T., M. "Transformation of opium poppy (*Papaver somniferum* L.) with antisense berberine bridge enzyme 1 (anti-bbe 1) via somatic embryogenesis results in an altered ratio of alkaloids in latex but not in roots." Transgenic Research in press. 2004

Gates, M. and Tsundi, G., "The synthesis of morphine." Journal of the American Chemical Society 74:1109-1110. 1952.

Gerardy, R. and Zenk, M., H. "Formation of salutaridine from (*R*)-reticuline by a membrane-bound cytochrome P-450 enzyme from *Papaver somniferum*." Phytochemistry 32: 79-86. 1993a

Gerardy, R. and Zenk, M., H. "Purification and characterization of salutaridine: NADPH 7-oxidoreductase from *Papaver somniferum*." Phytochemistry 34: 125-132. 1993b

Getz, H.P. "Sucrose transport in tonoplast vesicles of red beet roots is linked to ATP hydrolysis." Planta 185: 261-268. 1991

Ginsburg, D. and Elad, D., "The synthesis of morphine." Journal of the American Chemical Society 76: 312-313. 1954.

Girke, T., Todd, J., Ruuska, S., White, J., Benning, C. and Ohlrogge, J. "Microarray analysis of developing *Arabidopsis* seeds." Plant Physiology 124: 1570-1581. 2000

Goldstein, A., Fischli, W., Lowney, L.I., Hunkapiller, and M., Hood, L. " Porcine pituitary dynorphin: complete amino acid sequence of the biologically active heptadecapeptide." Proceedings of the National Academy of Sciences USA 78: 7219-7223. 1981

Goldstein, A., Barrett, R.W., James, I.F., Lowney, L.I. and Weiz, C.J. "Morphine and other opiates from beef brain and adrenal." Proceedings of the National Academy of Sciences USA 82: 5203-5207. 1985

Gollwitzer, J., Lenz, R., Hampp, N. and Zenk, M.H. "The transformation of neopinone to codeinone in morphine biosynthesis proceeds non-enzymaticaly." Tetrahedron Letters 34 (36): 5703-5706. 1993

Gomez-Serranillos, M.P., Palomino, O.M., Carretero, E. and Villar, A. "Analytical study and analgesic activity of oripavine from *Papaver somniferum* L." Phytotherapy Research 12 (5): 346-349. 1998

Goossens, A., Häkkinen, S.T., Laakso, I., Seppänen-Laakso, T., Biondi, S., De Sutter, V., Lammertyn, F., Nuutila, A.M., Söderlund, H., Zabeau, M., Inze, D. and Oksman-Caldentey, K-M. "A functional genomics approach toward the understanding of secondary metabolism in plant cells." Proceedings of the National Academy of Sciences USA (14) 100: 8595-8600. 2003

Granado, F., Olmedilla, B., Gil-Martinez, E. and Blanco, I. "A fast, reliable and low-cost saponification protocol for analysis of carotenoids in vegetables." Journal of Food Composition And Analysis 14: 479-489. 2001

Gross, C., Kelleher, M., Iyer, V.R., Brown, P.O., Dennis, R. and Winge, D.R. "Identification of the copper regulon in *Saccharomyces cerevisiae* by DNA microarrays." Journal of Biological Chemistry 275: 32310-32316. 2000

Grothe, T., Lenz, R. and Kutchan, T.M. "Molecular characterization of the salutaridinol 7-*O*-acetyltransferase involved in morphine biosynthesis in opium poppy *Papaver somniferum*." Journal of Biological Chemistry 276, 33: 30717-30723. 2001

Gundlach, H., Müller, M.J., Kutchan, T.M. and Zenk, M.H. "Jasmonic acid is a signal transducer in elicitor-induced plant cell cultures." Proceedings of the National Academy of Sciences USA 89: 2389-2393. 1992

Hager, A. and Hermsdorf, P. "A H/Ca-antiporter in membranes of microsomal of vesicles from maize coleoptiles, a secondary energized Ca pump." Zeitschrift fur Naturforschung 36c: 1009-1012. 1981

Hahlbrock, K. and Grisebach, H. "Enzymic controls in the biosynthesis of lignin and flavonoids." Annual Review of Plant Physiology 30: 105-130. 1979

Hanahan, D., Jessee, J. and Bloom, F.R. "Plasmid transformation of *Escherichia coli* and other bacteria." Methods in Enzymology 204: 63-113. 1991

Hara, M., Yazaki, K., Tanaka, S. and Tabata, M. "*S*-adenosyl-L-methionine: norcoclaurine 6-*O*-methyltransferase from *Thalictrum* minus cell cultures." Phytochemistry 38: 1131-1135. 1995

Harmer, S.L., Hogenesch, J.B., Straume, M., Chang, H.S., Han, B, Zhu, T., Wang, X., Kreps, J.A. and Kay, S.A. "Orchestrated transcription of key pathways in *Arabidopsis* by the circadian clock." Science 290: 2110-2113. 2000

Hartwig, U.A., Joseph, C.M. and Phillips, D.A. "Flavonoids released naturally from alfalfa seeds enhance growth rate of *Rhizobium meliloti*." Plant Physiology 95 (3): 797-803. 1991

Hashimoto, T. and Yamada, Y. "Alkaloid biogenesis: Molecular aspects." Annual Review of Plant Physiology and Plant Molecular Biology 45: 257-285. 1994

Hashimoto, T. and Yamada, Y. "New genes in alkaloid metabolism and transport." Current Opinion in Biotechnology 14: 163-168. 2003

Hauser, M.T. and Wink, M. "Uptake of alkaloids by latex vesicles and isolated mesophyll vacuoles of *Chelidonium majus* (*Papaveraceae*)." Zeitschrift fur Naturforschung 45: 949-957. 1990

Hauschild, K., Pauli, H.H. and Kutchan, T.M. "Isolation and analysis of a gene *bbe1* encoding the berberine bridge enzyme from the California poppy *Eschscholzia californica*." Plant Molecular Biology 36: 473-478. 1998

He, X-Z, and Dixon, R.A. "Genetic manipulation of isoflavone 7-*O* methyltransferase enhances biosynthesis of 4'-*O*-methylated isoflavonoid phytoalexins and disease resistance in alfalfa." Plant Cell 12: 1689-1702. 2000

Heese-Peck, A. and Raikhel, N.T. "A glycoprotein modified with terminal *N*-acetylglucosamine and localized at the nuclear rim shows sequence similarity to aldose-1-epimerases." The Plant Cell 10: 599-612. 1998

Heidrich, R., Kurkdjian, A., Guern, J. and Flugge, U.I. "Comparative studies on the electrical properties of the H translocating ATPase and pyrophosphatase of the vacuolar/lysosomal compartment." EMBO Journal 8: 2835-2841. 1989

Helliwell, C. A., Wesley, S. V., Wielopolska, A. J. and Waterhouse, P. M. "High-throughput vectors for efficient gene silencing in plants." Functional Plant Biology 29: 1217-1225. 2002

Hemm, M. R.; Ruegger, M. O. and Chapple, C. "The *Arabidopsis Ref2* mutant is defective in the gene encoding *Cyp83a1* and shows both phenylpropanoid and glucosinolate phenotypes." Plant Cell 15(1): 179-194. 2003

Hibberd, J. M. and Quick, W. P."Characteristics of C4 photosynthesis in stems and petioles of C3 flowering plants." Nature 415 (6870): 451-454. 2002

Higgins, T.J.V., Zwar, J.A. and Jacobsen, J.V. "Gibberellic acid enhances the level of translatable mRNA for α-amylase in barley aleurone layers." Nature 260: 166-169. 1976

Hirata, K., Poeaknapo, C., Schmidt, J. and Zenk, M.H. "1,2-Dehydroreticuline synthase, the branch point enzyme opening the morphinan biosynthetic pathway." Phytochemistry 65: 1039-1046. 2004

Hodges, C.C. and Rapoport, H. "Enzymic reduction of codeinone in vitro cell-free systems from *Papaver somniferum* and *P. bracteatum*." Phytochemistry19: 1681-1684. 1980

Homeyer, B.C. and Roberts, M.F. "Alkaloid sequestration by *Papaver somniferum* L. latex." Zeitschrift fur Naturforschung 39c: 876-881. 1984

Hughes, J., Smith, T., Kosterlitz, H., Fothergill, L., Morgan, B. and Morris, H. " Identification of two related pentapeptides from the brain with potent opiate agonist activity.' Nature 258: 577-579. 1975

Hughes, E.H. and Shanks, J.V. "Metabolic engineering of plants for alkaloid production." Metabolic Engineering 4: 41-48. 2002

Huang, F.C. and Kutchan, T.M. "Distribution of morphinan and benzo[c]phenanthridine alkaloid gene transcript accumulation in *Papaver somniferum*." Phytochemistry 53: 555-564. 2000

Huang, S. and Weir, B.S. "Estimating the total number of alleles using a sample coverage method." Genetics 159: 1365-1373. 2001

Ihaka, R. and Gentleman, R. "A language for data analysis and graphics." Journal of Computational Graph. Statistics 5; 299-314. 1996

Ikonomovski, K. "Colorimetric determination of thebaine in *Papaver bracteatum*." Journal of Pharmaceutical Sciences 70 (1): 102-103. 1981

Ilieva, S., Krusheva, R., (Krousheva, R.,) and Mateeva, D. "The morphological nature of some modifications of the flowers of opium poppy." Rasteniev"dni Nauki 14 (9): 60-66. 1977

International Narcotics Control Board. Annual Report 2003.

Ivanova, R. M. "The use of experimental mutagenesis in breeding the oil poppy." Genetika 8 (1): 30-37. 1972

Janssen, S. "A guide to the Lhc genes and their relatives in *Arabidopsis*." Trends in Plant Science 4(6): 236-240. 1999

Jones, P.G., Allaway, D., Gilmour, D.M., Harris, C., Rankin, D., Retzel, E.R. and Jones, C. "Gene discovery and microarray analysis of cacao (*Theobroma cacao* L.) varieties." Planta 216: 255-264. 2002

Jung, S.H., Lee, J.Y. and Lee, D.H. "Use of SAGE technology to reveal gene expression in *Arabidopsis* leaves undergoing cold stress." Plant Molecular Biology 52 (3): 553-567.2003

Kadereit, J. W. and Sytsma, K. J. "Disassembling *Papaver*: a restriction site analysis of chloroplast DNA." Nordic Journal of Botany 12 (2): 205-217. 1992

Kametani, T., Nemoto, H., Kobari, T. and Takano, S. "Enzymic phenol oxidation of *N*-methylcoclaurine. Studies on the syntheses of heterocyclic compounds. Part CCCXXXII." Journal of Heterocyclic Chemistry 7: 181-186. 1970.

Kametani, T., Mizushima, M., Takano, S. and Fukumoto, K. "Studies on the syntheses of heterocyclic compounds – DXXIII. Enzymic oxidation of homoorientaline with homogenised potato peelings." Tetrahedron: 29: 2031-2033. 1973

Kermekchiev, M. and Ivanova, L. "Ribin, a protein encoded by a message complementary to rRNA, modulates ribosomal transcription and cell proliferation." Molecular and Cellular Biology (24) 21: 8255-8263.

Khanna, K. R. and Singh, U.P., "Correlation studies in *Papaver somniferum* L. and their bearing on yield improvement." Planta Medica 28: 92-96. 1975

Kodaira, H. and Spector, S. "Transformation of thebaine to oripavine, codeine, and morphine by rat liver, kidney, and brain microsomes." Proceedings of the National Academy of Sciences USA 85: 1267-1271. 1988

Kodaira, H., Lisek, C.A., Jardine, I., Arimura, A. and Spector, S. "Identification of the convulsant opiate thebaine in mammalian brain." Proceedings of the National Academy of Sciences USA 86: 716-719. 1989

Kosterlitz, H.W. "Biosynthesis of morphine in the animal kingdom." Nature 330: 606. 1987

Kriegler, M. In: Gene Transfer And Expression: A Laboratory Manual. Stockton Press New York, NY. 1990

Kuhar, M.J., Pert, C.B. and Snyder, S.H. "Regional distribution of opiate receptor binding in monkey and human brain." Nature 245: 447-450. 1973

Kumar, A., Harrison, P.M., Cheung, K-H., Lan, N., Echols, N., Bertone, P., Miller, P., Gerstein, M.B. and Snyder, M. "An integrated approach for finding overlooked genes in yeast." Nature Biotechnology 20: 58-63. 2002

Kumar, S., Tamura, K. and Nei, M. "MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment." Briefings in Bioinformatics 5 (2): 150-163. 2004

Küppers, F.J.E.M., Salemink, C.A., Bastart, M. and Paris, M. "Alkaloids of *Papaver bracteatum*: Presence of codeine, neopinone and alpinine." Phytochemistry 15: 444-445. 1976

Kutchan, T.M. "Alkaloid biosynthesis-The basis for metabolic engineering in medicinal plants." The Plant Cell 7: 1059-1070. 1995

Kutchan, T.M. "Heterologous expression of alkaloid biosynthetic genes- a review." Gene 179: 73-81. 1996

Kutchan, T.M. "Molecular genetics of plant alkaloid biosynthesis." The Alkaloids 50: 257-316. 1998

Kutchan, T.M. "Ecological arsenal and development and developmental dispatcher. The paradigm of secondary metabolism." Plant Physiology 125: 58-60. 2001

Kutchan, T.M. and Dittrich, H. "Characterization and mechanism of the berberine bridge enzyme, a covalently flavinylated oxidase of benzophenanthridine alkaloid biosynthesis in plants." The Journal of Biological Chemistry 270 (41): 24475-24481. 1995

Kutchan, T.M. and Zenk, M.H. "Enzymology and molecular biology of benzophenanthridine alkaloid biosynthesis." Journal of Plant Research Special Issue 3: 165-173. 1993

Kuznetsov, A. I. "A genetic and physiological study of mutant opium poppy plants." Stabil'nost' i izmenchivost' genoma." 81-89. 1985

Lalezari, I., Nasseri, P. and Ashgarian, R. "*Papaver bracteatum* Lindl.: Population *Arya II*." Journal Pharmacological Sciences 63: 1331. 1974

Lavie, D., Berger-Josephs, H., Yehezkel, T., Gottlieb, H.E. and Levy, E.C. "Alpinigenine from *Papaver Bracteatum* Lindl. Restricted rotation in an unusual oxidation product." Journal Chemical Society Perkin I.: 1019-1022. 1981

Lee, J. and Pederson, H. "Stable genetic transformation of *Eschscholzia californica* expressing synthetic green fluorescent proteins." Biotechnology Progress 17: 247-251. 2001

Leigh, R.A. and Walker, R.R. "ATPase and acid phosphatase activities associated with vacuoles isolated from storage roots of red beet." Planta 150: 222-229. 1980

Lenz, R. and Zenk, M.H. "Closure of the oxide bridge in morphine biosynthesis." Tetrahedron Letters 35, No. 23: 3897-3900. 1994

Lenz, R. and Zenk, M.H. "Acetyl Coenzyme A: Salutaridinol-7-*O*-Acetyltransferase from *Papaver somniferum* plant cell cultures." The Journal of Biological Chemistry 270 (52): 31091-31096, 1995a.

Lenz, R. and Zenk, M.H. "Stereoselective reduction of codeinone, the penultimate enzymic step during morphine biosynthesis in *Papaver somniferum*." Tetrahedron Letters 36; 2449-2452. 1995b

Lenz, R. and Zenk, M.H. "Purification and properties of codeinone reductase (NADPH) from *Papaver somniferum* cell cultures and differentiated plants." European Journal Biochemistry 233: 132-139. 1995c.

Levy, A. and Milo, J. "Genetics and breeding of Papaver somniferum Poppy." In: Poppy: The Genus Papaver. : xiv. Ed: Bernath, J.Harwood Academic Publishers; Amsterdam; Netherlands. 1998. Pages 93-104.

Loeffler, S., Stadler, R., Nagakura, N. and Zenk, M.H. "Norcoclaurine as biosynthetic precursor of thebaine and morphine." Journal of the Chemical Societies Chemical Communications 15: 1160-1162. 1987

Luan, S. "Protein phosphatases in plants." Annual Review of Plant Biology 54 (1): 63-92. 2003

Lüttge, U., Smith, J. A. C., Marigo, G. and Osmond, C.B. "Energetics of malate accumulation in the vacuoles of *Kalanchoe tubiflora* cells." FEBS letters 126: 81-84. 1981

Maldonado-Mendoza, I.E., López-Meyer, M., Galef, J,R., Burnett, R.J. and Nessler, C.L. "Molecular analysis of a new member of the opium poppy tyrosine/3,4-dihydroxyphenylalanine decarboxylase gene family." Plant Physiology 110: 43-49.

1996

Marques, I.A. and Brodelius, P.E. "Elicitor induced L-tyrosine decarboxylase from plant cell suspension cultures." Plant Physiology 88: 46-51. 1988

Marrs, K. A., Alfineto, M. R., Lloyd, A.M. and Walbot, V. "A glutathione *S*-transferase involved in vacuolar transfer encoded by the maize gene Bronze-2." Nature 375: 397-400. 1995

Matile, P. "Localisation of alkaloids and mechanism of their accumulation in vacuoles of *Chelidonium majus* laticifers." Nova Acta Leopold. Supplement 7: 137-156. 1976

Matile, P. "Das toxische kompartiment der pflanzenelle." Naturwissenschaften 71: 18-24. 1984

Meijerink, W.H.J.H.J., Molina, P.E. and Abumrad, N.N. "Mammalian opiate alkaloid synthesis: lessons derived from plant biochemistry." Shock 12, 3: 165-173. 1999

Mele, G., Ori, N., Sato, Y. and Hake, S. "The *knotted1*-like homebox gene *BREVIPEDICELLUS* regulates cell differentiation by modulating metabolic pathways." Genes & Development 17: 2088-2093. 2003

Memelink, J., Verpoorte, R. and Kijne, J.W. "ORCAnization of jasmonate-responsive gene expression in alkaloid metabolism." Trends in Plant Science 6 (5): 212-219. 2001

Mende, P. and Wink, M. "Uptake of the quinolizidine alkaloid lupanine by protoplasts and vacuoles of *Lupinus polyphyllus* cell suspension cultures." Journal Plant Physiology 129: 229-242. 1987

Menke, F.L.H., Parchmann, S., Mueller, M.J., Kijne, J.W. and Memelink, J. "Involvement of the octadecanoid pathway and protein phosphorylation in fungal elicitor-induced expression of terpenoid indole alkaloid biosynthetic genes in *Catharanthus roseus*." Plant Physiology 199: 1289-1296. 1999a

Menke, F.L.H., Champion, A., Kijne, J.W. and Memelink, J. "A novel jasmonate- and elicitor-responsive element in the periwinkle secondary metabolic biosynthetic gene Str interacts with a jasmonate- and elicitor-inducible AP2-domain transcription factor, ORCA2." The EMBO Journal 18: 4455-4463. 1999b

Meshulam, H. and Lavie, D. "The alkaloidal constituents of *Papaver bracteatum* Arya II." Phytochemistry 19 (12): 2633-2635. 1980

Mikus, G., Somogyi, A.A., Bochner, F. and Eichelbaum, M. "Thebaine *O*-demethylation to oripavine - genetic-differences between 2 rat strains." Xenobiotica 21: 1501-1509. 1991

Millgate, A.G., Pogson, B.J., Wilson, I.W., Kutchan, T.M., Zenk, M.H., Gerlach, W.L., Fist, A.J. and Larkin, P.J. "Morphine-pathway block in *top1* poppies." Nature 431: 413-414. 2004

Milo, J., Levy, A., Palevitch, D. and Ladizinsky, G. "Thebaine content and yield in induced tetraploid and triploid plants of *Papaver bracteatum* Lindl." Euphytica 36 (2): 361-367. 1987

Montgomery, C.T., Cassels, B.K. and Shamma, M. "The rhoeadine alkaloids." Journal of Natural Products 46 (4): 441-453. 1983

Morishige, T., Tsujita, T., Yamada, Y. and Sato, F. "Molecular characterization of the *S*-adenosyl-L-methionine: 3'-hydroxy-*N*-methylcoclaurine 4'-*O*-methyltransferase of isoquinoline alkaloid biosynthesis in *Coptis japonica*." Journal of Biological Chemistry 275: 23398-23405. 2000

Mueller, U.G. and Wolfenbarger, L.L. "AFLP genotyping and fingerprinting." Tree 14 (10): 389-394. 1999

Myeonghyeon, W. and Hakik, R. "Improved Technique for Isolating RNA From Tobacco Tissues." Plant Molecular Biology Reporter 19, 187a-187f (2001)

Nesi, N., Jond, C., Debeaujon, I., Caboche, M. and Lepiniec, L. "The Arabidopsis *TT2* gene encodes an R2R3 MYB domain protein that acts as a key determinant for proanthocyanidin accumulation in developing seed." The Plant Cell 13: 2099-2114. 2001

Nesi, N., Debeaujon, I., Jond, C., Stewart, A.J., Jenkins, G.I., Caboche, M. and Lepiniec, L. "The *TRANSPARENT TESTA16* locus encodes the ARABIDOPSIS BSISTER MADS domain protein and is required for proper development and pigmentation of the seed coat." The Plant Cell 14: 2463-2479. 2002

Nessler, C.L., Allen, R.D. and Galewsky, S. "Identification and characterization of latex-specific proteins in opium poppy." Plant Physiology 79: 499-504. 1985

Nessler, C.L. and Mahlberg, P.G. "Ultrastructure of laticifers in redifferentiated organs on callus from *Papaver somniferum* (Papaveraceae)." Canadian Journal of Botany 57 (6): 675-685. 1979

Neumann, D., Krauss, G., Kiete, M. and Groger, D. "Indole alkaloid formation and storage in cell suspension cultures of *Catharanthus roseus*." Planta Medica 48: 20-23. 1983

Neuwinger, H.D. "Alkaloids in arrow poisons". In: Alkaloids: Biochemistry, Ecology, And Medicinal Applications, Eds: Roberts M.F. and Wink M. Plenum Press, New York. 1998. Pages 45-84

Nevkrytaya, N. V. Andreev, V. S., Belyaeva, R. G. and Nikiforov, Yu. L. "Study of pistilloid mutant plants of poppy." Botanicheskii Zhurnal 66 (6): 826-833. 1981a

Nevkrytaya, N. V. Andreev, V. S., Belyaeva, R. G. and Nikiforov, Yu. "Flower developmemt in mutant pistilloid poppy plants. II. Development and structure of the pistillodes." Ontogenez 12 (1): 79-86. 1981b

Nevkrytaya, N. V. Andreev, V. S., Belyaeva, R. G. and Nikiforov, Yu. L. "Flower development in mutant pistilloid poppy plants. I. Characteristics of morphogenesis." Ontogenez 12 (1): 72-78. 1981c

Nigam, K.B., Kandalkar, V.S. and Dhumale, D.B. "Induced mutations in opium poppy (*Papaver somniferum*)." Indian Journal of Agricultural Science 60 (4): 267-268. 1990

Noller, H.F., Hoffarth V. and Zimniak, L. "Unusual resistance of peptidyl transferase to protein extraction procedures." Science 256: 1416-1419. 1992.

Norris, S.R., Barrette, T.R. and Della Penna, D. "Genetic dissection of carotenoid synthesis in *Arabidopsis* defines plastoquinone as an essential component of phytoene desaturation." Plant Cell 7: 2139-2149. 1995

Notredame, C., Higgins, D. and Heringa, J. "T-Coffee: A novel method for multiple sequence alignments." Journal of Molecular Biology 302: 205-217. 2000

Nyman, U. and Hall, O. "Some varieties of *Papaver somniferum* L. with changed morphinane alkaloid content." Hereditas 84: 69-76. 1976

Nyman, U. and Hansson, B. "Morphine content variation in *Papaver somniferum* L. as affected by the presence of some isoquinoline alkaloids." Hereditas. 88:17-26. 1978

Nyman, U. "Selection for high thebaine/low morphine content (cpv. Morph:The) in *Papaver somniferum* L." Hereditas 89: 43-50. 1978

Nyman, U. "Selection for high thebaine/low morphine content (cpv. Morph:The) in *Papaver somniferum* L. II. Studies of the I4 and I5 generations from a spontaneous mutant." Hereditas 93: 121-124. 1980

O'Brien, C.P., "Drug addiction and drug abuse." In: The Pharmacological Basis of Therapeutics, 10th ed., (edit. Hardman, J.G. and Limbird, L.E.), McGraw-Hill, New

York NY. 2001. Pages 621-642

Ogita, S., Uefuji, H., Yamaguchi, Y., Koizumi, N. and Sano, H. "RNA interference - Producing decaffeinated coffee plants." Nature 423 (6942): 823-823. 2003

Oka, K., Kantrowitz, J.D. and Spector, S. "Isolation of morphine from toad skin." Proceedings of the National Academy of Sciences USA 82: 1852-1854. 1985

Ounaroon, A., Decker, G., Schmidt, J., Lottspeich, F. and Kutchan, T.M. "(*R*, *S*)-reticuline 7-*O*-methyltransferase and (*R*, *S*)-norcoclaurine 6-*O*-methyltransferase of *Papaver somniferum*- cDNA cloning and characterization of methyl transfer enzymes of alkaloid biosynthesis in opium poppy." Plant Journal 36: 808-819. 2003

Panicot, M., Minguet, E.G., Ferrando, A., Alcazar, R., Blazquez, M.A., Carbonell, J., Altabella, T., Koncz, C. and Tiburcio, A.F. "A Polyamine Metabolon Involving Aminopropyl Transferase Complexes in Arabidopsis." The Plant Cell 14: 2539–2551. 2002

Park, S., Johnson, A.G., Penzes-Yost, C. and Facchini, P.J. "Analysis of promoters from tyrosine/dihydroxyphenylalanine decarboxylase and berberine bridge enzyme genes involved in benzylisoquinoline alkaloid biosynthesis in opium poppy." Plant Molecular Biology 40: 121-131. 1999

Park, S, and Facchini, P.J. "High-efficiency somatic embryogenesis and plant regeneration in California poppy, *Eschscholzia californica* Cham." Plant Cell Reports 19: 421-426. 1999

Park, S, and Facchini, P.J. "*Agrobacterium*-mediated transformation of opium poppy, *Papaver somniferum*, via shoot organogenesis." J. Plant Phys. 157:207-214. 2000a

Park, S, and Facchini, P.J. "*Agrobacterium rhizogenes*-mediated transformation of opium poppy, *Papaver somniferum* L. and California poppy, *Eschscholzia californica* Cham., root cultures." Journal of Experimental Botany 51: 1005-1016. 2000b

Park, S, and Facchini, P.J. "Somatic embryogenesis from embryogenic cell suspension cultures of California poppy, *Eschscholzia californica* Cham." In Vitro Cellular Developmental Biology-Plant 37: 35-39. 2001

Park, S., Yu, M. and Facchini, P.J. "Antisense RNA-mediated suppression of benzophenanthridine alkaloid biosynthesis in transgenic cell cultures of California poppy." Plant Physiology 128: 696-706. 2002

Park, S., Yu, M. and Facchini, P.J. "Modulation of berberine bridge enzyme levels in transgenic root cultures of California poppy alters the accumulation of benzophenanthridine alkaloids." Plant Molecular Biology 51: 153-164. 2003

Parker, H. I., Blaschke, G. and Rapoport, H. "Biosynthetic conversion of thebaine to codeine." American Chemical Society 94 (4): 1276-1282. 1972

Patra, N.K. and Chauhan, S.P. "Morphophysiology and genetics of induced mutants expressed in M1 generation in opium poppies." Journal of Heredity, 81 (5): 347-350. 1990

Patra, N. K., Chauhan, S. P., Ram, R. S. and Kandpal, A. "High alkaloid yielding big capsulated dominant mutant in opium poppy." Mutation Breeding Newsletter (No. 42): 10-11. 1996

Pauli, H.H. and Kutchan, T.M. "Molecular cloning and functional heterologous expression of two alleles encoding (*S*)-*N*-methylcoclaurine 3'-hydroxylase (CYP80B1), a new methyl jasmonate-inducible cytochrome P-450-dependent mono-oxygenase of benzylisoquinoline alkaloid biosynthesis." The Plant Journal 13 (6): 793-801. 1998

Pert, C.B. and Snyder, S.H. "Opiate receptor: Demonstration in nervous tissue." Science 179: 1011-1014. 1973.

Peters, N.K., Frost, J.W. and Long, S.R. "A plant flavone, luteolin, induces expression of *Rhizobium meliloti* nodulation genes." Science 233: 977-980. 1986

Pham, T.D.T. and Roberts, M.F. "Quantitative characterization of the contents of *Papaver somniferum* latex vacuoles." Phytochemical Analysis 2: 68-73. 1991

Prajapati, S., Bajpai, S., Gupta, M.M. and Kumar, S. "The floral androcarpel organ (*ACO*) mutation permits high alkaloid yields in opium poppy *Papaver somniferum*." Current Science 81 (8): 1109-1112. 2001

Preininger, V., Novák, J. and Santavy, F. "Isolation and chemistry of the alkaloids from plants of the *Papaveraceae*. LXXXI. *Glauca*, a new section of the genus *Papaver*." Planta Medica 41:119-123. 1981

Raskin, I., Ribnicky, D.M., Komarnytsky, S., Llic, N., Poulev, A., Borisjuk, N., Brinker, A., Moreno, D.A., Ripoll, C., Yakoby, N., O'Neal, J.M., Cornwell, T., Pastor, I. and Fridlender, B. "Plants and Human health in the twenty-first century." Trends in Biotechnology 20: 522-531. 2002

Rasmussen, O. F., Stummann, B. M. and Henningsen, K. W. "Nucleotide sequence of a 1.1 kb fragment of the pea chloroplast genome containing three tRNA genes,

one of which is located within an open reading frame of 91 codons." Nucleic Acids Research 12 (23): 9143-9153. 1984

Rea, P. and Sanders, D. "Tonoplast energizations: Two H-pumps, one membrane." Physiologia Plantarum 71: 131-141. 1987

Rea, P.A., Li, Z., Lu, Y., Drozdowicz, Y.M. and Martinoia, E. "From vacuolar GS-X pumps to multispecific ABC transporters." Annual Reviews Plant Physhiology and Plant Molecular Biology 49: 727-760. 1998.

Reinboth, S., Mollenhauer, B. and Reinbothe, C. "JIPs and RIPs: the regulation of plant gene expression by jasmonates in response to environmental cues and pathogens." Plant Cell 6 (9): 1197-1209. 1994

Rice, K.C. "The development of a practical total synthesis of natural and unnatural codeine, morphine and thebaine." In: Phillipson et. al., Eds: The chemistry and biology of Isoquinoline Alkaloids. Berlin-Heidelberg: Springer-Verlag. 1985. Pages 191-203

Roberts, M.F., McCarthy, D., Kutchan, T.M. and Coscia, C.J. "Localization of enzymes and alkaloidal metabolites in *Papaver* Latex." Archives of Biochemistry and Biophysics 222 (2): 599-609. 1983

Roberts, M.F. "*Papaver* latex and alkaloid storage vacuoles." In: Plant Vacuoles and Their Importance in Plant Cell Compartmentations and Their Applications in Biotechnology. Nato ASI series, Plenum Press, New York: 513-528. 1987

Roberts, M.F., Homeyer, B.C. and Pham, T.D.T. "Further studies of sequestration of alkaloids in *Papaver somniferum* latex vacuoles." Zeitschrift fur Naturforschung 46c: 377-388. 1991

Roberts, M.F. and Wink, M. In "Alkaloids: Biochemistry, Ecology, and Medicinal Applications". Eds: Roberts, M. F. and Wink, M. Plenum Press, New York. 1998. Pages 1-7.

Roos, W., Evers, S., Hieke, M., Tschöpe, M. and Schumann, B. "Shifts of intracellular pH distribution as a part of the signal mechanism leading to the elicitation of benzophenanthridine alkaloids." Plant Physiology 118: 349-364. 1998

Rosco, A., Pauli, H.H., Priesner, W. and Kutchan, T.M. "Cloning and heterologous expression of NADPH-Cytochrome P450 reductases from the *Papaveraceae*." Archives of Biochemistry and Biophysics 348: 369-377. 1997

Rueffer, M., El-Shagi, H., Nagahura, N. and Zenk, M.H. "(*S*)-Norlaudanosoline synthase: the first enzyme in the benzylisoquinoline biosynthetic pathway." FEBS Letters 129: 5-9. 1981

Rueffer, M. and Zenk, M.H. "Enzymatic formation of protopines by a microsomal cytochrome P-450 system of *Corydalis vaginans*." Tetrahedron Letters 28: 5307-5310. 1987

Rueffer, M., Nagakura, N. and Zenk, M.H. "Partial purification and properties of *S*-adenosylmethionine: (*R*), (*S*)-norlaudanosiline-6-*O*-methyltransferase from *Argemone platyceras* cell cultures." Planta Medica 49 (3): 131-137. 1983

Samanani, N. and Facchini, P.J. "Isolation and partial characterization of norcoclaurine synthase, the first committed step in benzylisoquinoline alkaloid biosynthesis, from opium poppy." Planta 213: 898-906. 2001

Samanani, N. and Facchini, P.J. "Purification and characterization of norcoclaurine synthase." The Journal of Biological Chemistry 277 (37): 33878-33883. 2002

Samanani, N., Yeung, E.C. and Facchini, P.J. "Cell type-specific protoberberine alkaloid accumulation in *Thalictrum flavum*." Journal of Plant Physiology 159: 1189-1196. 2002

Santavy, F. In: "The Alkaloids," Vol 12, Ed: Manske, R.H.F. Academic Press, New York 1970. Pages 333-454

Santavy, F. In: "The Alkaloids," Vol 17, Ed: Manske, R.H.F. Academic Press, New York 1979. Pages 385-544

Sato, F., Tsujita, T., Katagiri, Y., Yoshida, S. and Yamada, Y. "Purification and characterization of *S*-adenosyl-L-methionine: norcoclaurine 6-*O*-methyltransferase from cultured *Coptis japonica* cells." European Journal of Biochemistry 225: 125-131. 1994

Schena, M. Shalon, D., Davis R.W. and Brown, P.O. "Quantitative monitoring of gene expression patterns with a complementary DNA microarray." Science 270, 5235: 467-470. 1995

Schenk, P.M., Kazan, K., Wilson, I. Anderson, J.P., Richmond, T., Somerville, S.C. and Manners, J.M. "Coordinated plant defense responses in *Arabidopsis* revealed by microarray analysis." Proceedings of the National Academy of Sciences USA 97 (21): 11655-60. 2000

Schiff, P.J. Jr. "Opium and its alkaloids." American Journal of Pharmaceutical Education 66: 186-194. 2002

Schmeller, T. and Wink, M. "Utilization of Alkaloids in Modern Medicine." Chapter 18. In: Alkaloids- Biochemistry, Ecology, and Medicinal Applications. Eds: Roberts,

M. F. and Wink, M. Plenum Press, New York: 1998. Pages 435-459.

Schultz, C.J. and Coruzzi, G.M. "The aspartate aminotransferase gene family of *Arabidopsis* encodes isoenzymes localized to three distinct subcellular compartments." The Plant Journal 7 (1): 61-75. 1995

Schultz, J., Milpetz, F., Bork, P. and Ponting,C.P. "SMART, a simple modular architecture research tool: Identification of signaling domains." Proceedings of the National Academy of Sciences USA 95: 5857-5864. 1998.

Schumacher, H.M., Gundlach, H., Fiedler, F. and Zenk, M.H. "Elicitation of benzophenanthridine alkaloid synthesis in *Eschscholzia* cell cultures." Plant Cell Reports 6: 410-413. 1987

Seki, M., Narusaka, M., Abe, H., Kasuga, M., Yamagachi-Shinozaki, K., Carninci, P., Hayashizaki, Y. and Shinozaki, K. "Monitoring the expression pattern of 1300 *Arabidopsis* genes under drought and cold stresses by using full-length cDNA microarray." The Plant Cell 13: 61-72. 2001

Sembdner, G. and Parthier, B. "The biochemistry and the physiological and molecular actions of jasmonates." Annual Review of Plant Physiology and Plant Molecular Biology 44: 569-589. 1993

Shamma, M. and Töke, L. "Convenient route to the rhoeadanes." Journal of Chemical Society, Chemical Communications: 740-741. 1973

Shamma, M. and Töke, L. "A new route to the benzazapines and a general synthesis of rhoeadanes." Tetrahedron 31: 1991-1996. 1975

Shamma, M. and Moniot, J.L. "Isoquinoline Alkaloids Research, 1972-1977," Plenum Press, New York. 1978. Pages 337-353

Sharma, J.R. and Singh, O.P. "Genetics and genetic improvement." In: The Opium Poppy. Medicinal and Aromatic Plants Series1. Eds.: Akhtan-Husain and Sharma, J.R. CIMAP, Lucknow, India, 39-68. 1983

Sharma, J.R., Lal, R.K., Gupta, P., Misra, H.O., Pant, V., Singh, N.K. and Pandey, V. "Development of a non-narcotic (opiumless and alkaloid-free) opium poppy, *Papaver somniferum*." Plant Breeding 118: 449-452. 1999a

Sharma, J.R., Lal, R.K., Misra, H.O., Naqvi, A.A. and Patra, D.D. "Combating opium-linked global abuses and supplementing the production of edible seed and seed oil: A novel non-narcotic var. Sujata of opium poppy (*Papaver somniferum* L.) Current Science 77: 1584-1589. 1999b

Shen, Z-W., Fisinger, U., Poulev, A., Eisenreich, W., Werner, I., Pleiner, E., Bacher, A. and Zenk, M.H. "Tracer studies with 13C-labeled carbohydrates in cultured plant cells. Retrobiosynthetic analysis of chelidonic acid biosynthesis." Phytochemistry 57: 33-42. 2001

Shirley, B., Kubasek, W.L., Storz, G., Bruggemannn, E., Koorneef, M., Ausubel, F.M. and Goodman, H.M. "Analysis of *Arabidopsis* mutants deficient in flavonoid biosynthesis." The Plant Journal 8: 659-671. 1995

Simon, E.J., Hiller, J.M. and Edelman, I. "Stereo-specific binding of the potent narcotic analgesic [^{3}H] etorphine to rat-brain homogenate." Proceedings of the National Academy of Sciences USA 70: 1947-1949. 1973

Single Convention on Narcotic Drugs. The "1961 Records", for the Official Records of the United Nations Conference for the Adoption of a Single Convention on Narcotic Drugs, documents É/CONF.34/24, and E/CONF.34/24/Add.1, reproduced respectively in United Nations publications, Sales Nos. 63.XI.4 and 63.XI.5.

Smith, N.A., Singh, S.P., Wang, M., Stoutjesdijk, P.A., Green, A.G. and Waterhouse, P.M. "Total silencing by intron-spliced hairpin RNAs." Nature 407: 319-320. 2000

Somerville, C. and Somerville, S. "Plant functional genomics." Science 285: 380-383. 1999

Spector, S., Kantrowitz, J.D. and Oka, K. "Presence of endogenous morphine in toad skin." Progress in Clinical Biological Research. 192: 329-332. 1985

St-Pierre, B., Vazquez-Flota, F.A. and De Luca, V. "Multicellular Compartmentation of *Catharanthus Roseus* Alkaloid Biosynthesis Predicts Intercellular Translocation of a Pathway Intermediate." Plant Cell 11: 887-900. 1999

Stadler, R., Kutchan, T.M., Loeffler, S., Nagakura, N., Cassels, B. and Zenk, M.H. "Revision of the early steps of reticuline biosynthesis." Tetrahedron Letters 28: 1251-54. 1987

Stadler, R., Kutchan, T.M. and Zenk, M.H. "Norcoclaurine is the central intermediate in benzylisoquinoline alkaloid biosynthesis." Phytochemistry 28: 1083-86. 1989

Strassner, J., Schaller, F., Frick, U.B., Howe, G.A., Weiler, E.W., Amrhein, N., Macheroux, P. and Schaller, A. " Characterization and cDNA-microarray expression analysis of 12-oxophytodienoate reductases reveals differential roles for octadecanoid biosynthesis in the local versus the systemic wound response." The Plant Journal 32: 585-601. 2002

Sugiura, M., Sakaki, T., Yabusaki, Y. and Ohkawa, H. "Cloning and expression in *Escherichia coli* and *Saccharomyces cerevisiae* of a novel tobacco cytochrome *P*-450-like cDNA." Biochemica et Biophysics Acta 1308: 231-240. 1996

Sze, H. "H-translocating ATPase: Advances using membrane vesicles." Annual Review Plant Physiology 36: 175-208. 1985

Taiz, L. "The plant vacuole." Journal Experimental Biology 172: 113-122. 1992

Tamai, I. and Safa, A.R. "Azidopine noncompetitively interacts with vinblastine and cyclosporin A binding to P-glycoprotein in multidrug resistant cells." Journal of Biological Chemistry 266:16796-16800. 1991

Terenius, L. "Characteristics of the receptor for narcotic analgesics in synoptic plasma membrane fraction from rat brain." Acta Pharmacol Toxicology 33: 377-384. 1973

Tétényi, P. "Opium poppy (*Papaver somniferum*): Botany and Horticulture". Horticultural Reviews 19: 373-410. 1997

The *Arabidopsis* Genome Initiative." Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*." Nature 408: 796-815. 2000

Theodoulou, F.L. "Plant ABC transporters." Biochemica et Biophysica Acta 1465: 79-103. 2000

Theuns, H.G., Theuns, H.L. and Lousberg, R.J.J.C. "Search for new natural sources of morphinans." Economic Botany 40 (4): 485-497. 1986

Thoden, J.B., Kim, J., Raushel, F.M. and Holden, H.M. "Structural and kinetic studies of sugar binding to galactose mutarotase from *Lactococcus lactis*." The Journal of Biological Chemistry 277 (47): 45458-45465. 2002

Thom, M. and Komor, E. "H-sugar antiport as the mechanism of sugar uptake by sugar cane vacuoles." FEBS Letters 173: 1-4. 1984

Thom, M. and Komor, E. "Electrogenic proton translocation by the ATPase of sugarcane vacuoles." Plant Physiology 77: 329-334. 1985

Thompson J.D., Higgins D.G. and Gibson T.J. "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. 1994

United Nations. Office on Drugs and Crime. World Drug Report: Chapter 2 "Trends" http://www.unodc.org/pdf/WDR_2004/volume_1.pdf. 2004

Unterlinner, B., Lenz, R. and Kutchan, T.M. "Molecular cloning and functional expression of codeinone reductase: the penultimate enzyme in morphine biosynthesis in the opium poppy *Papaver somniferum*." The Plant Journal 18 (4): 1-11. 1999

Uprety, H., Bhakuni, D.S. and Kapil, R.S. "Biosynthesis of papaverine." Phytochemistry 14 (7): 1535-1537. 1975

Van Der Fits, L. and Memelink, J. "ORCA3, a jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism. Science 289, 295-297. 2000

Van Der Fits, L. and Memelink, J. "The jasmonate-inducible AP2/ERF-domain transcription factor ORCA3 activates gene expression via interaction with a jasmonate-responsive promoter element." The Plant Journal 25: 43-53. 2001

Verpoorte, R., van der Heijden, R. and Memelink, J. "Engineering the plant cell factory for secondary metabolite production." Transgenic Research 9: 323-343. 2000

Vincent, P.G., Bare, C.E. and Gentner, W.A. "Rapid semi-quantitative spot test for determination of thebaine and differentiation of *Papaver bracteatum* from *P. orientale* and *P. pseudo-orientale*." Lloydia 39 (6): 76-78. 1976

Vom Endt, D., Kijne, J.W. and Memelink, J. "Transcription factors controlling plant secondary metabolism: what regulates the regulators?" Phytochemistry 61: 107-114. 2002

Vos, P., Vos, P., Hogers, R., Bleeker, M., Reijans, M., Lee, T. van de, Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. "AFLP: a new technique for DNA fingerprinting." Nucleic Acids Research 23: 4407-4414. 1995

Vuoristo, J.T., Berrettini, W.H. and Ala-Kokko, L. "C18orf2, a novel, highly conserved intronless gene within intron 5 of the GNAL gene on chromosome 18p11." Cytogenetic Cell Genetics 93 (1-2): 19-22. 2001

Wagner, H., Bladt, S. and Zgainski, E.M. in: Plant Drug Analysis: A Thin Layer Chromatography Atlas. Berlin, Springer-Verlag. 1984. Pages 78-79

Wang, J., Zhang, H. and Goodman, H.M. "An Arabidopsis cab gene homologous to Cab-8 of tomato." Plant Physiology 104: 297. 1994

Wang, M. B. and Waterhouse, P. M. "Application of gene silencing in plants." Current Opinion in Plant Biology 5: 146-150. 2002

Wasternack, C. and Parthier, B. "Jasmonate-signalled plant gene expression." Trends in Plant Science 2: 302-307. 1997

Wat, C.K., Steffens, P. and Zenk, M.H. "Partial purification and characterisation of *S*-adenosyl-L-methionine: norreticuline *N*-methyltransferase from *Berberis* cell suspension cultures." Zeitschrift für Naturforschung 41c: 126-134. 1986

Waterhouse, P.M., Helliwell, C.A. "Exploring plant genomes by RNA-induced gene silencing." Nature Reviews Genetics 4 (1): 29-38. 2003

Weid, M., Ziegler, J. and Kutchan, T.M. "The roles of latex and the vascular bundle in morphine biosynthesis in the opium poppy, *Papaver somniferum*." Proceedings of the National Academy of Sciences 101 (38): 13957-13962. 2004

Weiss, D. "Regulation of flower pigmentation and growth: Multiple signalling pathways control anthocyanin synthesis in expanding petals." Physiologia Plantarum 110 (2): 152-157. 2000

Weitz, C.J., Lowney, L.I., Faull, K.F., Feistner, G., Goldstein, A. "Morphine and codeine from mammalian brain." Proceedings of the National Academy of Sciences USA 83: 9784-9788. 1986

Weitz, C.J., Faull, K.F., Goldstein, A. "Synthesis of the skeleton of the morphine molecule by mammalian liver." Nature 330: 674-677. 1987

Wesley, S.V., Helliwell, C.A., Smith, N.A., Wang, M.B., Rouse, D.T., Liu, Q., Gooding, P.S., Singh, S.P., Abbott, D., Stoutjesdijk, P.A., Robinson, S.P., Gleave, A.P., Green, A.G. and Waterhouse, P.M. "Construct design for efficient, effective and high-throughput gene silencing in plants." Plant Journal 27 (6): 581-590. 2001

Wilhelm, R. and Zenk, M.H. "Biotransformation of thebaine by cell cultures of *Papaver somniferum* and *Mahonia nervosa*." Phytochemistry 46: 701-708. 1997

Williams R.D. and Ellis, B.E. "Age and tissue distribution of alkaloids in *Papaver somniferum*." Phytochemistry, 28 (8): 2085-2088. 1989

Wilson, D.L., Buckley, M.J., Helliwell, C.A. and Wilson, I.W. "New normalization methods for cDNA microarray data." Bioinformatics 19 (11): 1325-1332. 2003

Wingender, R., Rohrig, H., Horicke, C., Wing, D. and Schell, J. "Differential regulation of soybean chalcone synthase genes in plant defence, symbiosis and upon environmental stimuli." Molecular and General Genetics 218, 2: 315-322. 1989

Wink, M., Hartmann, T. and Witte, L. "Biotransformation of cadaverine and potential biogenetic intermediates of lupanine biosynthesis by plant cell suspension cultures" Planta Medica 40 (1): 31-39. 1980

Wink, M. and Hartmann, T. "Production of quinolizidine alkaloids by photomixotrophic cell suspension cultures: biochemical and biogenetic aspects"

Plant Medica 40 (2): 149-155. 1980

Wink, M., Hartmann, T., Witte, L. and Rheinheimer, J "Interrelationship between quinolizidine alkaloid producing legumes and infesting insects: exploitation of the alkaloid-containing phloem sap of *Cytisus scoparius* by the broom aphid *Aphis cytisorum*." Zeitschrift fur Naturforschung 37c: 1081. 1982

Wink, M. and Hartmann, T. "Physiological and biochemical aspects of quinolizidine alkaloid formation in cell suspension cultures." Plant tissue culture 1982. Japanese Association for Plant Tissue Culture; Tokyo; Japan: Pages 333-334.

Wink, M. "Quinolizidine alkaloids: Biochemistry, metabolism, and function in plants and cell suspension cultures." Planta Medica 53 (6): 509-514. 1987

Wink, M. "A short history of alkaloids." Chapter 2. In: Alkaloids: Biochemistry, Ecology, And Medicinal Applications, Eds: Roberts M.F. and Wink M. Plenum Press, New York, 1998. Pages 11-44.

Wink, M. "Plant secondary metabolites from higher plants: biochemistry, function and biotechnology." In: Biochemistry of Plant Secondary Metabolism, Annual Plant Reviews, Ed. M. Wink, 2. Sheffield: Sheffield Academic. 1999. Pages 1-16.

Wink, M., Hofer, A., Bilfinger, M., Englert, E., Martin, M. and Schneider, D. "Geese and plant dietary allelochemicals- Food palatability and geophagy." Chemoecology 4: 93-107. 1993

Wink, M. "Phytochemical Diversity Of Secondary Metabolites." In: Encyclopaedia of Plant And Crop Science Marcel Dekker Inc. New York. 2004. Pages 915-919

Winkle-Shirley, B. "Flavonoid biosynthesis. A colourful model for genetics, biochemistry, cell biology and biotechnology." Plant Physiology 126, 485-493. 2001

Yang, Y.H., Dudoit, S., Luu. and Speed, T.P. "Normalization of cDNA Microarray Data." SPIE Biomedical Optics Symposium. San Jose, CA, USA. 2001

Yazaki, K., Shitan, N., Takamatsu, H., Ueda, K. and Sato, F. "A novel *Coptis japonica* multidrug-resistant protein preferentially expressed in the alkaloid-accumulating rhizome." Journal of Experimental Botany 52 (357): 877-879. 2001

Ye, K., Ke, Y., Keshava, N., Shanks, J., Kapp, J.A., Tekmal, R.R., Petros, J. and Joshi, H.C. "Opium alkaloid noscapine is an antitumor agent that arrests metaphase and induces apoptosis in dividing cells." Proceedings of the National Academy of Sciences USA 95: 1601-1606. 1998

Yu, M. and Facchini, P.J. "Molecular cloning and characterisation of a type III glutathione *S*-transferase from cell suspension cultures of opium poppy treated with a

fungal elicitor." Physiologia Plantarum 108: 101-109. 2000

Yu, G., Nishimura, M., Arawaka, S., Levitan, D., Zhang, L., Tandon, A., Song, Y.Q., Rogaeva, E., Chen, F., Kawarai, T., Supala, A., Levesque, L., Yu, H., Yang, D.S., Holmes, E., Milman, P., Liang, Y., Zhang, D.M., Xu, D.H., Sato, C., Rogaev, E., Smith, M., Janus, C., Zhang, Y., Aebersold, R., Farrer, L.S., Sorbi, S., Bruni, A., Fraser, P. and St George-Hyslop, P. "Nicastrin modulates presenilin-mediated notch/glp-1 signal transduction and betaAPP processing." Nature 407 (6800): 48-54. 2000

Yun, D-J., Hashimoto, T. and Yamada, Y. "Metabolic engineering of medicinal plants: Transgenic *Atropa belladonna* with an improved alkaloid composition." Proceedings of the National Academy of Sciences USA 89: 11799-11803. 1992

Zenk, M.H. "The formation of benzophenanthridine alkaloids." Pure and Applied Chemistry, 66: 2023-2028. 1994

Zenk, M.H. In: Organic Reactivity: Physical And Biological Aspects. Eds: Maskill, Howard: Golding, Bernard T.; Griffin, Roger J. Royal Society of Chemistry (Great Britain) Cambridge Perkin Division. 1995 pages 89-109.

Zhu, W., Ma, Yuliang, M., Cadet, P., Yu, D., Bilfinger, T.V., Bianchi, E. and Stefano, G.B. "Presence of reticuline in rat brain: a pathway for morphine biosynthesis." Molecular Brain Research 117: 83-90. 2003

Appendix 1

List of BLAST x values of the randomly sequenced 1056 clones from the *P. somniferum* library.

SEQUENCE ID	Homology	Accession Number	Blast_E Value
SI00232762	Coffea arabica mRNA for DCL protein	AJ420083	2.02E-24
SI00232763	Pyrus communis PVP3 mRNA for vacuolar proton-inorganic pyrophosphatase, complete cds	AB097115	1.94E-58
SI00232764	Unknown	No Hit	null
SI00232765	At1g23740/F5O8_27 [Arabidopsis thaliana] >gi|15724193|gb|AAL06488.1|AF411799_1 At1g23740/F5O8_27 [Arabidopsis thaliana] >gi|15220854|ref|NP_173786.1| putative auxin-induced protein; protein id: At1g23740.1, supported by cDNA: 155143., supported by cDNA: gi_15724192, supported by cDNA: gi_20334863 [	AAM16188	2.17E-85
SI00232766	Unknown	No Hit	null
SI00232767	N.sylvestris psaDa gene for PSI-D2	X60008	6.15E-58
SI00232768	D.stramonium mRNA for S-adenosylmethionine decarboxylase	Y07768	7.85E-21
SI00232769	ribosomal protein L12 [Nicotiana sylvestris=tobacco, leaf, mRNA Chloroplast, 766 nt]	S93166	5.41E-29
SI00232770	putative Sec31p [Oryza sativa (japonica cultivar-group)]	BAC16134	1.55E-72
SI00232771	L.esculentum mRNA for 33kDa precursor protein of oxygen-evolving complex	Z11999	1.97E-61
SI00232772	unknown protein [Arabidopsis thaliana] >gi|15220697|ref|NP_174313.1| hypothetical protein; protein id: At1g30220.1, supported by cDNA: gi_18377758 [Arabidopsis thaliana] >gi|12320850|gb|AAG50560.1|AC073506_2 hypothetical protein [Arabidopsis thaliana] >gi|25309027|pir||D86426 hypothetical protein F	AAL67029	8.57E-96
SI00232773	unknown protein [Arabidopsis thaliana] >gi|15220697|ref|NP_174313.1| hypothetical protein; protein id: At1g30220.1, supported by cDNA: gi_18377758 [Arabidopsis thaliana] >gi|12320850|gb|AAG50560.1|AC073506_2 hypothetical protein [Arabidopsis thaliana] >gi|25309027|pir||D86426 hypothetical protein F	AAL67029	1.65E-119
SI00232774	CLC-Nt2 protein [Nicotiana tabacum]	AAD29679	2.30E-112
SI00232775	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 68	AL161572	1.08E-11
SI00232776	N.tabacum mRNA for precursor of photosysten II 22 kDa protein	X84225	3.73E-35
SI00232777	L.esculentum mRNA for 10kDa polypeptide precursor of photosystem II	X95987	1.03E-15
SI00232778	Capsicum chinense bZIP transcription factor (PPI1) mRNA, complete cds	AF430372	1.97E-21
SI00232779	S.oleracea mRNA for 6.1 kDa polypeptide of photosystem II	X85038	3.15E-16
SI00232780	Oryza sativa chromosome 12, . BAC OJ1587_D05 of library Monsanto from chromosome 12 of cultivar Nipponbare of ssp. japonica of Oryza sativa (rice), complete sequence	AL713906	1.77E-12
SI00232781	Nicotiana tabacum beta-carbonic anhydrase (CA) mRNA, complete cds; nuclear gene for chloroplast product	AF454759	7.60E-18
SI00232782	Unknown protein [Arabidopsis thaliana] >gi|25344189|pir||B86304 hypothetical protein F6I1.12 [imported] - Arabidopsis thaliana >gi|9802776|gb|AAF99845.1|AC051629_12 Unknown protein [Arabidopsis thaliana] >gi|18394414|ref|NP_564010.1| expressed protein; protein id: At1g16880.1, supported by cDNA: gi	AAK62433	1.48E-51
SI00232783	Pisum sativum mRNA for Rieske iron-sulfur protein Tic55	AJ000520	5.08E-25
SI00232784	Unknown	No Hit	null

SI00232785	AT4g31540/F3L17_110 [Arabidopsis thaliana] >gi|28416523|gb|AAO42792.1| AT4g31540/F3L17_110 [Arabidopsis thaliana]	AAM98083	1.20E-39
SI00232786	adenosylmethionine decarboxylase (EC 4.1.1.50) - maize >gi|1532073|emb|CAA69075.1| S-adenosylmethionine decarboxylase [Zea mays] >gi|3913427|sp|O24575|DCAM_MAIZE S-adenosylmethionine decarboxylase proenzyme (AdoMetDC) (SamDC) [Contains: S-adenosylmethionine decarboxylase alpha chain; S-adenosylmeth	T03947	8.90E-89
SI00232787	Arabidopsis thaliana chromosome 5 CHR5v07142002 genomic sequence	NM_125800	5.19E-25
SI00232788	Nicotiana tabacum mRNA for hydroxycinnamoyl transferase (hct gene)	AJ507825	7.26E-50
SI00232789	Bruguiera gymnorhiza HPR mRNA for hydroxypyruvate reductase, complete cds	AB060810	0
SI00232790	Arabidopsis thaliana At4g04350/T19B17_7 mRNA, complete cds	BT002249	1.14E-13
SI00232791	putative protein; protein id: At5g15170.1, supported by cDNA: gi_17381097 [Arabidopsis thaliana] >gi|23297734|gb|AAN13014.1| unknown protein [Arabidopsis thaliana]	NP_197021	4.17E-41
SI00232792	ribulose-1,5-bisphosphate carboxylase/oxygenase activase 1 [Gossypium hirsutum]	AAG61120	5.29E-125
SI00232793	Oncorhynchus mykiss polyubiquitin mRNA, complete cds	AF361365	0
SI00232794	AT5g59750/mth12_150 [Arabidopsis thaliana] >gi|15215752|gb|AAK91421.1| AT5g59750/mth12_150 [Arabidopsis thaliana] >gi|9758840|dbj|BAB09512.1| GTP cyclohydrolase II; 3,4-dihydroxy-2-butanone-4-phoshate synthase [Arabidopsis thaliana] >gi|27363454|gb|AAO11646.1| At5g59750/mth12_150 [Arabidopsis thali	AAK32847	7.21E-104
SI00232795	N.tabacum mRNA for chloroplast Rieske FeS precursor protein 1	X66009	1.14E-59
SI00232796	Unknown	No Hit	null
SI00232797	Arabidopsis thaliana putative RNA-binding protein (At1g09340) mRNA, complete cds	AY070022	4.37E-50
SI00232798	Vitis vinifera mRNA for zinc finger protein (SINA1p)	Y18471	6.08E-77
SI00232799	Lotus japonicus genomic DNA, chromosome 1, clone:LjT06I17, TM0017, complete sequence	AP004484	5.12E-25
SI00232800	Arabidopsis thaliana chromosome 2 CHR2v07142002 genomic sequence	NM_129039	6.78E-21
SI00232801	Unknown	No Hit	null
SI00232802	Unknown	No Hit	null
SI00232803	Spinacia oleracea heat shock C70 protein mRNA, complete cds	L26243	0
SI00232804	putative villin [Arabidopsis thaliana] >gi|3093294|emb|CAA73320.1| putative villin [Arabidopsis thaliana] >gi|5730126|emb|CAB52460.1| putative villin [Arabidopsis thaliana] >gi|15234646|ref|NP_194745.1| putative villin; protein id: At4g30160.1 [Arabidopsis thaliana] >gi|7269916|emb|CAB81009.1| puta	BAC41968	0
SI00232805	P.sativum atpC mRNA for gamma subunit of ATP synthase	X63604	7.45E-52
SI00232806	Nicotiana sylvestris Lhcb1*5, Lhcb1*6 genes for light harvesting chlorophyll a/b-binding protein, complete cds	AB012638	5.76E-42
SI00232807	ankyrin-like protein [Arabidopsis thaliana] >gi|25408188|pir||E84725 ankyrin-like protein [imported] - Arabidopsis thaliana >gi|15225141|ref|NP_180741.1| ankyrin-like protein; protein id: At2g31820.1 [Arabidopsis thaliana]	AAD32290	9.19E-94
SI00232808	monooxygenase 1 [imported] - Arabidopsis thaliana >gi|3426062|emb|CAA07574.1| monooxygenase [Arabidopsis thaliana]	T51603	2.19E-71
SI00232809	Arabidopsis thaliana unknown protein (At4g02790) mRNA, complete cds	AY065311	1.17E-13
SI00232810	Unknown	No Hit	null
SI00232811	Nicotiana tabacum FtsH-like protein Pftf precursor (Pftf) mRNA, nuclear gene encoding chloroplast protein, complete cds	AF117339	1.27E-96
SI00232812	gamma-terpinene synthase [Citrus limon]	AAM53943	8.93E-32
SI00232813	Unknown	No Hit	null
SI00232814	Unknown	No Hit	null

SI00232815	Castor bean chloroplast beta-ketoacyl-ACP synthase (50 kDa synthase) mRNA, complete cds	L13242	2.59E-79
SI00232816	Unknown	No Hit	null
SI00232817	Unknown	No Hit	null
SI00232818	hypothetical protein F25O24.10 - Arabidopsis thaliana >gi|7269744|emb|CAB81477.1| putative protein [Arabidopsis thaliana] >gi|4972078|emb|CAB43903.1| putative protein [Arabidopsis thaliana] >gi|15235482|ref|NP_194618.1| putative protein; protein id: At4g28890.1 [Arabidopsis thaliana]	T08944	7.14E-17
SI00232819	Arabidopsis thaliana chromosome 1 CHR1v07142002 genomic sequence	NM_106605	6.52E-43
SI00232820	Pisum sativum mRNA for P protein, a part of glycine cleavage complex	X59773	0
SI00232821	Zantedeschia aethiopica glycolate oxidase (gox) mRNA, complete cds	AY173074	2.10E-104
SI00232822	Vernicia fordii negatively light-regulated protein mRNA, complete cds	AF061157	5.92E-21
SI00232823	Arabidopsis thaliana chromosome 1 CHR1v07142002 genomic sequence	NM_104698	9.60E-49
SI00232824	Barley mRNA for toxin alpha-hordothionin	X05901	3.27E-27
SI00232825	Vigna radiata mRNA for alpha-1,3-fucosyltransferase (FucT c3)	Y18529	3.18E-23
SI00232826	zinc transporter protein ZIP1 [Glycine max]	AAK37761	3.59E-51
SI00232827	Unknown	No Hit	null
SI00232828	Arabidopsis thaliana clone 157482 mRNA, complete sequence	AY085561	1.51E-37
SI00232829	Unknown	No Hit	null
SI00232830	Unknown	No Hit	null
SI00232831	Pisum sativum serine hydroxymethyltransferase mRNA, complete cds	M87649	1.46E-49
SI00232832	Lotus japonicus genomic DNA, chromosome 1, clone:LjT13O04, TM0016, complete sequence	AP004483	8.95E-33
SI00232833	hypothetical protein; protein id: At1g34420.1 [Arabidopsis thaliana] >gi|12323858|gb|AAG51899.1|AC023913_7 hypothetical protein; 24606-21623 [Arabidopsis thaliana] >gi|8778255|gb|AAF79264.1|AC023279_13 F12K21.25 [Arabidopsis thaliana]	NP_174702	3.08E-62
SI00232834	Unknown	No Hit	null
SI00232835	Vigna radiata chlorophyll a/b binding protein CP29 (CipCp29) mRNA, complete cds; nuclear gene for chloroplast product	AF139466	1.69E-46
SI00232836	CRS2 [Zea mays]	AAF27939	4.72E-90
SI00232837	Nicotiana tabacum beta-carbonic anhydrase (CA) mRNA, complete cds; nuclear gene for chloroplast product	AF454759	1.82E-15
SI00232838	harpin-induced protein-like [Arabidopsis thaliana] >gi|6041819|gb|AAF02134.1|AC009918_6 unknown protein [Arabidopsis thaliana] >gi|9502172|gb|AAF88021.1|AF264697_1 NDR1/HIN1-Like protein 1 [Arabidopsis thaliana] >gi|18399308|ref|NP_566396.1| VAMP protein SEC22; protein id: At3g11660.1, supported by	AAM66116	1.03E-61
SI00232839	Citrus maxima citrate synthase (cit) mRNA, nuclear gene encoding mitochondrial protein, complete cds	U19481	1.27E-88
SI00232840	Arabidopsis thaliana chromosome 1 CHR1v07142002 genomic sequence	NM_106192	5.54E-31
SI00232841	Unknown	No Hit	null
SI00232842	chloroplast protein 12 [Nicotiana tabacum] >gi|1617197|emb|CAA96569.1| CP12 [Nicotiana tabacum] >gi|7489109|pir||T02941 CP12 protein precursor, chloroplast - common tobacco	AAK49535	9.05E-29
SI00232843	Unknown Protein [Arabidopsis thaliana] >gi|25518632|pir||F86302 hypothetical protein F17F16.7 [imported] - Arabidopsis thaliana	AAG09086	4.19E-107
SI00232844	Unknown	No Hit	null
SI00232845	Arabidopsis thaliana unknown protein (At5g67370) mRNA, complete cds	AY051028	3.04E-23

SI00232846	Unknown	No Hit	null
SI00232847	Cucumis melo aminotransferase 1 mRNA, complete cds	AY066012	7.53E-58
SI00232848	Unknown	No Hit	null
SI00232849	Arabidopsis thaliana putative mitochondrial NAD-dependent malate dehydrogenase (At1g53240) mRNA, complete cds	AF339684	7.31E-55
SI00232850	G.hirsutum cab gene for chlorophyll ab binding protein	X54090	3.19E-97
SI00232851	aromatic rich glycoprotein, putative; protein id: At1g70370.1, supported by cDNA: gi_15912220, supported by cDNA: gi_1762583 [Arabidopsis thaliana] >gi|3176680|gb|AAC18803.1| Identical to polygalacuronase isoenzyme 1 beta subunit homolog mRNA gb|U63373. EST gb|AA404878 comes from this gene. [Arabi	NP_177194	2.96E-84
SI00232852	Unknown	No Hit	null
SI00232853	Cicer arietinum mRNA for ribosomal protein L24	AJ225027	9.90E-19
SI00232854	Petunia x hybrida aquaporin-like protein (PIP1;2) mRNA, complete cds	AF452011	3.81E-75
SI00232855	Cotton mRNA for cottonseed catalase subunit 1 (EC 1.11.1.6)	X52135	8.77E-73
SI00232856	unknown [Arabidopsis thaliana]	AAM65297	1.48E-59
SI00232857	L.esculentum psbX mRNA for photosystem II 23 kDa protein	X63007	7.09E-18
SI00232858	thylakoid lumenal 16.5 kDa protein, chloroplast precursor [Arabidopsis thaliana]	AAM66113	1.37E-41
SI00232859	putative ubiquitin activating enzyme [Arabidopsis thaliana] >gi|25408233|pir||G84732 probable ubiquitin activating enzyme [imported] - Arabidopsis thaliana >gi|15225241|ref|NP_180800.1| putative ubiquitin activating enzyme; protein id: At2g32410.1 [Arabidopsis thaliana]	AAC69937	5.56E-20
SI00232860	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	1.22E-31
SI00232861	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	7.56E-30
SI00232862	Lotus japonicus genomic DNA, chromosome 6, clone:LjT19B18, TM0139, complete sequence	AP004960	1.16E-13
SI00232863	Beta vulgaris mitochondrial genomic DNA, complete sequence, section 2/2	AP000397	2.07E-26
SI00232864	Arabidopsis thaliana glutamate:glyoxylate aminotransferase 2 (GGT2) mRNA, complete cds	AF479640	1.15E-84
SI00232865	P.sativum mRNA for fructose-1, 6-biphosphate aldolase (clone aldcyt1)	X89828	3.81E-35
SI00232866	Arabidopsis thaliana chromosome 1 CHR1v07142002 genomic sequence	NM_101628	2.07E-21
SI00232867	At1g23740/F5O8_27 [Arabidopsis thaliana] >gi|15724193|gb|AAL06488.1|AF411799_1 At1g23740/F5O8_27 [Arabidopsis thaliana] >gi|15220854|ref|NP_173786.1| putative auxin-induced protein; protein id: At1g23740.1, supported by cDNA: 155143., supported by cDNA: gi_15724192, supported by cDNA: gi_20334863 [	AAM16188	5.31E-65
SI00232868	Putative ribosomal protein L19 [Arabidopsis thaliana] >gi|15218602|ref|NP_171777.1| ribosomal protein L19, putative; protein id: At1g02780.1, supported by cDNA: 2906., supported by cDNA: gi_15983411, supported by cDNA: gi_18087584 [Arabidopsis thaliana] >gi|15983412|gb|AAL11574.1|AF424580_1 At1g027	AAF02889	5.83E-30
SI00232869	hypothetical protein; protein id: At1g54210.1, supported by cDNA: gi_19912168 [Arabidopsis thaliana] >gi|19912169|dbj|BAB88396.1| autophagy 12a [Arabidopsis thaliana] >gi|21636954|gb|AAM70187.1|AF492758_1 autophagy APG12 [Arabidopsis thaliana]	NP_175823	3.78E-40
SI00232870	Lycopersicon esculentum mRNA binding protein precursor, mRNA, nuclear gene encoding chloroplast protein, complete cds	AF106660	6.62E-80
SI00233790	putative FK506-binding protein [Oryza sativa]	AAK38500	3.32E-61
SI00233791	Unknown	No Hit	null
SI00232871	Spinacea oleracea mRNA for 20 kDa protein of CP24	Z25886	4.86E-53
SI00232872	Arabidopsis thaliana chromosome 5 CHR5v07142002 genomic sequence	NM_120696	9.15E-30
SI00232873	Unknown	No Hit	null

SI00232874	Arabidopsis thaliana chromosome 1 CHR1v07142002 genomic sequence	NM_104451	2.12E-39
SI00232875	Arabidopsis thaliana ACR4 (ACR4) mRNA, complete cds	AF528060	1.36E-25
SI00233792	Unknown	No Hit	null
SI00233793	Arabidopsis thaliana clone 25569 mRNA, complete sequence	AY086521	2.64E-16
SI00232876	Unknown	No Hit	null
SI00232877	Arabidopsis thaliana chromosome 3 CHR3v07142002 genomic sequence	NM_112302	5.59E-105
SI00232878	Oryza sativa genomic DNA, chromosome 4, BAC clone: OSJNBa0029H02, complete sequence	AL606594	2.15E-12
SI00232879	unknown protein [Arabidopsis thaliana] >gi|25372890|pir||E96760 hypothetical protein T9L24.40 [imported] - Arabidopsis thaliana >gi|11120790|gb|AAG30970.1|AC012396_6 hypothetical protein [Arabidopsis thaliana] >gi|15219459|ref|NP_177482.1| hypothetical protein; protein id: At1g73390.1, supported by	AAL61919	1.26E-64
SI00233794	light harvesting chlorophyll a/b-binding protein [Nicotiana sylvestris]	BAA25389	2.14E-46
SI00232880	Nicotiana tabacum mRNA for thioredoxin peroxidase	AJ309009	5.84E-77
SI00233795	wall-associated serine/threonine kinase (EC 2.7.1.-) 2 [imported] - Arabidopsis thaliana >gi|24417340|gb|AAN60280.1| unknown [Arabidopsis thaliana] >gi|4826399|emb|CAB42872.1| wall-associated kinase 2 [Arabidopsis thaliana]	T52588	1.29E-55
SI00232881	Arabidopsis thaliana chromosome 3 CHR3v07142002 genomic sequence	NM_115132	6.00E-37
SI00232882	Unknown	No Hit	null
SI00232883	thioredoxin m4 [Arabidopsis thaliana] >gi|15232567|ref|NP_188155.1| thioredoxin m4; protein id: At3g15360.1, supported by cDNA: 42151., supported by cDNA: gi_14030704, supported by cDNA: gi_16974518, supported by cDNA: gi_6539613 [Arabidopsis thaliana] >gi|21593734|gb|AAM65701.1| thioredoxin m4 [Ar	AAF35402	7.65E-47
SI00232884	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	1.84E-36
SI00232885	Unknown	No Hit	null
SI00232886	Unknown	No Hit	null
SI00232887	Hevea brasiliensis phosphomevalonate kinase mRNA, complete cds	AF429385	7.26E-15
SI00232888	Mesembryanthemum crystallinum water channel protein MipK (MipK) mRNA, complete cds	AF133532	4.23E-44
SI00233796	similar to hypothetical protein FLJ20511 [Homo sapiens] [Rattus norvegicus]	XP_226467	3.85E-38
SI00232889	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	6.13E-30
SI00232890	Arabidopsis thaliana clone RAFL15-09-E23 (R20450) At3g23790 mRNA, complete sequence	BT004051	4.79E-13
SI00232891	Arabidopsis thaliana chromosome 5 CHR5v07142002 genomic sequence	NM_121315	5.10E-56
SI00233797	Lotus japonicus genomic DNA, chromosome , clone:LjT35I07, TM0121b, complete sequence	AP004948	5.77E-49
SI00232892	DnaJ protein homolog - kidney bean (fragment) >gi|1684851|gb|AAB36543.1| DnaJ-like protein [Phaseolus vulgaris]	T11855	2.24E-32
SI00232893	S.nigra mRNA for metallothionein-like protein	X83439	4.69E-12
SI00232894	hypothetical protein [Spinacia oleracea] >gi|11497575|ref|NP_054982.1| hypothetical protein [Spinacia oleracea] >gi|11497590|ref|NP_054997.1| hypothetical protein [Spinacia oleracea] >gi|7636156|emb|CAB88778.1| hypothetical protein [Spinacia oleracea]	CAB88793	9.89E-25
SI00232895	Oryza sativa genomic DNA, chromosome 4, BAC clone: OSJNBa0029H02, complete sequence	AL606594	3.38E-17
SI00232896	AT3g15840 [Arabidopsis thaliana]	AAG40376	2.51E-23
SI00232897	extensin-like protein [Arabidopsis thaliana] >gi|15230120|ref|NP_189091.1| disease resistance protein, putative; protein id: At3g24480.1 [Arabidopsis thaliana]	BAB01951	1.81E-73

SI00232898	Unknown	No Hit	null
SI00232899	Unknown	No Hit	null
SI00232900	Arabidopsis thaliana chromosome 2 clone T1B8 map TEn5, complete sequence	U78721	1.60E-29
SI00232901	Cucurbita maxima Pugf mRNA for glutathione S-transferase, complete cds	AB059484	6.33E-56
SI00232902	Arabidopsis thaliana AT4g05180/C17L7_100 mRNA, complete cds	AF372897	5.01E-13
SI00232903	cell division related protein-like [Arabidopsis thaliana] >gi|15239928|ref|NP_196229.1| cell division related protein-like; protein id: At5g06110.1 [Arabidopsis thaliana]	BAA98200	1.00E-37
SI00232904	G.hirsutum cab gene for chlorophyll ab binding protein	X54090	6.58E-80
SI00232905	P.sativum GA mRNA (clone F)	X65155	5.78E-40
SI00232906	Crataegus crus-galli putative plastidic glutamine synthetase (GS2) mRNA, complete cds; nuclear gene for chloroplast product	AY162465	4.65E-87
SI00232907	Unknown	No Hit	null
SI00232908	T17H3.9 [Arabidopsis thaliana]	AAD45997	9.63E-65
SI00232909	Unknown	No Hit	null
SI00232910	hypothetical protein F24G16.180 - Arabidopsis thaliana >gi|20466798|gb|AAM20716.1| putative protein [Arabidopsis thaliana] >gi|15232209|ref|NP_191550.1| putative protein; protein id: At3g59910.1, supported by cDNA: gi_20466797 [Arabidopsis thaliana] >gi|7019685|emb|CAB75810.1| putative protein [Ara	T47815	9.71E-30
SI00232911	Arabidopsis thaliana clone RAFL05-19-M23 (R10494) unknown protein (At1g11480) mRNA, complete cds	BT000706	9.84E-22
SI00232912	putative protein; protein id: At3g62770.1, supported by cDNA: gi_14517419, supported by cDNA: gi_16323361 [Arabidopsis thaliana]	NP_567132	9.26E-68
SI00232913	Unknown	No Hit	null
SI00232914	Nicotiana tabacum beta-carbonic anhydrase (CA) mRNA, complete cds; nuclear gene for chloroplast product	AF454759	1.86E-15
SI00232915	glutamine synthetase GS2 [Beta vulgaris]	AAK07678	2.16E-63
SI00232916	unknown [Arabidopsis thaliana] >gi|28393243|gb|AAO42050.1| unknown protein [Arabidopsis thaliana] >gi|18420050|ref|NP_568382.1| Expressed protein; protein id: At5g19855.1, supported by cDNA: 42402. [Arabidopsis thaliana]	AAM65713	6.12E-51
SI00232917	unknown [Prunus armeniaca]	AAD38147	8.16E-80
SI00232918	Glycine max aspartokinase-homoserine dehydrogenase (AK-HSDH) mRNA, complete cds	AF049706	1.27E-28
SI00232919	Arabidopsis thaliana chromosome 4 CHR4v07142002 genomic sequence	NM_118534	5.73E-23
SI00232920	B.juncea mRNA for chlorophyll a/b-binding protein	X95727	4.79E-53
SI00232921	Arabidopsis thaliana sugar transporter - like protein (At5g18840/F17K4_90) mRNA, complete cds	BT000608	1.18E-16
SI00232922	Unknown	No Hit	null
SI00232923	Zea mays PCO092528 mRNA sequence	AY107881	3.18E-20
SI00232924	Arabidopsis thaliana unknown protein (At4g33410/F17M5_170) mRNA, complete cds	BT000484	4.05E-41
SI00233798	Unknown	No Hit	null
SI00232925	Pisum sativum chloroplast photosystem I 24 kDa light harvesting protein (lhca3) mRNA, complete cds	L19651	7.14E-49
SI00232926	nine-cis-epoxycarotenoid dioxygenase4 [Pisum sativum]	BAC10552	1.71E-85
SI00232927	Arabidopsis thaliana putative phenylalanine-tRNA synthetase (At3g58140) mRNA, complete cds	AY063063	1.25E-56
SI00232928	L.esculentum mRNA for 10kDa polypeptide precursor of photosystem II	X95987	4.25E-12
SI00232929	Tomato CAB-8 gene (constucted from mRNA and DNA) for type III chlorophyll a/b binding polypeptide of photosystem I	X15258	1.87E-52

SI00232930	Nicotiana paniculata mRNA for plastidic aldolase NPALDP1, complete cds	AB027001	5.56E-99
SI00232931	NTGP4 [Nicotiana tabacum]	AAD09518	2.16E-53
SI00232932	Unknown	No Hit	null
SI00232933	Arabidopsis thaliana chromosome 2 CHR2v07142002 genomic sequence	NM_130301	7.93E-45
SI00232934	Pisum sativum ubiquitin conjugating enzyme (UBC4), complete cds	L29077	3.39E-75
SI00232935	Arabidopsis thaliana unknown protein (At1g55620) mRNA, complete cds	AY080722	2.66E-82
SI00233799	Unknown	No Hit	null
SI00233800	Carica papaya ACC oxidase mRNA, complete cds	AF254125	7.42E-43
SI00233801	Grp94 [Xerophyta viscosa]	AAN34791	1.19E-19
SI00233802	Nicotiana tabacum beta-carbonic anhydrase (CA) mRNA, complete cds; nuclear gene for chloroplast product	AF454759	1.86E-15
SI00232936	P.sativum Cab-8 gene for photosystemm II chlorophyll a/b binding protein	X56538	4.40E-68
SI00233803	70 kDa heat shock protein [Arabidopsis thaliana] >gi|9294373|dbj|BAB02269.1| 70 kDa heat shock protein [Arabidopsis thaliana] >gi|15809846|gb|AAL06851.1| AT3g12580/T2E22_110 [Arabidopsis thaliana] >gi|15809832|gb|AAL06844.1| AT3g12580/T2E22_110 [Arabidopsis thaliana] >gi|12321973|gb|AAG51030.1|AC06	AAL24367	1.95E-42
SI00232937	Unknown	No Hit	null
SI00232938	Arabidopsis thaliana unknown protein (At5g65840) mRNA, complete cds	BT000333	2.87E-14
SI00232939	Unknown	No Hit	null
SI00232940	Arabidopsis thaliana chromosome 5 CHR5v07142002 genomic sequence	NM_124561	5.92E-77
SI00232941	Unknown	No Hit	null
SI00232942	Unknown	No Hit	null
SI00232943	hypothetical protein, similar to H. pilory nifS like protein [Oryza sativa]	AAF34417	5.54E-56
SI00232944	Arabidopsis thaliana DNA chromosome 3, BAC clone F2K15	AL132956	4.06E-11
SI00232945	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	8.06E-39
SI00232946	sucrose-phosphate synthase [Medicago sativa]	AAK09427	4.90E-17
SI00232947	Pisum sativum mRNA for precursor for 33-kDa protein of photosystem II, complete cds	D13297	7.92E-21
SI00232948	P.sativum Cab-8 gene for photosystemm II chlorophyll a/b binding protein	X56538	6.11E-74
SI00232949	Arabidopsis thaliana chromosome 2 CHR2v07142002 genomic sequence	NM_127659	1.82E-15
SI00232950	L.esculentum mRNA for 10kDa polypeptide precursor of photosystem II	X95987	1.69E-14
SI00232951	Lettuce mRNA for light-harvesting chlorophyll a/b-binding protein (LHCP), complete cds	D14002	9.60E-116
SI00232952	nine-cis-epoxycarotenoid dioxygenase4 [Pisum sativum]	BAC10552	2.15E-79
SI00232953	P.amygdalus mRNA for alpha-tubulin	X67162	0
SI00232954	D.stramonium mRNA for S-adenosylmethionine decarboxylase	Y07768	7.87E-21
SI00232955	pEARLI 1-like protein; protein id: At4g12520.1 [Arabidopsis thaliana] >gi|18413820|ref|NP_567391.1| pEARLI 1-like protein; protein id: At4g12510.1 [Arabidopsis thaliana] >gi|27765002|gb|AAO23622.1| At4g12520 [Arabidopsis thaliana] >gi|4725950|emb|CAB41721.1| pEARLI 1-like protein [Arabidopsis thali	NP_567392	1.44E-30
SI00232956	Nicotiana paniculata mRNA for plastidic aldolase NPALDP1, complete cds	AB027001	6.66E-80
SI00232957	light-harvesting complex protein [Arabidopsis thaliana]	AAM65689	8.35E-90

SI00232958	Arabidopsis thaliana chromosome 3 CHR3v07142002 genomic sequence	NM_111128	1.77E-12
SI00232959	Solanum tuberosum mRNA for plastidic phosphoglucomutase	AJ240053	7.70E-92
SI00232960	Arabidopsis thaliana putative 23 kDa polypeptide of oxygen-evolving complex (OEC) (At1g06680) mRNA, complete cds	AY096487	1.19E-19
SI00232961	Auxin-binding protein ABP19a precursor >gi|4098517|gb|AAD00295.1| auxin-binding protein ABP19 [Prunus persica]	Q9ZRA4	8.24E-70
SI00232962	Lycopersicon esculentum 40S ribosomal protein S17 (Gi1) mRNA, complete cds	AF161704	3.73E-25
SI00233804	N.tabacum mRNA for chloroplast Rieske FeS precursor protein 1	X66009	2.58E-47
SI00232963	Arabidopsis thaliana chromosome 5 CHR5v07142002 genomic sequence	NM_121731	7.26E-15
SI00232964	Berberis stolonifera calreticulin mRNA, complete cds	AF052040	0
SI00232965	CONSTANS B-box zinc finger family protein; protein id: At1g68520.1, supported by cDNA: 108109., supported by cDNA: gi_15451089, supported by cDNA: gi_20148424 [Arabidopsis thaliana] >gi|15451090|gb|AAK96816.1| putative B-box zinc finger protein [Arabidopsis thaliana] >gi|20148425|gb|AAM10103.1| put	NP_564932	1.12E-14
SI00232966	Zea mays PCO109946 mRNA sequence	AY104216	4.32E-44
SI00232967	Arabidopsis thaliana clone RAFL15-43-N23 (R50060) putative alternative oxidase 1c precursor (At3g27620) mRNA, complete cds	BT004153	4.86E-15
SI00232968	Vitis berlandieri x Vitis rupestris putative aquaporin PIP2-1 (PIP2-1) mRNA, complete cds	AF141642	1.89E-52
SI00232969	Coptis japonica triosphosphate isomerase mRNA, complete cds	J04121	1.98E-67
SI00232970	CONSTANS-like protein 1 [Malus x domestica]	AAC99309	7.08E-25
SI00232971	Solanum tuberosum cv. Desire mRNA for P-protein	Z99770	1.47E-118
SI00232972	lipoxygenase [Citrus jambhiri]	BAB84352	8.95E-26
SI00232973	Nicotiana tabacum poly(A)-binding protein (PABP) mRNA, complete cds	AF190655	3.32E-26
SI00232974	photosystem I reaction center IV >gi|21303|emb|CAA31523.1| PSI subunit IV preprotein (AA -77 to 154) [Spinacia oleracea] >gi|131187|sp|P12355|PSAF_SPIOL Photosystem I reaction centre subunit III, chloroplast precursor (Light-harvesting complex I 17 kDa protein) (PSI-F) >gi|72681|pir||F1SP3 photosys	1413236A	6.23E-29
SI00232975	At1g15670/F7H2_1 [Arabidopsis thaliana] >gi|12083254|gb|AAG48786.1|AF332423_1 unknown protein [Arabidopsis thaliana] >gi|15146332|gb|AAK83649.1| At1g15670/F7H2_1 [Arabidopsis thaliana] >gi|8927646|gb|AAF82137.1|AC034256_1 Contains similarity to Keap1 from Mus musculus gb|AB020063 and contains two K	AAL15368	5.13E-75
SI00232976	Arabidopsis thaliana chromosome 5 CHR5v07142002 genomic sequence	NM_122767	0
SI00232977	Unknown	No Hit	null
SI00232978	Vitis berlandieri x Vitis rupestris putative aquaporin PIP2-1 (PIP2-1) mRNA, complete cds	AF141642	4.68E-50
SI00232979	S.oleracea atpG mRNA for chloroplast CF(o)II ATP synthase subunit 9	X71397	9.05E-11
SI00232980	Unknown	No Hit	null
SI00232981	Unknown	No Hit	null
SI00232982	Arabidopsis thaliana chromosome 4 CHR4v07142002 genomic sequence	NM_117846	1.19E-13
SI00232983	Arabidopsis thaliana chromosome 4 CHR4v07142002 genomic sequence	NM_117846	9.31E-14
SI00232984	Unknown	No Hit	null
SI00232985	S.tuberosum mRNA for T subunit of glycine decarboxylase multi-enzyme complex	Z25862	2.62E-79
SI00232986	AT3g11400/F24K9_7 [Arabidopsis thaliana] >gi|18377870|gb|AAL67121.1| AT3g11400/F24K9_7 [Arabidopsis thaliana] >gi|16226341|gb|AAL16140.1|AF428308_1 AT3g11400/F24K9_7 [Arabidopsis thaliana] >gi|12322907|gb|AAG51445.1|AC008153_18 putative eukaryotic translation initiation factor 3 subunit; 21071-2290	AAM91455	4.63E-13

SI00232987	Pyrus communis PVP3 mRNA for vacuolar proton-inorganic pyrophosphatase, complete cds	AB097115	0
SI00232988	Unknown	No Hit	null
SI00232989	ORF 3 >gi|897638|gb|AAA84680.1| unknown protein >gi|7436257|pir||T02946 hypothetical protein 3 - common tobacco chloroplast	1102209C	4.24E-25
SI00232990	Pyrus communis cytochrome P450 mRNA, complete cds	AF386512	3.57E-62
SI00232991	Arabidopsis thaliana glutamate:glyoxylate aminotransferase 1 (GGT1) mRNA, complete cds	AF479639	4.18E-78
SI00232992	Pisum sativum mRNA for translation initiation factor	Y17186	2.22E-101
SI00232993	Unknown	No Hit	null
SI00232994	2-Cys-peroxiredoxin [Riccia fluitans]	CAB82860	1.74E-16
SI00232995	D.carota D2 mRNA for glycine rich protein	X72384	2.68E-11
SI00232996	Arabidopsis thaliana putative photosystem I subunit V precursor (At1g55670) mRNA, complete cds	AF339692	1.48E-27
SI00232997	Solanum tuberosum mRNA for plastidic ATP/ADP-transporter	Y10821	1.91E-55
SI00232998	polygalacturonase inhibitor protein - tomato >gi|469457|gb|AAA53547.1| polygalacturonase inhibitor protein	S47965	5.04E-71
SI00232999	Unknown	No Hit	null
SI00233000	Mesembryanthemum crystallinum protein phosphatase 2C homolog (PP2C) mRNA, complete cds	AF097667	8.16E-24
SI00233001	CPRD46 protein [Vigna unguiculata] >gi|7433160|pir||T11578 probable lipoxygenase (EC 1.13.11.12) CPRD46, drought-inducible - cowpea	BAA13542	5.46E-52
SI00233002	L.esculentum psbX mRNA for photosystem II 23 kDa protein	X63007	7.78E-18
SI00233003	Unknown	No Hit	null
SI00233004	endoxyloglucan transferase, putative [Arabidopsis thaliana] >gi|25313534|pir||B86446 probable endoxyloglucan transferase [imported] - Arabidopsis thaliana >gi|10801367|gb|AAG23439.1|AC084165_5 endoxyloglucan transferase, putative [Arabidopsis thaliana] >gi|15222593|ref|NP_174496.1| xyloglucan endot	AAL32776	7.78E-105
SI00233005	Pea ferredoxin I (Fed-1)gene, complete cds	M31713	9.71E-26
SI00233006	Arabidopsis thaliana chromosome 4 CHR4v07142002 genomic sequence	NM_119002	2.53E-25
SI00233007	Unknown	No Hit	null
SI00233008	Unknown	No Hit	null
SI00233009	unknown [Arabidopsis thaliana] >gi|18415084|ref|NP_567555.1| expressed protein; protein id: At4g18400.1, supported by cDNA: 14794. [Arabidopsis thaliana]	AAM62591	2.30E-19
SI00233010	Unknown	No Hit	null
SI00233011	Pisum sativum PSI light-harvesting antenna chlorophyll a/b-binding protein (lhcA-P4) mRNA, complete cds	AF002248	1.15E-53
SI00233012	Unknown	No Hit	null
SI00233013	Arabidopsis thaliana chromosome 5 CHR5v07142002 genomic sequence	NM_122045	2.13E-27
SI00233014	Arabidopsis thaliana chromosome 4 CHR4v07142002 genomic sequence	NM_120141	4.70E-16
SI00233015	Fagus crenata mRNA for chlorophyll a/b-binding protein, complete cds	AB006081	1.30E-96
SI00233016	Spinacea oleracea mRNA for 20 kDa protein of CP24	Z25886	7.98E-58
SI00233017	Arabidopsis thaliana unknown protein (At2g46820; F19D11.10) mRNA, complete cds	AY065156	1.12E-28
SI00233018	unknown [Arabidopsis thaliana] >gi|11357988|pir||T48031 hypothetical protein T12C14.90 - Arabidopsis thaliana >gi|7340710|emb|CAB82953.1| putative protein [Arabidopsis thaliana] >gi|15228743|ref|NP_191798.1| putative protein; protein id: At3g62390.1, supported by cDNA: 126100. [Arabidopsis thaliana]	AAM61621	4.87E-106

SI00233019	Unknown	No Hit	null
SI00233020	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	1.24E-43
SI00233021	N.tabacum mRNA for precursor of photosysten II 22 kDa protein	X84225	4.98E-22
SI00233022	expressed protein; protein id: At3g04560.1, supported by cDNA: gi_15027888, supported by cDNA: gi_20259284 [Arabidopsis thaliana] >gi|20259285|gb|AAM14378.1| unknown protein [Arabidopsis thaliana] >gi|15027889|gb|AAK76475.1| unknown protein [Arabidopsis thaliana] >gi|6175161|gb|AAF04887.1|AC011437_	NP_566231	4.35E-64
SI00233023	Unknown	No Hit	null
SI00233024	Castor bean chloroplast beta-ketoacyl-ACP synthase mRNA, complete cds	L13241	1.24E-59
SI00233025	Unknown	No Hit	null
SI00233026	Lycopersicon esculentum cryptochrome 1 mRNA, complete cds	AF130423	2.86E-51
SI00233027	chlorophyll a/b-binding protein precursor - spinach >gi|14240|emb|CAA32526.1| chlorophyll a/b binding protein precursor [Spinacia oleracea] >gi|115780|sp|P12333|CB21_SPIOL Chlorophyll A-B binding protein, chloroplast precursor (LHCII type I CAB) (LHCP)	JQ0020	1.57E-129
SI00233028	Arabidopsis thaliana unknown protein (At1g62750) mRNA, complete cds	AY142646	4.02E-78
SI00233029	Zea mays PCO077463 mRNA sequence	AY105541	5.30E-28
SI00233030	Arabidopsis thaliana clone 16846 mRNA, complete sequence	AY085669	0
SI00233031	Arabidopsis thaliana chromosome 3 CHR3v07142002 genomic sequence	NM_116077	6.47E-43
SI00233032	putative serine carboxypeptidase I [Arabidopsis thaliana] >gi|14334758|gb|AAK59557.1| putative serine carboxypeptidase I [Arabidopsis thaliana] >gi|8699619|gb|AAF78760.1|AF275313_1 sinapoylglucose:malate sinapoyltransferase [Arabidopsis thaliana] >gi|25289789|pir||C84619 probable serine carboxypept	AAK93737	5.04E-76
SI00233033	P.sativum mRNA for type II chlorophyll a/b binding protein	X81962	1.69E-83
SI00233034	Arabidopsis thaliana putative heat shock factor protein hsf8 (At1g12800) mRNA, complete cds	AY045991	3.84E-18
SI00233035	Capsicum annuum putative acyl-CoA synthetase mRNA, complete cds	AF354454	1.12E-13
SI00233036	flavonoid 3',5'-hydroxylase homolog T4C9.150 [similarity] - Arabidopsis thaliana >gi|7267932|emb|CAB78274.1| flavonoid 3', 5'-hydroxylase-like protein [Arabidopsis thaliana] >gi|5281042|emb|CAB45978.1| flavonoid 3', 5'-hydroxylase-like protein [Arabidopsis thaliana] >gi|15234517|ref|NP_192968	T48141	2.92E-87
SI00233037	Arabidopsis thaliana chromosome 3 CHR3v07142002 genomic sequence	NM_112594	1.11E-95
SI00233038	Oryza sativa (japonica cultivar-group) OsTATC mRNA, complete cds	AB050885	3.13E-20
SI00233039	Zea mays PCO122996 mRNA sequence	AY105822	1.11E-47
SI00233805	Unknown	No Hit	null
SI00233040	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	1.11E-31
SI00233041	P.sativum atpC mRNA for gamma subunit of ATP synthase	X63604	1.03E-44
SI00233042	Arabidopsis thaliana unknown protein (At5g55120) mRNA, complete cds	AY091285	2.23E-30
SI00233043	S-adenosylmethionine decarboxylase leader [Narcissus pseudonarcissus]	AAO43185	6.81E-12
SI00233806	putative dTDP-6-deoxy-L-mannose-dehydrogenase; protein id: At4g00560.1 [Arabidopsis thaliana] >gi|7486553|pir||T01220 hypothetical protein F6N23.17 - Arabidopsis thaliana >gi|7267395|emb|CAB80865.1| putative dTDP-6-deoxy-L-mannose-dehydrogenase [Arabidopsis thaliana] >gi|3047109|gb|AAC13620.1| F6N2	NP_191965	4.64E-53
SI00233044	spore coat protein-like protein [Arabidopsis thaliana]	AAM67200	1.32E-81
SI00233045	Glycine max 14-.3.3 protein mRNA, complete cds	AF532628	3.43E-66
SI00233046	Nicotiana tabacum chloroplast mRNA for FtsZ-like protein (ftsZ gene), clone FtsZ2-2	AJ311847	2.73E-48

SI00233807	Nicotiana tabacum beta-carbonic anhydrase (CA) mRNA, complete cds; nuclear gene for chloroplast product	AF454759	7.59E-18
SI00233047	BEL1-related homeotic protein 30 [Solanum tuberosum]	AAN03627	9.32E-12
SI00233808	Unknown	No Hit	null
SI00233048	expressed protein [Arabidopsis thaliana] >gi|18396349|ref|NP_565335.1| expressed protein; protein id: At2g06520.1, supported by cDNA: 14272., supported by cDNA: gi_13926200 [Arabidopsis thaliana] >gi|13926201|gb|AAK49579.1|AF370573_1 Unknown protein [Arabidopsis thaliana] >gi|7488015|pir||S70489 ph	AAD25151	3.20E-20
SI00233049	hypothetical protein; protein id: At1g48570.1, supported by cDNA: gi_17473843 [Arabidopsis thaliana] >gi|23197718|gb|AAN15386.1| unknown protein [Arabidopsis thaliana] >gi|17473844|gb|AAL38346.1| unknown protein [Arabidopsis thaliana] >gi|25372840|pir||D96525 protein T1N15.19 [imported] - Arabidops	NP_175290	9.76E-46
SI00233050	Unknown	No Hit	null
SI00233809	Fagus crenata mRNA for chlorophyll a/b-binding protein, complete cds	AB006081	6.79E-83
SI00233051	Lettuce mRNA for light-harvesting chlorophyll a/b-binding protein (LHCP), complete cds	D14002	3.87E-75
SI00233052	Unknown	No Hit	null
SI00233053	Solanum tuberosum mRNA for putative beta-subunit of K+ channel	AJ000999	4.39E-87
SI00233054	Arabidopsis thaliana chromosome 2 CHR2v07142002 genomic sequence	NM_127571	9.63E-39
SI00233055	unknown protein; protein id: At3g07560.1, supported by cDNA: gi_18700168, supported by cDNA: gi_20147310 [Arabidopsis thaliana] >gi|20147311|gb|AAM10369.1| AT3g07560/F21O3_27 [Arabidopsis thaliana] >gi|18700169|gb|AAL77696.1| AT3g07560/F21O3_27 [Arabidopsis thaliana] >gi|6041851|gb|AAF02160.1|AC009	NP_187412	1.19E-27
SI00233056	Unknown	No Hit	null
SI00233057	Arabidopsis thaliana chromosome 3 CHR3v07142002 genomic sequence	NM_115553	1.78E-83
SI00233058	Elaeis guineensis mRNA for metallothionein-like protein	AJ236913	3.43E-12
SI00233059	Glycine max SG-05 gene for thiamin biosynthetic enzyme, complete cds	AB030490	6.68E-43
SI00233060	Arabidopsis thaliana chromosome 4 CHR4v07142002 genomic sequence	NM_117080	2.93E-14
SI00233061	Unknown	No Hit	null
SI00233062	Medicago truncatula clone mth2-24h4, complete sequence	AC122164	4.62E-15
SI00233063	G.hirsutum cab gene for chlorophyll ab binding protein	X54090	1.85E-89
SI00233810	Pea ferredoxin I (Fed-1)gene, complete cds	M31713	9.89E-29
SI00233811	Unknown	No Hit	null
SI00233064	Cicer arietinum ADP-glucose pyrophosphorylase mRNA, complete cds	AF356002	4.95E-59
SI00233065	Arabidopsis thaliana chromosome 4 CHR4v07142002 genomic sequence	NM_117835	4.16E-44
SI00233066	Tabacco Cab21 mRNA for major chlorophyll a/b binding protein	X52743	3.15E-37
SI00233067	unknown protein [Arabidopsis thaliana] >gi|18416995|ref|NP_567775.1| putative protein; protein id: At4g27450.1, supported by cDNA: 35207., supported by cDNA: gi_14030722, supported by cDNA: gi_15028056, supported by cDNA: gi_20259058 [Arabidopsis thaliana] >gi|21593019|gb|AAM64968.1| unknown [Arabi	AAK76559	5.33E-71
SI00233068	Unknown	No Hit	null
SI00233069	Fagus crenata mRNA for chlorophyll a/b-binding protein, complete cds	AB006081	1.91E-89
SI00233070	putative oxalyl-CoA decarboxylase [Oryza sativa (japonica cultivar-group)]	BAB33274	8.78E-119
SI00233071	Unknown	No Hit	null

SI00233072	Arabidopsis thaliana chromosome 1 CHR1v07142002 genomic sequence	NM_105732	3.05E-14
SI00233073	Arabidopsis thaliana chromosome 1 CHR1v07142002 genomic sequence	NM_104023	4.35E-29
SI00233074	Arabidopsis thaliana putative ribose phosphate pyrophosphokinase (At2g42910) mRNA, complete cds	BT000083	3.10E-57
SI00233075	P.sativum Cab-8 gene for photosystemm II chlorophyll a/b binding protein	X56538	0
SI00233076	Arabidopsis thaliana chromosome 1 CHR1v07142002 genomic sequence	NM_106656	1.64E-80
SI00233077	unknown protein [Arabidopsis thaliana] >gi|15232373|ref|NP_188718.1| disease resistance protein family (LRR); protein id: At3g20820.1, supported by cDNA: gi_17380931 [Arabidopsis thaliana] >gi|9294409|dbj|BAB02490.1| polygalacturonase inhibitor-like protein [Arabidopsis thaliana] >gi|21436417|gb|AA	AAL36278	3.39E-83
SI00233078	Arabidopsis thaliana putative acetyltransferase (At3g13930) mRNA, complete cds	BT001223	2.36E-36
SI00233079	ribosomal S29-like protein - Arabidopsis thaliana >gi|11908082|gb|AAG41470.1|AF326888_1 putative ribosomal S29 protein [Arabidopsis thaliana] >gi|26452256|dbj|BAC43215.1| putative ribosomal S29 subunit [Arabidopsis thaliana] >gi|22136560|gb|AAM91066.1| AT3g43980/T15B3_120 [Arabidopsis thaliana] >gi	T48952	1.08E-23
SI00233080	Arabidopsis thaliana chromosome 3 CHR3v07142002 genomic sequence	NM_113616	9.36E-33
SI00233081	Arabidopsis thaliana chromosome 4 CHR4v07142002 genomic sequence	NM_117349	3.19E-26
SI00233082	Spinach mRNA for photosystem I subunit IV	X13133	8.34E-27
SI00233083	protein phosphatase 2C (PP2C); protein id: At3g02750.1, supported by cDNA: gi_17064839 [Arabidopsis thaliana] >gi|6728987|gb|AAF26985.1|AC018363_30 putative protein phosphatase-2C (PP2C) [Arabidopsis thaliana]	NP_186924	2.09E-11
SI00233084	Nicotiana tabacum beta-carbonic anhydrase (CA) mRNA, complete cds; nuclear gene for chloroplast product	AF454759	1.83E-15
SI00233085	lipid transfer protein [Helianthus annuus] >gi|6093512|sp|Q39950|NLTP_HELAN NONSPECIFIC LIPID-TRANSFER PROTEIN PRECURSOR (LTP) (NSLTP) (SDI-9)	CAA63340	1.78E-35
SI00233086	Nicotiana tabacum mRNA for ethylene responsive element binding factor, complete cds	AB024575	1.27E-22
SI00233087	P.sativum Cab-8 gene for photosystemm II chlorophyll a/b binding protein	X56538	0
SI00233088	N.tabacum mRNA (T23-5B) for PSII 23-kDa polypeptide	X62426	3.04E-20
SI00233812	unknown protein [Arabidopsis thaliana] >gi|15810277|gb|AAL07026.1| unknown protein [Arabidopsis thaliana] >gi|18395868|ref|NP_566143.1| expressed protein; protein id: At3g01590.1, supported by cDNA: 41354., supported by cDNA: gi_16612257, supported by cDNA: gi_20258972 [Arabidopsis thaliana] >gi|21	AAM14202	2.54E-17
SI00233089	Arabidopsis thaliana At5g59380 gene, complete cds	BT004606	1.71E-11
SI00233090	At1g79600/F20B17_3 [Arabidopsis thaliana] >gi|18412437|ref|NP_565214.1| expressed protein; protein id: At1g79600.1, supported by cDNA: gi_14532489, supported by cDNA: gi_18086324 [Arabidopsis thaliana] >gi|18086325|gb|AAL57626.1| At1g79600/F20B17_3 [Arabidopsis thaliana] >gi|7715610|gb|AAF68128.1|A	AAM98256	0
SI00233091	Arabidopsis thaliana chromosome 1 CHR1v07142002 genomic sequence	NM_105792	1.33E-15
SI00233092	Unknown	No Hit	null
SI00233093	S-adenosyl-L-methionine decarboxylase [Dendrobium crumenatum]	AAL16065	9.88E-95
SI00233094	Prunus cerasus expansin (EXP2) mRNA, complete cds	AF350937	2.34E-70
SI00233095	Unknown	No Hit	null
SI00233096	Arabidopsis thaliana At1g53500/F22G10_13 mRNA, complete cds	BT002264	5.36E-28
SI00233097	Vigna radiata chlorophyll a/b binding protein CP29 (CipCp29) mRNA, complete cds; nuclear gene for chloroplast product	AF139466	2.77E-48
SI00233098	pyruvate dehydrogenase E1 beta subunit isoform 2 [Zea mays]	AAC72193	2.17E-127
SI00233099	Arabidopsis thaliana clone 94409 mRNA, complete sequence	AY088757	1.48E-37

SI00233100	Unknown	No Hit	null
SI00233101	Arabidopsis thaliana At3g06670/T8E24.10 mRNA, complete cds	AY094413	5.56E-40
SI00233102	putative protein; protein id: At5g19380.1, supported by cDNA: 158575. [Arabidopsis thaliana]	NP_568373	9.79E-23
SI00233103	Arabidopsis thaliana chromosome 3 CHR3v07142002 genomic sequence	NM_115209	9.44E-33
SI00233104	Arabidopsis thaliana chromosome 3 CHR3v07142002 genomic sequence	NM_111597	9.17E-33
SI00233105	Fagus crenata mRNA for chlorophyll a/b-binding protein, complete cds	AB006081	1.25E-93
SI00233106	emb|CAB70981.1~gene_id:MVE11.3~similar to unknown protein [Arabidopsis thaliana]	BAB01793	5.71E-22
SI00233107	S.oleracea mRNA for 6.1 kDa polypeptide of photosystem II	X85038	1.75E-14
SI00233108	Arabidopsis thaliana cytosolic IMP-GMP specific 5-nucleotidase, putative (At1g75210) mRNA, complete cds	BT003378	1.34E-25
SI00233109	gene_id:MCO15.1~unknown protein [Arabidopsis thaliana] >gi|15240453|ref|NP_200317.1| unknown protein; protein id: At5g55060.1 [Arabidopsis thaliana]	BAB08575	1.31E-54
SI00233110	Unknown	No Hit	null
SI00233111	ADP-glucose pyrophosphorylase small subunit [Brassica napus] >gi|17865468|sp|Q9M462|GLGS_BRANA Glucose-1-phosphate adenylyltransferase small subunit, chloroplast precursor (ADP-glucose synthase) (ADP-glucose pyrophosphorylase) (AGPASE B) (Alpha-D-glucose-1-phosphate adenyl transferase)	CAB89863	4.71E-116
SI00233112	unknown protein [Arabidopsis thaliana] >gi|21592833|gb|AAM64783.1| unknown [Arabidopsis thaliana] >gi|26453214|dbj|BAC43681.1| unknown protein [Arabidopsis thaliana] >gi|20260264|gb|AAM13030.1| unknown protein [Arabidopsis thaliana] >gi|12323736|gb|AAG51832.1|AC016163_21 unknown protein; 39989-3874	AAM91334	7.03E-78
SI00233113	Unknown	No Hit	null
SI00233114	remorin - potato >gi|1881585|gb|AAB49425.1| remorin [Solanum tuberosum]	T07780	1.72E-36
SI00233115	ORF 3 >gi|897638|gb|AAA84680.1| unknown protein >gi|7436257|pir||T02946 hypothetical protein 3 - common tobacco chloroplast	1102209C	5.87E-24
SI00233116	M.domestica ribulose-1,5-bisphosphate carboxylase/oxygenase activase mRNA	Z21794	7.85E-30
SI00233117	expressed protein [Arabidopsis thaliana] >gi|23197966|gb|AAN15510.1| expressed protein [Arabidopsis thaliana]	AAM97011	2.01E-40
SI00233118	superoxide dismutase (EC 1.15.1.1) (Cu-Zn) 2, chloroplast - wheat >gi|1572627|gb|AAB67991.1| Cu/Zn superoxide dismutase [Triticum aestivum]	T06800	4.50E-68
SI00233119	Vigna radiata chlorophyll a/b binding protein CP29 (CipCp29) mRNA, complete cds; nuclear gene for chloroplast product	AF139466	1.00E-41
SI00233120	P.sativum Cab-8 gene for photosystemm II chlorophyll a/b binding protein	X56538	0
SI00233121	60S ribosomal protein L4-B (L1) [Arabidopsis thaliana]	AAM65510	3.94E-29
SI00233122	Arabidopsis thaliana glutamate:glyoxylate aminotransferase 2 (GGT2) mRNA, complete cds	AF479640	7.05E-89
SI00233123	Arabidopsis thaliana chromosome 3 CHR3v07142002 genomic sequence	NM_116155	4.50E-21
SI00233124	Lotus japonicus genomic DNA, chromosome , clone:LjT35I07, TM0121b, complete sequence	AP004948	1.23E-46
SI00233125	Arabidopsis thaliana chromosome 2 CHR2v07142002 genomic sequence	NM_130092	6.46E-117
SI00233126	Arabidopsis thaliana chromosome 1 BAC F9L1 sequence, complete sequence	AC007591	1.76E-12
SI00233127	Arabidopsis thaliana chromosome 1 CHR1v07142002 genomic sequence	NM_103931	2.20E-67
SI00233128	Unknown	No Hit	null
SI00233129	Unknown	No Hit	null
SI00233130	plastid-lipid associated protein PAP3 [Brassica rapa]	AAK57563	5.31E-56
SI00233131	Lotus japonicus genomic DNA, chromosome , clone:LjT35I07, TM0121b, complete sequence	AP004948	1.98E-51

SI00233132	Spinach mRNA for photosystem I subunit IV	X13133	3.52E-41
SI00233133	Arabidopsis thaliana glyoxalase II cytoplasmic isozyme (Glx2-2) mRNA, complete cds	U90929	7.31E-27
SI00233134	glutathione S-transferase 2 [Papaver somniferum] >gi|6652870|gb|AAF22517.1|AF118924_1 glutathione S-transferase 1 [Papaver somniferum]	AAF22518	5.84E-119
SI00233135	Unknown	No Hit	null
SI00233136	Unknown	No Hit	null
SI00233137	B.juncea mRNA for chlorophyll a/b-binding protein	X95727	7.78E-58
SI00233138	Unknown	No Hit	null
SI00233139	Spinacea oleracea mRNA for 20 kDa protein of CP24	Z25886	6.00E-58
SI00233140	Arabidopsis thaliana chromosome 3 CHR3v07142002 genomic sequence	NM_115049	4.91E-22
SI00233141	glutathione S-transferase 2 [Papaver somniferum] >gi|6652870|gb|AAF22517.1|AF118924_1 glutathione S-transferase 1 [Papaver somniferum]	AAF22518	2.38E-101
SI00233142	Actin (AA 1-377) [Oryza sativa]	CAA34356	7.22E-29
SI00233143	Cucumis melo aminotransferase 1 mRNA, complete cds	AY066012	7.70E-58
SI00233144	Pisum sativum nced1 mRNA for nine-cis-epoxycarotenoid dioxygenase1, complete cds	AB080191	3.97E-78
SI00233145	B.juncea mRNA for chlorophyll a/b-binding protein	X95727	6.76E-58
SI00233146	Spinacia oleracea Alwaro ribosomal protein 5 precursor (Psrp-5) mRNA, complete cds; nuclear gene for chloroplast product	AF261940	6.16E-11
SI00233147	expressed protein; protein id: At3g15840.1, supported by cDNA: gi_13358213, supported by cDNA: gi_14334703 [Arabidopsis thaliana] >gi|11994356|dbj|BAB02315.1| gene_id:MSJ11.24~unknown protein [Arabidopsis thaliana] >gi|22136938|gb|AAM91813.1| unknown protein [Arabidopsis thaliana] >gi|14334704|gb|A	NP_566528	1.22E-85
SI00233148	Unknown	No Hit	null
SI00233149	Unknown	No Hit	null
SI00233150	D.carota D2 mRNA for glycine rich protein	X72384	1.01E-13
SI00233151	Unknown	No Hit	null
SI00233152	Unknown	No Hit	null
SI00233153	putative pectinesterase [Arabidopsis thaliana] >gi|15238729|ref|NP_200149.1| pectinesterase; protein id: At5g53370.1, supported by cDNA: gi_13507548, supported by cDNA: gi_15293286 [Arabidopsis thaliana] >gi|9759184|dbj|BAB09799.1| pectinesterase [Arabidopsis thaliana] >gi|15293287|gb|AAK93754.1| p	AAK28637	1.26E-126
SI00233154	Unknown	No Hit	null
SI00233155	ornithine decarboxylase [Datura stramonium] >gi|1706323|sp|P50134|DCOR_DATST ORNITHINE DECARBOXYLASE (ODC) >gi|2118242|pir||S64704 ornithine decarboxylase (EC 4.1.1.17) - jimsonweed	CAA61121	4.41E-107
SI00233156	Arabidopsis thaliana chromosome 3 CHR3v07142002 genomic sequence	NM_111840	1.12E-13
SI00233157	Lycopersicon esculentum mRNA binding protein precursor, mRNA, nuclear gene encoding chloroplast protein, complete cds	AF106660	1.39E-31
SI00233158	Unknown	No Hit	null
SI00233159	Arabidopsis thaliana PEX14 mRNA, complete cds	AB037539	5.21E-23
SI00233160	Pea ferredoxin I (Fed-1)gene, complete cds	M31713	6.83E-33
SI00233161	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	7.58E-30

SI00233162	hypothetical protein - Arabidopsis thaliana >gi|7268395|emb|CAB78687.1| hypothetical protein [Arabidopsis thaliana] >gi|2245001|emb|CAB10421.1| hypothetical protein [Arabidopsis thaliana] >gi|18414665|ref|NP_567500.1| expressed protein; protein id: At4g16450.1, supported by cDNA: 16090. [Arabidopsi	C71431	4.37E-17
SI00233163	Unknown	No Hit	null
SI00233164	AT3g24190/MUJ8_17 [Arabidopsis thaliana] >gi|11994238|dbj|BAB01360.1| gene_id:MUJ8.9~similar to unknown protein~sp|Q55680 [Arabidopsis thaliana] >gi|15294250|gb|AAK95302.1|AF410316_1 AT3g24190/MUJ8_17 [Arabidopsis thaliana] >gi|18404075|ref|NP_566745.1| expressed protein; protein id: At3g24190.1, s	AAM47311	4.92E-44
SI00233165	Unknown	No Hit	null
SI00233166	Nicotiana tabacum beta-carbonic anhydrase (CA) mRNA, complete cds; nuclear gene for chloroplast product	AF454759	7.77E-18
SI00233167	Lettuce mRNA for light-harvesting chlorophyll a/b-binding protein (LHCP), complete cds	D14002	0
SI00233168	Tabacco Cab16 mRNA for major chlorophyll a/b binding protein	X52741	4.67E-90
SI00233169	Arabidopsis thaliana chromosome 1 CHR1v07142002 genomic sequence	NM_102943	4.77E-19
SI00233170	Lycopersicon esculentum mRNA for beta-alanine synthase	Y19104	7.59E-95
SI00233171	Zea mays PCO128171 mRNA sequence	AY108767	8.19E-27
SI00233172	Cicer arietinum ADP-glucose pyrophosphorylase mRNA, complete cds	AF356002	1.55E-40
SI00233173	Unknown	No Hit	null
SI00233174	AT3g21670/MIL23_23 [Arabidopsis thaliana] >gi|15233123|ref|NP_188804.1| nitrate transporter; protein id: At3g21670.1, supported by cDNA: 111089., supported by cDNA: gi_13937208 [Arabidopsis thaliana] >gi|11994403|dbj|BAB02362.1| nitrate transporter [Arabidopsis thaliana] >gi|23506051|gb|AAN28885.1|	AAK50097	5.30E-100
SI00233175	N.sylvestris psaDa gene for PSI-D2	X60008	6.17E-58
SI00233176	unknown protein [Arabidopsis thaliana] >gi|17104651|gb|AAL34214.1| unknown protein [Arabidopsis thaliana] >gi|18378973|ref|NP_563656.1| unknown protein; protein id: At1g02475.1, supported by cDNA: gi_13878058, supported by cDNA: gi_17104650 [Arabidopsis thaliana]	AAK44107	1.86E-51
SI00233177	AtPH1-like protein [Arabidopsis thaliana]	AAM61053	5.34E-63
SI00233178	P.sativum mRNA for actin	X67666	0
SI00233179	Arabidopsis thaliana chromosome 5 CHR5v07142002 genomic sequence	NM_125800	1.64E-23
SI00233180	Capsicum annuum mRNA for partial glyceraldehyde-3-phosphate dehydrogenase, subunit GapB	AJ246010	1.60E-77
SI00233181	Phaseolus vulgaris cyclic nucleotide-gated channel C (CNGC-C) mRNA, partial cds	AF492818	6.83E-12
SI00233182	L.polyphyllus pPLB08 mRNA	X54702	4.23E-12
SI00233183	Zea mays PCO098989 mRNA sequence	AY103717	9.96E-45
SI00233184	N.tabacum mRNA for precursor of photosysten II 22 kDa protein	X84225	3.15E-23
SI00233185	Lotus japonicus genomic DNA, chromosome , clone:LjT35I07, TM0121b, complete sequence	AP004948	2.06E-51
SI00233186	CONSTANS B-box zinc finger family protein; protein id: At5g24930.1, supported by cDNA: gi_15450658 [Arabidopsis thaliana]	NP_197875	1.28E-50
SI00233187	AT3g52740/F3C22_140 [Arabidopsis thaliana] >gi|15450655|gb|AAK96599.1| AT3g52740/F3C22_140 [Arabidopsis thaliana] >gi|11289959|pir||T49027 hypothetical protein F3C22.140 - Arabidopsis thaliana >gi|7669948|emb|CAB89235.1| putative protein [Arabidopsis thaliana] >gi|18409685|ref|NP_566972.1| expresse	AAL38610	5.59E-21
SI00233188	S.tuberosum mRNA for glycine hydroxymethyltransferase	Z25863	5.58E-24
SI00233189	methionine adenosyltransferase [Petunia x hybrida] >gi|1084428|pir||S49491 methionine adenosyltransferase (EC 2.5.1.6) - garden petunia >gi|1346523|sp|P48498|METK_PETHY S-adenosylmethionine synthetase (Methionine adenosyltransferase) (AdoMet synthetase)	CAA57696	2.96E-108

SI00233190	glutamine synthetase GS2 [Beta vulgaris]	AAK07678	3.67E-70
SI00233191	Arabidopsis thaliana genomic DNA, chromosome 3, BAC clone:T6J22	AP001314	7.67E-30
SI00233192	L.esculentum mRNA for 33kDa precursor protein of oxygen-evolving complex	Z11999	9.68E-39
SI00233193	Elaeis guineensis mRNA for metallothionein-like protein	AJ236913	3.59E-12
SI00233194	Unknown	No Hit	null
SI00233195	RRM-containing protein; protein id: At2g05160.1 [Arabidopsis thaliana] >gi|25344225|pir||E84465 hypothetical protein At2g05160 [imported] - Arabidopsis thaliana >gi|4755187|gb|AAD29054.1| hypothetical protein [Arabidopsis thaliana]	NP_178588	3.65E-15
SI00233196	Unknown	No Hit	null
SI00233197	Arabidopsis thaliana chromosome 1 CHR1v07142002 genomic sequence	NM_100484	3.60E-43
SI00233198	EST gb|N65787 comes from this gene. [Arabidopsis thaliana] >gi|25511669|pir||B86400 T17H3.1 protein - Arabidopsis thaliana	AAD45989	3.33E-59
SI00233199	P.sativum Cab-8 gene for photosystemm II chlorophyll a/b binding protein	X56538	0
SI00233200	putative WD-repeat protein [Arabidopsis thaliana] >gi|21689757|gb|AAM67522.1| putative WD-repeat protein [Arabidopsis thaliana]	AAL87251	1.16E-72
SI00233201	Unknown	No Hit	null
SI00233202	M.truncatula mRNA for glutamine synthetase, 1488bp	Y10268	2.61E-86
SI00233203	Unknown	No Hit	null
SI00233204	Medicago sativa subsp. falcata mRNA for beta-tubulin	AJ319667	4.61E-16
SI00233205	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	1.06E-31
SI00233206	Arabidopsis thaliana At1g05910/T20M3_16 gene, complete cds	BT002728	3.25E-29
SI00233207	N.tabacum NeIF-4A11 mRNA	X79136	4.05E-121
SI00233813	Unknown	No Hit	null
SI00233208	Unknown	No Hit	null
SI00233209	S.oleracea 3-ketoacyl-acyl carrier protein synthase III (KAS III)	Z22771	2.41E-14
SI00233210	Capsicum annuum mRNA for partial glyceraldehyde-3-phosphate dehydrogenase, subunit GapA2	AJ246009	4.20E-84
SI00233211	Hevea brasiliensis DnaJ protein mRNA, complete cds	AF085275	2.02E-23
SI00233212	Unknown	No Hit	null
SI00233213	zinc-finger protein [Petunia x hybrida]	BAA05079	9.50E-35
SI00233214	Capsicum annuum mRNA for partial glyceraldehyde-3-phosphate dehydrogenase, subunit GapB	AJ246010	5.64E-71
SI00233215	Nicotiana tabacum beta-carbonic anhydrase (CA) mRNA, complete cds; nuclear gene for chloroplast product	AF454759	1.78E-15
SI00233216	Arabidopsis thaliana clone 37506 mRNA, complete sequence	AY087667	8.02E-64
SI00233217	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	3.40E-41
SI00233218	Nicotiana tabacum Avr9/Cf-9 rapidly elicited protein 132 (ACRE132) mRNA, complete cds	AF211532	1.74E-18
SI00233219	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 68	AL161572	2.55E-25
SI00233220	hypothetical protein F20O9.30 - Arabidopsis thaliana	T04605	4.55E-24
SI00233221	Arabidopsis thaliana putative heme A:farnesyltransferase (At2g44520) mRNA, complete cds	AY045952	4.63E-11
SI00233222	Arabidopsis thaliana chromosome 5 CHR5v07142002 genomic sequence	NM_123672	3.77E-35

SI00233223	Daucus carota mRNA for citrate synthase, complete cds	AB017159	7.73E-64
SI00233224	chlorophyll a/b-binding protein precursor - garden petunia >gi|169214|gb|AAA33711.1| chlorophyll binding protein precursor >gi|115766|sp|P13869|CB12_PETHY Chlorophyll A-B binding protein, chloroplast precursor (LHCI type II CAB) >gi|226259|prf||1503272A chlorophyll binding protein	S00442	7.12E-28
SI00233225	Beta vulgaris phophate translocator (pt) mRNA, partial cds; nuclear gene for chloroplast product	AF173679	8.97E-67
SI00233226	hypothetical protein At2g21740 [imported] - Arabidopsis thaliana >gi|15227092|ref|NP_179766.1| hypothetical protein; protein id: At2g21740.1 [Arabidopsis thaliana] >gi|4417269|gb|AAD20394.1| hypothetical protein [Arabidopsis thaliana]	F84604	7.87E-28
SI00233227	unknown protein [Arabidopsis thaliana] >gi|25054983|gb|AAN71963.1| unknown protein [Arabidopsis thaliana]	AAL86002	1.12E-107
SI00233228	Unknown	No Hit	null
SI00233229	Arabidopsis thaliana chromosome 2 CHR2v07142002 genomic sequence	NM_126349	6.60E-12
SI00233230	cytochrome P450 [Solanum tuberosum]	CAC24711	9.27E-73
SI00233231	Unknown	No Hit	null
SI00233232	putative p68 RNA helicase [Arabidopsis thaliana] >gi|19347812|gb|AAL86356.1| putative p68 RNA helicase [Arabidopsis thaliana] >gi|15222526|ref|NP_174479.1| DEAD/DEAH box RNA helicase, putative; protein id: At1g31970.1, supported by cDNA: gi_19347811 [Arabidopsis thaliana] >gi|12321459|gb|AAG50784.1	AAM51373	0
SI00233233	L.esculentum psbX mRNA for photosystem II 23 kDa protein	X63007	7.93E-18
SI00233234	Arabidopsis thaliana chromosome 5 CHR5v07142002 genomic sequence	NM_124033	1.47E-105
SI00233235	Unknown	No Hit	null
SI00233236	putative mitochondrial processing peptidase alpha subuunit, mitochondrial recursor(ALPHA-MPP) [Oryza sativa (japonica cultivar-group)]	BAB55500	4.07E-93
SI00233237	Nicotiana tabacum homeodomain protein Hfi22 (Hfi22) mRNA, partial cds	AY101610	2.09E-24
SI00233238	Arabidopsis thaliana putative protein phosphatase homolog PPH1 (At4g27800) mRNA, complete cds	AY114060	7.12E-12
SI00233239	Vitis berlandieri x Vitis rupestris putative aquaporin PIP2-1 (PIP2-1) mRNA, complete cds	AF141642	2.68E-57
SI00233240	myo-inositol-1-phosphate synthase [Suaeda maritima subsp. salsa]	AAL28131	3.89E-27
SI00233241	unknown [Arabidopsis thaliana]	AAM63735	1.80E-34
SI00233242	Capsicum annuum mRNA for partial glyceraldehyde-3-phosphate dehydrogenase, subunit GapA2	AJ246009	1.72E-83
SI00233243	Unknown	No Hit	null
SI00233244	putative threonine dehydratase/deaminase [Oryza sativa]	AAK18849	1.49E-81
SI00233245	Arabidopsis thaliana HDA8 mRNA, complete cds	AF510167	1.84E-15
SI00233246	putative protein [Arabidopsis thaliana] >gi|21389651|gb|AAM48024.1| putative protein [Arabidopsis thaliana] >gi|22326731|ref|NP_196706.2| oxygen-evolving complex related protein; protein id: At5g11450.1, supported by cDNA: gi_18252954 [Arabidopsis thaliana]	AAL62404	9.06E-42
SI00233247	copper/zinc-superoxide dismutase [Ananas comosus] >gi|13431904|sp|Q9SQL5|SODC_ANACO Superoxide dismutase [Cu-Zn]	CAB60191	2.42E-18
SI00233248	NADPH-dependent codeinone reductase [Papaver somniferum]	AAF13736	5.97E-123
SI00233249	Unknown	No Hit	null
SI00233250	Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, PAC clone:P0698A04	AP002868	5.35E-12
SI00233251	N.tabacum NeIF-4A11 mRNA	X79136	4.77E-127
SI00233252	A.majus gene for protoporphyrin IX: Mg Chelatase, subunit	X73144	2.61E-79

SI00233253	Lycopersicon esculentum mRNA binding protein precursor, mRNA, nuclear gene encoding chloroplast protein, complete cds	AF106660	3.86E-75
SI00233254	Unknown	No Hit	null
SI00233255	AMP1 [Macadamia integrifolia] >gi|3121752|sp|P80915|AMP1_MACIN Antimicrobial peptide 1 precursor (AMP1)	CAA71842	3.80E-17
SI00233256	Arabidopsis thaliana chromosome 2 CHR2v07142002 genomic sequence	NM_129305	1.19E-16
SI00233257	Unknown	No Hit	null
SI00233258	Arabidopsis thaliana chromosome 1 CHR1v07142002 genomic sequence	NM_106135	1.78E-36
SI00233259	putative protein; protein id: At3g51520.1, supported by cDNA: 36721., supported by cDNA: gi_13605701, supported by cDNA: gi_18700271 [Arabidopsis thaliana] >gi|13605702|gb|AAK32844.1|AF361832_1 AT3g51520/F26O13_160 [Arabidopsis thaliana] >gi|18700272|gb|AAL77746.1| AT3g51520/F26O13_160 [Arabidopsis	NP_566952	4.63E-92
SI00233260	Arabidopsis thaliana chromosome V Lhcb3 protein (Lhcb3) mRNA, complete cds	AF134126	1.92E-89
SI00233261	Ajuga reptans mRNA for galactinol synthase, isoform GolS-1	AJ237693	1.43E-28
SI00233262	Zn transporter >gi|7268765|emb|CAB78971.1| Fe(II) transport protein [Arabidopsis thaliana] >gi|1353266|gb|AAB01678.1| Fe(II) transport protein >gi|3250678|emb|CAA19686.1| Fe(II) transport protein [Arabidopsis thaliana] >gi|7488126|pir||T04750 probable iron transport protein T16H5.50 - Arabidopsis t	2211425A	5.05E-61
SI00233263	Tobacco (N.tabacum) GapB mRNA encoding B-subunit of chloroplast glyceradehyde-3-phosphate dehydrogenase	M14418	2.75E-35
SI00233264	Unknown	No Hit	null
SI00233265	Fagus crenata mRNA for chlorophyll a/b-binding protein, complete cds	AB006081	3.09E-94
SI00233266	putative CER3 [Oryza sativa (japonica cultivar-group)]	BAB43991	6.14E-61
SI00233267	ATP synthase C chain (Lipid-binding protein) (Subunit III) >gi|11465941|ref|NP_054483.1| ATP synthase CF0 C chain [Nicotiana tabacum] >gi|67898|pir||LWNTA H+-transporting two-sector ATPase (EC 3.6.3.14) lipid-binding protein - common tobacco chloroplast >gi|27807856|dbj|BAC55331.1| ATPase III subun	P06286	2.84E-25
SI00233268	B.pendula mRNA for nitrite reductase	X60093	5.80E-34
SI00233269	L.esculentum mRNA for 33kDa precursor protein of oxygen-evolving complex	Z11999	3.22E-57
SI00233270	putative serine carboxypeptidase II; protein id: At3g17180.1 [Arabidopsis thaliana]	NP_188343	1.09E-92
SI00233271	Phaseolus vulgaris mRNA for peroxiredoxin (2-Cys PRx gene)	AJ288895	7.88E-92
SI00233272	hypothetical protein [Arabidopsis thaliana] >gi|25454673|pir||T52460 hypothetical protein RXW8 [imported] - Arabidopsis thaliana >gi|6573289|dbj|BAA88270.1| RXW8 [Arabidopsis thaliana] >gi|22330304|ref|NP_683440.1| Expressed protein; protein id: At1g58520.1, supported by cDNA: gi_6573288 [Arabidops	BAB83877	3.26E-60
SI00233273	Arabidopsis thaliana putative protein (At5g46340) mRNA, complete cds	AY136339	1.34E-25
SI00233274	osmotic stress-activated protein kinase [Nicotiana tabacum]	AAL89456	9.04E-87
SI00233275	Fagus crenata mRNA for chlorophyll a/b-binding protein, complete cds	AB006081	5.19E-96
SI00233276	putative Hs1pro-1-like receptor [Glycine max]	AAG44839	1.26E-64
SI00233277	N.tabacum mRNA for chloroplast Rieske FeS precursor protein 1	X66009	1.30E-59
SI00233278	Unknown	No Hit	null
SI00233814	Mesembryanthemum crystallinum methionine synthase (MetE) mRNA, complete cds	U84889	5.77E-111
SI00233279	P0018C10.15 [Oryza sativa (japonica cultivar-group)] >gi|22202681|dbj|BAC07339.1| P0471B04.23 [Oryza sativa (japonica cultivar-group)]	BAC06214	5.98E-27
SI00233815	Unknown	No Hit	null
SI00233280	Arabidopsis thaliana chromosome 4 CHR4v07142002 genomic sequence	NM_118709	7.16E-12

SI00233281	putative beta-1,3-glucanase [Arabidopsis thaliana] >gi|20197850|gb|AAM15281.1| putative beta-1,3-glucanase [Arabidopsis thaliana] >gi|25316444|pir||E84471 probable beta-1,3-glucanase [imported] - Arabidopsis thaliana >gi|15224979|ref|NP_178637.1| glycosyl hydrolase family 17; protein id: At2g05790.	AAD26909	2.32E-103
SI00233282	S.tuberosum mRNA for aconitase/aconitate hydratase	X97012	6.31E-114
SI00233283	hypothetical protein; protein id: At1g25510.1, supported by cDNA: gi_20466515 [Arabidopsis thaliana] >gi|25518510|pir||D86385 hypothetical protein F2J7.6 - Arabidopsis thaliana >gi|20466516|gb|AAM20575.1| unknown protein [Arabidopsis thaliana] >gi|12321511|gb|AAG50814.1|AC079281_16 hypothetical pro	NP_173922	7.57E-86
SI00233284	putative class I chitinase [Arabidopsis thaliana] >gi|14334488|gb|AAK59442.1| putative class I chitinase [Arabidopsis thaliana] >gi|15221283|ref|NP_172076.1| glycosyl hydrolase family 19 (chitinase); protein id: At1g05850.1, supported by cDNA: gi_12083323, supported by cDNA: gi_14334487, supported	AAG48821	2.14E-109
SI00233285	putative capsid protein [Rhizoctonia solani virus] >gi|7211416|gb|AAF40300.1|AF133291_1 putative capsid protein [Rhizoctonia solani virus]	NP_620660	2.98E-55
SI00233286	Unknown	No Hit	null
SI00233287	Unknown	No Hit	null
SI00233288	alpha-tubulin [Daucus carota]	AAG02564	1.07E-113
SI00233289	Prunus persica unknown mRNA	AY012684	4.29E-47
SI00233290	Lotus japonicus genomic DNA, chromosome 1, clone:LjT43B20, TM0117b, complete sequence	AP004945	1.49E-34
SI00233291	contains similarity to isoamyl acetate-hydrolyzing esterase~gene_id:K15I22.12 [Arabidopsis thaliana] >gi|15242539|ref|NP_199404.1| putative protein; protein id: At5g45920.1 [Arabidopsis thaliana]	BAB09320	1.65E-69
SI00233292	enhanced disease susceptibility 5 [Arabidopsis thaliana]	AAM63262	2.71E-72
SI00233293	contains similarity to 30s ribosomal protein s1~gene_id:MYM9.4 [Arabidopsis thaliana] >gi|22655022|gb|AAM98102.1| At3g23700/MYM9_3 [Arabidopsis thaliana] >gi|18403854|ref|NP_566737.1| expressed protein; protein id: At3g23700.1, supported by cDNA: gi_15146283 [Arabidopsis thaliana] >gi|15146284|gb|A	BAB01847	7.23E-27
SI00233294	Ipomoea nil PNIL34 mRNA, complete cds	U37437	7.78E-54
SI00233295	F.pringlei gdcsPA gene for P-protein of the glycine cleavage system	Z36879	1.08E-81
SI00233296	unknown protein F14G6.19 [imported] - Arabidopsis thaliana >gi|12323979|gb|AAG51950.1|AC015450_11 unknown protein; 77280-78196 [Arabidopsis thaliana]	A96794	1.52E-25
SI00233297	Pichia anomala PAEXG2 gene, strain K	AJ222862	4.25E-23
SI00233298	cytochrome P450 [Arabidopsis thaliana] >gi|13878375|sp|Q9FLC8|C792_ARATH Cytochrome P450 79A2	BAB09969	1.48E-46
SI00233299	Unknown	No Hit	null
SI00233300	N.tabacum mRNA for precursor of photosysten II 22 kDa protein	X84225	3.67E-35
SI00233301	Arabidopsis thaliana chromosome 3 CHR3v07142002 genomic sequence	NM_113243	2.92E-14
SI00233302	unknown protein [Arabidopsis thaliana]	AAK59627	1.87E-36
SI00233303	Unknown	No Hit	null
SI00233304	Lettuce mRNA for light-harvesting chlorophyll a/b-binding protein (LHCP), complete cds	D14002	0
SI00233305	SDAT(At) [Arabidopsis thaliana]	CAC19843	8.11E-82
SI00233306	Arabidopsis thaliana clone 6135 mRNA, complete sequence	AY088358	0
SI00233307	Arabidopsis thaliana clone C104705 putative myosin (At3g19960) mRNA, complete cds	BT001941	5.27E-25
SI00233308	Unknown	No Hit	null
SI00233309	G.hirsutum cab gene for chlorophyll ab binding protein	X54090	4.50E-87

SI00233310	putative AP2 domain transcription factor [Arabidopsis thaliana] >gi|2642430|gb|AAB87098.1| putative AP2 domain transcription factor [Arabidopsis thaliana] >gi|7487249|pir||T00498 probable AP2 domain transcription factor At2g23340 [imported] - Arabidopsis thaliana >gi|28372974|gb|AAO39969.1| At2g233	BAC43099	5.29E-36
SI00233311	Unknown	No Hit	null
SI00233312	Populus tremula x Populus tremuloides pumilio domain-containing protein PPD1 (PPD1) mRNA, complete cds	AF153276	2.73E-82
SI00233313	P0460C04.11 [Oryza sativa (japonica cultivar-group)]	BAB92919	8.01E-72
SI00233314	Nicotiana tabacum mRNA for hydroxycinnamoyl transferase (hct gene)	AJ507825	5.41E-43
SI00233315	Nicotiana tabacum beta-carbonic anhydrase (CA) mRNA, complete cds; nuclear gene for chloroplast product	AF454759	1.76E-15
SI00233316	Paulownia kawakamii superoxide dismutase (SOD5) mRNA, complete cds	AF037359	2.66E-57
SI00233317	Arabidopsis thaliana chromosome 4 CHR4v07142002 genomic sequence	NM_118937	7.01E-14
SI00233318	hypothetical protein [Trichodesmium erythraeum IMS101]	ZP_00075008	4.22E-30
SI00233319	Arabidopsis thaliana chromosome 5 CHR5v07142002 genomic sequence	NM_122251	3.19E-13
SI00233320	P.sativum atpC mRNA for gamma subunit of ATP synthase	X63604	1.79E-49
SI00233321	unknown protein [Arabidopsis thaliana] >gi|15229471|ref|NP_187413.1| hypothetical protein; protein id: At3g07565.1 [Arabidopsis thaliana]	AAO22562	5.64E-19
SI00233322	unknown [Arabidopsis thaliana]	AAM65279	1.89E-22
SI00233323	Spinacia oleracea chloroplast ribosomal protein S1 (rpS1) mRNA, complete cds	M82923	1.81E-86
SI00233324	P.sativum Cab-8 gene for photosystemm II chlorophyll a/b binding protein	X56538	0
SI00233325	hypothetical protein F28J12.200 - Arabidopsis thaliana >gi|7268647|emb|CAB78856.1| hypothetical protein [Arabidopsis thaliana] >gi|2832659|emb|CAA16734.1| hypothetical protein [Arabidopsis thaliana] >gi|15233920|ref|NP_193589.1| hypothetical protein; protein id: At4g18540.1 [Arabidopsis thaliana]	T04550	3.51E-59
SI00233326	Nicotiana tabacum beta-carbonic anhydrase (CA) mRNA, complete cds; nuclear gene for chloroplast product	AF454759	1.92E-15
SI00233327	Unknown protein [Arabidopsis thaliana] >gi|25344189|pir||B86304 hypothetical protein F6I1.12 [imported] - Arabidopsis thaliana >gi|9802776|gb|AAF99845.1|AC051629_12 Unknown protein [Arabidopsis thaliana] >gi|18394414|ref|NP_564010.1| expressed protein; protein id: At1g16880.1, supported by cDNA: gi	AAK62433	1.53E-78
SI00233328	Lotus japonicus genomic DNA, chromosome 6, clone:LjT14B06, TM0045, complete sequence	AP004517	2.98E-14
SI00233329	Unknown	No Hit	null
SI00233330	Arabidopsis thaliana chromosome 5 CHR5v07142002 genomic sequence	NM_122767	3.07E-103
SI00233331	Crocus sativus phytoene desaturase (pds) mRNA, complete cds; nuclear gene for chloroplast product	AY183118	1.15E-13
SI00233332	Fagus crenata mRNA for chlorophyll a/b-binding protein, complete cds	AB006081	1.28E-96
SI00233333	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	1.68E-29
SI00233334	Arabidopsis thaliana chromosome 4 CHR4v07142002 genomic sequence	NM_119300	5.09E-18
SI00233335	G.hirsutum cab gene for chlorophyll ab binding protein	X54090	5.34E-99
SI00233336	Unknown	No Hit	null
SI00233337	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	8.07E-39
SI00233338	probable glutaredoxin - Arabidopsis thaliana >gi|7268315|emb|CAB78609.1| glutaredoxin [Arabidopsis thaliana] >gi|2244923|emb|CAB10345.1| glutaredoxin [Arabidopsis thaliana] >gi|15234673|ref|NP_193302.1| glutaredoxin; protein id: At4g15670.1 [Arabidopsis thaliana]	G71421	1.43E-30
SI00233339	Arabidopsis thaliana clone 27681 mRNA, complete sequence	AY086788	1.48E-27

SI00233340	Lycopersicon esculentum mRNA binding protein precursor, mRNA, nuclear gene encoding chloroplast protein, complete cds	AF106660	6.84E-80
SI00233341	putative protein; protein id: At5g20190.1 [Arabidopsis thaliana]	NP_197519	6.21E-35
SI00233342	N.plumbaginifolia chloroplast glutamine synthetase gene, complete cds	M19055	1.95E-52
SI00233343	Arabidopsis thaliana At1g02640/T14P4_11 mRNA, complete cds	AY143952	5.48E-25
SI00233344	Arabidopsis thaliana At5g59380 gene, complete cds	BT004606	2.20E-11
SI00233345	Cotton mRNA for cottonseed catalase subunit 1 (EC 1.11.1.6)	X52135	0
SI00233346	Lotus japonicus genomic DNA, chromosome 2, clone:LjT10N22, TM0381, complete sequence	AP004577	1.15E-13
SI00233347	Zea mays PCO065577 mRNA sequence	AY109214	3.11E-21
SI00233348	Zea mays PCO127219 mRNA sequence	AY109207	5.08E-18
SI00233349	Unknown	No Hit	null
SI00233350	Zea mays PCO112338 mRNA sequence	AY105049	6.07E-71
SI00233351	Unknown	No Hit	null
SI00233352	Unknown	No Hit	null
SI00233353	Phaseolus vulgaris NBS-LRR resistance-like protein J78 (J78) gene, complete cds	AF306506	4.85E-13
SI00233354	unknown [Arabidopsis thaliana]	AAM64770	1.27E-65
SI00233355	N.tabacum LHC-I mRNA for photosystem I light-harvesting chlorophyll a/b-binding protein	X64198	7.44E-86
SI00233356	Beta vulgaris phophate translocator (pt) mRNA, partial cds; nuclear gene for chloroplast product	AF173679	6.19E-74
SI00233357	Arabidopsis thaliana clone RAFL15-09-E23 (R20450) At3g23790 mRNA, complete sequence	BT004051	4.47E-13
SI00233358	Glycine max biotin carboxyl carrier protein subunit precursor (accB-2) mRNA, complete cds; chloroplast gene for chloroplast product	AF271071	4.03E-20
SI00233359	V.radiata mRNA for pectinacetylesterase	X99348	7.88E-21
SI00233360	Glycine max mRNA for magnesium chelatase subunit	AJ001091	3.07E-17
SI00233361	Arabidopsis thaliana putative heat shock factor protein hsf8 (At1g12800) mRNA, complete cds	AY045991	1.63E-20
SI00233362	Vicia faba ferredoxin NADP+ reductase precursor (fnr) mRNA, complete cds	U14956	2.84E-119
SI00233363	putative protein [Arabidopsis thaliana] >gi|26452044|dbj|BAC43112.1| unknown protein [Arabidopsis thaliana] >gi|22531074|gb|AAM97041.1| unknown protein [Arabidopsis thaliana] >gi|23197924|gb|AAN15489.1| unknown protein [Arabidopsis thaliana]	CAC34500	3.06E-74
SI00233364	G.hirsutum cab gene for chlorophyll ab binding protein	X54090	1.32E-96
SI00233365	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	1.19E-31
SI00233366	Arabidopsis thaliana At2g19560 mRNA for unknown protein, complete cds, clone: RAFL21-23-K10	AK118897	3.34E-23
SI00233367	Arabidopsis thaliana At2g42600 mRNA sequence	AY074346	3.30E-97
SI00233368	hypothetical protein [Arabidopsis thaliana] >gi|25408647|pir||D84823 hypothetical protein At2g39950 [imported] - Arabidopsis thaliana >gi|15225598|ref|NP_181524.1| hypothetical protein; protein id: At2g39950.1 [Arabidopsis thaliana]	AAB95280	3.50E-54
SI00233369	Arabidopsis thaliana chromosome 5 CHR5v07142002 genomic sequence	NM_121242	7.83E-13
SI00233370	Petunia gene for chlorophyll a/b binding protein cab 25	X02358	5.48E-74
SI00233371	Lycopersicon pennellii 2-isopropylmalate synthase (lp-ipmsa) mRNA, complete cds	AF004165	7.79E-21
SI00233372	Unknown	No Hit	null

SI00233373	Unknown	No Hit	null
SI00233374	Unknown	No Hit	null
SI00233375	Arabidopsis thaliana chromosome 2 CHR2v07142002 genomic sequence	NM_129078	8.37E-24
SI00233376	Unknown	No Hit	null
SI00233377	G.hirsutum cab gene for chlorophyll ab binding protein	X54090	3.15E-94
SI00233378	Arabidopsis thaliana chromosome 1 CHR1v07142002 genomic sequence	NM_104543	5.87E-73
SI00233379	Lettuce mRNA for light-harvesting chlorophyll a/b-binding protein (LHCP), complete cds	D14002	0
SI00233380	Arabidopsis thaliana chromosome 3 BAC F11F8 genomic sequence, complete sequence	AC016661	9.75E-90
SI00233381	Unknown protein [Arabidopsis thaliana] >gi|22330957|ref|NP_187641.2| unknown protein; protein id: At3g10300.1, supported by cDNA: gi_17064843 [Arabidopsis thaliana]	AAL32576	2.60E-62
SI00233382	Fagus crenata mRNA for chlorophyll a/b-binding protein, complete cds	AB006081	1.35E-55
SI00233383	P.sativum Cab-8 gene for photosystemm II chlorophyll a/b binding protein	X56538	2.36E-72
SI00233384	Arabidopsis thaliana chromosome 1 CHR1v07142002 genomic sequence	NM_100563	3.52E-16
SI00233385	Bruguiera gymnorhiza HPR mRNA for hydroxypyruvate reductase, complete cds	AB060810	0
SI00233386	Lactuca sativa methionine sulfoxide reductase mRNA, partial cds	AF162204	6.89E-27
SI00233387	Populus tremula x Populus tremuloides thioredoxin H mRNA, complete cds	AF483265	4.22E-12
SI00233388	L.esculentum mRNA for 33kDa precursor protein of oxygen-evolving complex	Z11999	1.90E-52
SI00233389	Arabidopsis thaliana Unknown protein (At3g52060) mRNA, complete cds	BT002508	7.77E-18
SI00233390	Lettuce mRNA for light-harvesting chlorophyll a/b-binding protein (LHCP), complete cds	D14002	0
SI00233391	Unknown	No Hit	null
SI00233392	contains ESTs AU162548(R3914),AU032339(R3914)~similar to Arabidopsis thaliana chromosome 3,At3g12260~unknown protein [Oryza sativa (japonica cultivar-group)] >gi|27261102|dbj|BAC45215.1| OJ1340_C08.30 [Oryza sativa (japonica cultivar-group)]	BAC16465	1.14E-52
SI00233393	Unknown	No Hit	null
SI00233394	Unknown	No Hit	null
SI00233395	P.sativum Cab-8 gene for photosystemm II chlorophyll a/b binding protein	X56538	0
SI00233396	Unknown	No Hit	null
SI00233397	lipase-like protein [Arabidopsis thaliana] >gi|18413291|ref|NP_567350.1| lipase-like protein; protein id: At4g10050.1, supported by cDNA: 6822., supported by cDNA: gi_14423411, supported by cDNA: gi_20148342 [Arabidopsis thaliana] >gi|14423412|gb|AAK62388.1|AF386943_1 lipase-like protein [Arabidops	AAM10062	5.83E-89
SI00233398	Mesembryanthemum crystallinum MPC9 mRNA for protein phosphatase 2C, complete cds	AB083482	7.43E-89
SI00233399	Arabidopsis thaliana clone U10380 putative acetyl-CoA synthetase (At5g36880) mRNA, complete cds	BT002371	4.41E-84
SI00233400	Unknown	No Hit	null
SI00233401	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	1.11E-31
SI00233402	Unknown	No Hit	null
SI00233403	Spinacia oleracea heat shock C70 protein mRNA, complete cds	L26243	0
SI00233404	Unknown	No Hit	null

SI00233405	Arabidopsis thaliana clone C105095 putative lysyl-tRNA synthetase (At3g13490) mRNA, complete cds	BT002317	2.20E-30
SI00233406	gene_id:MYF5.3~similar to unknown protein~sp|P49224 [Arabidopsis thaliana] >gi|18405254|ref|NP_566809.1| expressed protein; protein id: At3g27160.1, supported by cDNA: gi_14532553, supported by cDNA: gi_18655398 [Arabidopsis thaliana]	BAB01933	4.83E-16
SI00233407	phosphoglycerate kinase (EC 2.7.2.3) precursor, chloroplast - common tobacco >gi|1161600|emb|CAA88841.1| phosphoglycerate kinase [Nicotiana tabacum] >gi|2499497|sp|Q42961|PGKH_TOBAC Phosphoglycerate kinase, chloroplast precursor	T03660	4.73E-98
SI00233408	H.helix mRNA for light harvesting chlorophyll a/b binding protein	X68333	1.54E-55
SI00233409	Nicotiana tabacum beta-carbonic anhydrase (CA) mRNA, complete cds; nuclear gene for chloroplast product	AF454759	1.83E-15
SI00233410	Arabidopsis thaliana ribosomal protein L17-like protein (At3g54210) mRNA, complete cds	BT000413	2.78E-29
SI00233411	heterotrimeric G-protein gamma subunit 2 [Arabidopsis thaliana] >gi|14625852|gb|AAK71536.1|AF347077_1 heterotrimeric G-protein gamma subunit 2 [Arabidopsis thaliana]	AAK71537	7.33E-17
SI00233412	Arabidopsis thaliana clone 94409 mRNA, complete sequence	AY088757	6.24E-40
SI00233413	Lettuce mRNA for light-harvesting chlorophyll a/b-binding protein (LHCP), complete cds	D14002	0
SI00233414	Arabidopsis thaliana chromosome 5 CHR5v07142002 genomic sequence	NM_122251	9.11E-11
SI00233415	Populus tremula x Populus tremuloides thioredoxin H mRNA, complete cds	AF483265	4.52E-12
SI00233416	Unknown	No Hit	null
SI00233417	Tabacco Cab16 mRNA for major chlorophyll a/b binding protein	X52741	1.85E-92
SI00233418	phosphoglycerate kinase (EC 2.7.2.3) precursor, chloroplast - common tobacco >gi|1161600|emb|CAA88841.1| phosphoglycerate kinase [Nicotiana tabacum] >gi|2499497|sp|Q42961|PGKH_TOBAC Phosphoglycerate kinase, chloroplast precursor	T03660	1.36E-110
SI00233419	Rhizoctonia solani virus putative capsid protein mRNA, complete cds	AF133291	2.84E-23
SI00233420	Contains similarity to an unknown protein F7A7_100 gi|7327817 from Arabidopsis thaliana BAC F7A7 gb|AL161946. ESTs gb|N65842, gb|F19836 and gb|AI993679 come from this gene >gi|25406953|pir||F86205 hypothetical protein [imported] - Arabidopsis thaliana	AAF82212	1.08E-52
SI00233421	Nicotiana tabacum bZIP transcription factor BZI-2 mRNA, complete cds	AY045570	5.06E-22
SI00233422	Arabidopsis thaliana chromosome 4 CHR4v07142002 genomic sequence	NM_119215	2.82E-11
SI00233423	Arabidopsis thaliana putative RNA-binding protein (At1g09340) mRNA, complete cds	AY070022	7.08E-49
SI00233424	Zea mays PCO062733 mRNA sequence	AY104180	3.01E-128
SI00233425	Arabidopsis thaliana chromosome 3 CHR3v07142002 genomic sequence	NM_148744	7.53E-39
SI00233426	Arabidopsis thaliana chromosome 4 CHR4v07142002 genomic sequence	NM_119954	4.15E-118
SI00233427	Arabidopsis thaliana unknown protein (At5g65840) mRNA, complete cds	BT000333	2.85E-14
SI00233428	Nicotiana tabacum beta-carbonic anhydrase (CA) mRNA, complete cds; nuclear gene for chloroplast product	AF454759	7.59E-18
SI00233429	Zea mays PCO098989 mRNA sequence	AY103717	1.72E-46
SI00233430	Tomato CAB-8 gene (constucted from mRNA and DNA) for type III chlorophyll a/b binding polypeptide of photosystem I	X15258	8.28E-27
SI00233431	Lotus japonicus phosphatidylinositol transfer-like protein III (LjPLP-III) mRNA, complete cds	AF367433	5.22E-28
SI00233432	Vicia faba CREB-like protein mRNA, complete cds	M81827	1.84E-15
SI00233433	Nicotiana tabacum beta-carbonic anhydrase (CA) mRNA, complete cds; nuclear gene for chloroplast product	AF454759	1.86E-15
SI00233434	Arabidopsis thaliana putative RNA-binding protein (At1g09340) mRNA, complete cds	AY070022	5.00E-56
SI00233435	Unknown	No Hit	null
SI00233436	Glycine max mRNA for magnesium chelatase subunit	AJ001091	2.30E-33

SI00233437	P.sativum Cab II gene for chlorophyll a/b-binding protein	X57082	1.27E-93
SI00233438	At5g21040/T10F18_70 [Arabidopsis thaliana] >gi|15081723|gb|AAK82516.1| T10F18_70/T10F18_70 [Arabidopsis thaliana] >gi|18420385|ref|NP_568408.1| F-box protein family; protein id: At5g21040.1, supported by cDNA: gi_15081722 [Arabidopsis thaliana]	AAO11609	5.08E-34
SI00233439	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	6.22E-30
SI00233440	Arabidopsis thaliana At2g31890/F20M17.7 mRNA, complete cds	AY065010	8.11E-16
SI00233441	Nicotiana tabacum beta-carbonic anhydrase (CA) mRNA, complete cds; nuclear gene for chloroplast product	AF454759	1.82E-15
SI00233442	unknown [Arabidopsis thaliana]	AAM63014	5.36E-28
SI00233443	D.stramonium mRNA for S-adenosylmethionine decarboxylase	Y07768	7.64E-21
SI00233444	hypothetical protein T2K10.4 [imported] - Arabidopsis thaliana >gi|4249378|gb|AAD14475.1|AAD14475 Identical to gb|AJ010471 mRNA for DEAD box RNA helicase (RH22) from Arabidopsis thaliana. EST gb|Y11191 comes from this gene >gi|3776015|emb|CAA09210.1| RNA helicase [Arabidopsis thaliana] >gi|1135880	B96624	4.55E-61
SI00233445	Unknown	No Hit	null
SI00233446	Spinacia oleracea heat shock C70 protein mRNA, complete cds	L26243	0
SI00233447	A.thaliana mRNA for carbonic anhydrase	X65541	2.45E-11
SI00233448	expressed protein; protein id: At2g34050.1, supported by cDNA: 23892., supported by cDNA: gi_15450477 [Arabidopsis thaliana] >gi|20196905|gb|AAB67623.2| expressed protein [Arabidopsis thaliana] >gi|15450478|gb|AAK96532.1| At2g34050/T14G11.17 [Arabidopsis thaliana] >gi|24797040|gb|AAN64532.1| At2g34	NP_565778	2.03E-73
SI00233449	Nicotiana paniculata mRNA for plastidic aldolase NPALDP1, complete cds	AB027001	5.00E-93
SI00233450	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	1.76E-43
SI00233451	Unknown	No Hit	null
SI00233452	Unknown	No Hit	null
SI00233453	unknown protein [Arabidopsis thaliana] >gi|22329926|ref|NP_174638.2| chloroplast lumen common protein family; protein id: At1g33780.1, supported by cDNA: gi_17065471, supported by cDNA: gi_20148514 [Arabidopsis thaliana] >gi|17065472|gb|AAL32890.1| Unknown protein [Arabidopsis thaliana]	AAM10148	3.05E-38
SI00233454	H+-transporting two-sector ATPase (EC 3.6.3.14) beta chain, mitochondrial - Para rubber tree >gi|18831|emb|CAA41401.1| mitochondrial ATP synthase beta-subunit [Hevea brasiliensis] >gi|231586|sp|P29685|ATP2_HEVBR ATP synthase beta chain, mitochondrial precursor	S20504	1.95E-35
SI00233455	Vigna radiata chlorophyll a/b binding protein CP29 (CipCp29) mRNA, complete cds; nuclear gene for chloroplast product	AF139466	2.26E-36
SI00233456	putative AIM1 protein [Arabidopsis thaliana] >gi|7269799|emb|CAB79659.1| AIM1 protein [Arabidopsis thaliana] >gi|20465649|gb|AAM20293.1| putative AIM1 protein [Arabidopsis thaliana] >gi|4972047|emb|CAB43915.1| AIM1 protein [Arabidopsis thaliana] >gi|16648891|gb|AAL24297.1| AIM1 protein [Arabidopsis	AAL59895	6.23E-94
SI00233457	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	7.33E-30
SI00233458	Sedum lineare prxQ mRNA for peroxiredoxin Q, partial cds	AB037598	2.03E-27
SI00233459	Unknown	No Hit	null
SI00233460	Prunus armeniaca porin (mPOR) mRNA, nuclear gene encoding mitochondrial product, complete cds	AF139498	2.55E-39
SI00233461	Gossypium hirsutum leucine-rich repeat resistance protein-like protein mRNA, complete cds	AY040533	2.44E-73
SI00233462	Putative non-LTR retroelement reverse transcriptase [Oryza sativa (japonica cultivar-group)]	AAM47629	4.45E-14
SI00233463	Unknown	No Hit	null
SI00233464	Arabidopsis thaliana chromosome 3 CHR3v07142002 genomic sequence	NM_111258	5.90E-18

SI00233465	Prunus persica mRNA for plasma membrane H+ ATPase (PPA2 gene)	AJ271438	3.83E-109
SI00233466	myo-inositol 1-phosphate synthase [Sesamum indicum] >gi|14548095|sp|Q9FYV1|INO1_SESIN Inositol-3-phosphate synthase (Myo-inositol-1-phosphate synthase) (MI-1-P synthase) (IPS)	AAG01148	3.21E-98
SI00233467	P.sativum ApxI mRNA for ascorbate peroxidase	X62077	8.30E-30
SI00233468	Spinacia oleracea mRNA for cytosolic glucose-6-phosphate isomerase	AJ000266	1.56E-40
SI00233469	OSJNBb0021A09.5 [Oryza sativa (japonica cultivar-group)]	BAB89453	7.35E-53
SI00233470	serine protease-like protein [Nicotiana tabacum]	BAC53929	2.23E-55
SI00233471	ribulose-1,5-bisphosphate carboxylase/oxygenase activase 1 [Gossypium hirsutum]	AAG61120	8.95E-128
SI00233472	Arabidopsis thaliana chromosome 4 CHR4v07142002 genomic sequence	NM_120173	2.40E-44
SI00233473	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	5.94E-39
SI00233474	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	4.78E-34
SI00233475	Unknown	No Hit	null
SI00233476	expressed protein; protein id: At2g35680.1, supported by cDNA: gi_15450955, supported by cDNA: gi_17528969 [Arabidopsis thaliana] >gi|20197531|gb|AAD15447.2| expressed protein [Arabidopsis thaliana] >gi|17528970|gb|AAL38695.1| unknown protein [Arabidopsis thaliana]	NP_565816	1.65E-41
SI00233477	unknown protein [Arabidopsis thaliana] >gi|18420198|ref|NP_568037.1| Expressed protein; protein id: At4g38225.1, supported by cDNA: 39781. [Arabidopsis thaliana] >gi|21593523|gb|AAM65490.1| unknown [Arabidopsis thaliana]	AAN12956	1.06E-69
SI00233478	Lettuce mRNA for light-harvesting chlorophyll a/b-binding protein (LHCP), complete cds	D14002	0
SI00233479	Arabidopsis thaliana unknown protein (At5g67370) mRNA, complete cds	AY051028	3.13E-23
SI00233480	Medicago sativa glyoxysomal malate dehydrogenase precursor (gmdh) mRNA, complete cds	AF020270	9.51E-73
SI00233481	Papaver somniferum NADPH-dependent codeinone reductase (cor1) mRNA, cor1-2 allele, complete cds	AF108433	1.84E-52
SI00233482	Unknown	No Hit	null
SI00233483	Unknown protein [Arabidopsis thaliana] >gi|22330957|ref|NP_187641.2| unknown protein; protein id: At3g10300.1, supported by cDNA: gi_17064843 [Arabidopsis thaliana]	AAL32576	2.54E-62
SI00233484	expressed protein; protein id: At4g29070.1, supported by cDNA: 249331. [Arabidopsis thaliana]	NP_567821	2.05E-24
SI00233485	putative protein [Arabidopsis thaliana] >gi|20466552|gb|AAM20593.1| putative protein [Arabidopsis thaliana] >gi|18409862|ref|NP_566986.1| chloroplast lumen common protein family; protein id: At3g53560.1, supported by cDNA: 36212., supported by cDNA: gi_20466551 [Arabidopsis thaliana]	AAN15593	2.35E-39
SI00233486	L.esculentum DNA for CAB11 gene	X57706	6.90E-20
SI00233487	AWJL236 [Triticum aestivum] >gi|1076764|pir||S49300 AWJL236 protein - wheat	CAA57135	1.94E-29
SI00233488	chlorophyll a/b-binding protein - garden pea >gi|20671|emb|CAA49149.1| chlorophyll a/b-binding protein [Pisum sativum]	S33775	1.32E-32
SI00233489	Unknown	No Hit	null
SI00233490	Arabidopsis thaliana chromosome 4 CHR4v07142002 genomic sequence	NM_119954	1.80E-126
SI00233491	Unknown	No Hit	null
SI00233492	Unknown	No Hit	null
SI00233493	Arabidopsis thaliana chromosome 4 CHR4v07142002 genomic sequence	NM_119288	1.07E-81
SI00233494	Putative ATFP3 [Oryza sativa]	AAM01036	3.42E-27

SI00233495	acetyl-CoA synthetase, putative; 45051-31547 [Arabidopsis thaliana] >gi|25405817|pir||D96595 probable acetyl-CoA synthetase, 45051-31547 [imported] - Arabidopsis thaliana >gi|15222694|ref|NP_175929.1| acetyl-CoA synthetase, putative; protein id: At1g55320.1 [Arabidopsis thaliana]	AAG51574	4.87E-19
SI00233496	expressed protein; protein id: At1g22700.1, supported by cDNA: 120133. [Arabidopsis thaliana]	NP_683322	6.49E-34
SI00233497	unknown [Arabidopsis thaliana]	AAM61533	1.10E-22
SI00233498	Cicer arietinum mRNA for Trans-Cinnamate 4-Monooxygenase	AJ007449	9.36E-36
SI00233499	Lettuce mRNA for light-harvesting chlorophyll a/b-binding protein (LHCP), complete cds	D14002	0
SI00233500	Tomato CAB-8 gene (constucted from mRNA and DNA) for type III chlorophyll a/b binding polypeptide of photosystem I	X15258	1.95E-57
SI00233501	polygalacturonase, putative; protein id: At3g62110.1, supported by cDNA: gi_13358184 [Arabidopsis thaliana] >gi|11762132|gb|AAG40344.1|AF324992_1 AT3g62110 [Arabidopsis thaliana] >gi|23397166|gb|AAN31866.1| unknown protein [Arabidopsis thaliana]	NP_567126	8.20E-17
SI00233502	Unknown	No Hit	null
SI00233503	Vigna radiata chlorophyll a/b binding protein CP29 (CipCp29) mRNA, complete cds; nuclear gene for chloroplast product	AF139466	1.62E-40
SI00233504	Unknown	No Hit	null
SI00233505	Medicago sativa glyoxysomal malate dehydrogenase precursor (gmdh) mRNA, complete cds	AF020270	5.35E-62
SI00233506	Nicotiana paniculata mRNA for plastidic aldolase NPALDP1, complete cds	AB027001	6.76E-80
SI00233507	unknown protein [Arabidopsis thaliana] >gi|18397852|ref|NP_566299.1| expressed protein; protein id: At3g07140.1, supported by cDNA: gi_15292900 [Arabidopsis thaliana] >gi|22136838|gb|AAM91763.1| unknown protein [Arabidopsis thaliana]	AAK92821	2.05E-83
SI00233508	ADP,ATP carrier protein, mitochondrial precursor (ADP/ATP translocase) (Adenine nucleotide translocator) (ANT) >gi|100422|pir||S21974 ADP,ATP carrier protein, ant - potato >gi|21407|emb|CAA44054.1| ADP /ATP translocator [Solanum tuberosum]	P25083	5.87E-104
SI00233509	Prunus persica mRNA for plasma membrane H+ ATPase (PPA2 gene)	AJ271438	0
SI00233510	L.esculentum S-adenosyl-L-methionine synthetase mRNA, complete CDS	Z24743	0
SI00233511	Haynaldia villosa clone kong32 mRNA	AF498269	2.07E-11
SI00233512	Arabidopsis thaliana chromosome 1 CHR1v07142002 genomic sequence	NM_103681	1.70E-55
SI00233513	Mesembryanthemum crystallinum water channel protein MipK (MipK) mRNA, complete cds	AF133532	4.63E-53
SI00233514	short chain alcohol dehydrogenase-like; protein id: At5g06060.1, supported by cDNA: 111427. [Arabidopsis thaliana] >gi|27754526|gb|AAO22710.1| putative short chain alcohol dehydrogenase [Arabidopsis thaliana] >gi|28394081|gb|AAO42448.1| putative short chain alcohol dehydrogenase [Arabidopsis thalia	NP_196225	2.20E-87
SI00233515	C.arietinum mRNA for beta-tubulin	X98406	0
SI00233516	Unknown	No Hit	null
SI00233517	Zea mays PCO091925 mRNA sequence	AY106973	8.40E-24
SI00233518	Petunia hybrida thiol protease homolog P21 (PeTh3) mRNA, complete cds	U31094	3.47E-26
SI00233519	Unknown	No Hit	null
SI00233520	Arabidopsis thaliana chromosome 3 CHR3v07142002 genomic sequence	NM_111363	9.41E-70
SI00233521	chloroplast phosphoglycerate kinase [Populus nigra]	BAA33803	3.04E-16
SI00233522	Mesembryanthemum crystallinum ferredoxin-NADP+ reductase precursor, mRNA, complete cds	M25528	1.66E-77
SI00233523	Arabidopsis thaliana chromosome 1 CHR1v07142002 genomic sequence	NM_101917	2.98E-14
SI00233524	Arabidopsis thaliana clone 11684 mRNA, complete sequence	AY084755	1.05E-32

SI00233525	Lotus japonicus genomic DNA, chromosome 1, clone:LjT13B22, TM0009, complete sequence	AP004474	4.41E-29
SI00233526	Unknown protein [Arabidopsis thaliana] >gi|22330957|ref|NP_187641.2| unknown protein; protein id: At3g10300.1, supported by cDNA: gi_17064843 [Arabidopsis thaliana]	AAL32576	8.24E-54
SI00233527	Unknown	No Hit	null
SI00233528	Unknown	No Hit	null
SI00233529	Nicotiana tabacum mRNA for thioredoxin peroxidase	AJ309009	7.70E-89
SI00233530	unknown [Arabidopsis thaliana] >gi|3395441|gb|AAC28773.1| expressed protein [Arabidopsis thaliana] >gi|15809748|gb|AAL06802.1| At2g38310/T19C21.20 [Arabidopsis thaliana] >gi|14517502|gb|AAK62641.1| At2g38310/T19C21.20 [Arabidopsis thaliana] >gi|7487162|pir||T02514 hypothetical protein At2g38310 [im	AAM64704	9.63E-71
SI00233531	Arabidopsis thaliana chromosome I BAC F1B16 genomic sequence, complete sequence	AC023754	4.20E-15
SI00233532	putative RNA helicase [Arabidopsis thaliana] >gi|9293865|dbj|BAB01768.1| DEAD-Box RNA helicase-like protein [Arabidopsis thaliana] >gi|15228722|ref|NP_188870.1| putative RNA helicase; protein id: At3g22310.1, supported by cDNA: gi_20268675 [Arabidopsis thaliana]	AAM14042	4.68E-59
SI00233533	Tomato CAB-8 gene (constucted from mRNA and DNA) for type III chlorophyll a/b binding polypeptide of photosystem I	X15258	3.23E-57
SI00233534	Nicotiana tabacum beta-carbonic anhydrase (CA) mRNA, complete cds; nuclear gene for chloroplast product	AF454759	1.91E-15
SI00233535	Unknown	No Hit	null
SI00233536	Nicotiana tabacum beta-carbonic anhydrase (CA) mRNA, complete cds; nuclear gene for chloroplast product	AF454759	1.91E-15
SI00233537	Arabidopsis thaliana chromosome 2 CHR2v07142002 genomic sequence	NM_128327	8.03E-58
SI00233538	Unknown	No Hit	null
SI00233539	N.tabacum mRNA (T23-5B) for PSII 23-kDa polypeptide	X62426	3.12E-20
SI00233540	putative protein; protein id: At5g52540.1, supported by cDNA: gi_17064941 [Arabidopsis thaliana] >gi|8953709|dbj|BAA98072.1| contains similarity to unknown protein~dbj|BAA90637.1~gene_id:F6N7.1 [Arabidopsis thaliana] >gi|24899799|gb|AAN65114.1| Unknown protein [Arabidopsis thaliana] >gi|17064942|gb	NP_200067	3.54E-52
SI00233541	Unknown	No Hit	null
SI00233542	Populus nigra PnChlPGK mRNA for chloroplast phosphoglycerate kinase, complete cds	AB018412	0
SI00233543	small GTP-binding protein >gi|547478|emb|CAA85733.1| guanine nucleotide regulatory protein [Vicia faba] >gi|1076545|pir||S49225 guanine nucleotide regulatory protein - fava bean	2115367E	7.82E-46
SI00233544	3-methyladenine DNA glycosylase, putative; 31680-30045 [Arabidopsis thaliana] >gi|25406423|pir||F96782 hypothetical protein F22H5.5 [imported] - Arabidopsis thaliana	AAG12687	9.89E-50
SI00233545	Unknown	No Hit	null
SI00233546	N.tabacum mRNA for precursor of photosysten II 22 kDa protein	X84225	3.55E-29
SI00233547	Expressed protein; protein id: At1g09815.1, supported by cDNA: 6642. [Arabidopsis thaliana]	NP_563854	2.36E-29
SI00233548	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	7.46E-30
SI00233549	A.hypochondriacus mRNA for seed specific protein of balanced nutritional quality	Z11577	1.15E-11
SI00233550	L.esculentum psbX mRNA for photosystem II 23 kDa protein	X63007	7.61E-18
SI00233551	putative protein; protein id: At5g19540.1, supported by cDNA: 120101. [Arabidopsis thaliana] >gi|21537101|gb|AAM61442.1| unknown [Arabidopsis thaliana] >gi|22136824|gb|AAM91756.1| unknown protein [Arabidopsis thaliana] >gi|15292823|gb|AAK92780.1| unknown protein [Arabidopsis thaliana]	NP_197455	1.47E-77
SI00233552	Arabidopsis thaliana unknown protein (At5g52560) mRNA, complete cds	AY040035	1.31E-22

SI00233553	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	3.22E-41
SI00233554	Castanea sativa IDS-4-like protein mRNA, partial cds	AY055748	5.80E-28
SI00233555	Unknown	No Hit	null
SI00233556	Nicotiana paniculata mRNA for plastidic aldolase NPALDP1, complete cds	AB027001	6.74E-80
SI00233557	D.stramonium mRNA for S-adenosylmethionine decarboxylase	Y07768	8.16E-21
SI00233558	Arabidopsis thaliana At1g15340/F9L1_28 mRNA, complete cds	AY094439	2.51E-30
SI00233559	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	7.22E-30
SI00233560	Tomato photosystem I (PS I) reaction center protein subunit II (psaD) mRNA, complete cds	M21344	1.60E-44
SI00233561	Unknown	No Hit	null
SI00233562	S-adenosylmethionine decarboxylase leader [Narcissus pseudonarcissus]	AAO43185	6.44E-12
SI00233563	unknown [Arabidopsis thaliana] >gi|20198171|gb|AAM15442.1| predicted protein [Arabidopsis thaliana] >gi|19699178|gb|AAL90955.1| At2g04039/At2g04039 [Arabidopsis thaliana] >gi|16604302|gb|AAL24157.1| unknown protein [Arabidopsis thaliana] >gi|18395648|ref|NP_565308.1| unknown protein; protein id: At	AAM63915	4.01E-36
SI00233564	P0479C08.37 [Oryza sativa (japonica cultivar-group)]	BAC24948	1.38E-18
SI00233565	Spinacia oleracea plastid-specific ribosomal protein 3 precursor (Psrp-3) mRNA, complete cds; nuclear gene for chloroplast product	AF239218	6.79E-27
SI00233566	Unknown	No Hit	null
SI00233567	EF-hand Calcium binding protein-like [Arabidopsis thaliana] >gi|22326598|ref|NP_196037.2| EF - hand Calcium binding protein - like; protein id: At5g04170.1, supported by cDNA: gi_19698990 [Arabidopsis thaliana] >gi|19698991|gb|AAL91231.1| EF-hand calcium binding protein-like [Arabidopsis thaliana]	CAC05499	8.88E-34
SI00233568	hypothetical protein; protein id: At1g25510.1, supported by cDNA: gi_20466515 [Arabidopsis thaliana] >gi|25518510|pir||D86385 hypothetical protein F2J7.6 - Arabidopsis thaliana >gi|20466516|gb|AAM20575.1| unknown protein [Arabidopsis thaliana] >gi|12321511|gb|AAG50814.1|AC079281_16 hypothetical pro	NP_173922	2.60E-66
SI00233569	Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:K8K14	AB007645	6.35E-85
SI00233570	unknown [Arabidopsis thaliana]	AAM65209	1.08E-38
SI00233571	unknown protein [Arabidopsis thaliana] >gi|15231121|ref|NP_190774.1| putative protein; protein id: At3g52060.1, supported by cDNA: gi_13430781, supported by cDNA: gi_13877702, supported by cDNA: gi_15293266 [Arabidopsis thaliana] >gi|13877703|gb|AAK43929.1|AF370610_1 putative protein [Arabidopsis t	AAK26013	5.52E-53
SI00233572	Unknown	No Hit	null
SI00233573	unknown protein [Arabidopsis thaliana]	BAC43364	8.50E-13
SI00233574	Arabidopsis thaliana RAFL04-14-M02 mRNA sequence	AY091157	9.93E-27
SI00233575	Unknown	No Hit	null
SI00233576	expressed protein [Arabidopsis thaliana] >gi|14596225|gb|AAK68840.1| Unknown protein [Arabidopsis thaliana] >gi|18398083|ref|NP_565386.1| expressed protein; protein id: At2g16350.1, supported by cDNA: gi_14596224 [Arabidopsis thaliana]	AAD22304	9.42E-15
SI00233577	putative protein [Arabidopsis thaliana] >gi|11283451|pir||T48306 hypothetical protein F9G14.160 - Arabidopsis thaliana >gi|15242555|ref|NP_195905.1| putative protein; protein id: At5g02850.1 [Arabidopsis thaliana]	CAB86039	1.74E-34
SI00233578	Arabidopsis thaliana putative J8 protein (At1g80920) mRNA, complete cds	AF332461	4.64E-13
SI00233579	Nicotiana sylvestris Lhcb1*8 gene for light harvesting chlorophyll a/b-binding protein, complete cds	AB012640	5.47E-99
SI00233580	Mesembryanthemum crystallinum methionine synthase (MetE) mRNA, complete cds	U84889	5.69E-108
SI00233581	Arabidopsis thaliana chromosome 2 CHR2v07142002 genomic sequence	NM_126585	1.23E-21

SI00233582	Arabidopsis thaliana clone C105458 auxin response factor 1 (At1g59750) mRNA, complete cds	BT002748	3.00E-17
SI00233583	Unknown	No Hit	null
SI00233584	Tobacco mRNA for 31 kDa chloroplast ribonucleoprotein	X53942	1.52E-34
SI00233585	Arabidopsis thaliana chromosome 1 BAC T2E12 genomic sequence, complete sequence	AC015986	1.29E-31
SI00233586	Arabidopsis thaliana chromosome 4 CHR4v07142002 genomic sequence	NM_119136	1.24E-22
SI00233587	Arabidopsis thaliana chromosome I BAC T28K15 genomic sequence, complete sequence	AC022522	4.73E-39
SI00233588	Nicotiana tabacum CTR mRNA, partial cds	AF247567	2.26E-28
SI00233589	Zantedeschia aethiopica glycolate oxidase (gox) mRNA, complete cds	AY173074	1.22E-90
SI00233590	Mus musculus, clone MGC:6623 IMAGE:3491066, mRNA, complete cds	BC002118	3.27E-60
SI00233591	N.tabacum LHC-I mRNA for photosystem I light-harvesting chlorophyll a/b-binding protein	X64198	6.80E-86
SI00233592	Unknown	No Hit	null
SI00233593	Unknown	No Hit	null
SI00233594	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	1.19E-31
SI00233595	RAB7D [Lotus japonicus]	CAA98171	4.09E-105
SI00233596	G.hirsutum cab gene for chlorophyll ab binding protein	X54090	1.09E-84
SI00233597	Bruguiera gymnorhiza HPR mRNA for hydroxypyruvate reductase, complete cds	AB060810	0
SI00233598	Arabidopsis thaliana DNA chromosome 3, BAC clone T8M16	AL390921	2.29E-18
SI00233599	Arabidopsis thaliana chromosome 3 CHR3v07142002 genomic sequence	NM_113790	2.13E-24
SI00233600	Unknown	No Hit	null
SI00233601	Arabidopsis thaliana At2g26500/T9J22.17 mRNA, complete cds	AY060541	8.72E-19
SI00233602	Medicago sativa NADH-glutamate synthase mRNA, comlete cds	L01660	3.87E-35
SI00233603	G.hirsutum cab gene for chlorophyll ab binding protein	X54090	1.71E-80
SI00233604	unknown protein [Arabidopsis thaliana] >gi|3702339|gb|AAC62896.1| expressed protein [Arabidopsis thaliana] >gi|18252159|gb|AAL61912.1| unknown protein [Arabidopsis thaliana] >gi|25341889|pir||E84898 hypothetical protein At2g46080 [imported] - Arabidopsis thaliana >gi|18406969|ref|NP_566063.1| expre	AAM47867	3.31E-65
SI00233605	Elaeis guineensis mRNA for metallothionein-like protein	AJ236913	3.55E-12
SI00233606	Cucumis melo aminotransferase 1 mRNA, complete cds	AY066012	8.18E-58
SI00233607	Unknown protein [Oryza sativa (japonica cultivar-group)]	AAN87739	3.53E-21
SI00233608	Zea mays PCO130447 mRNA sequence	AY108298	7.52E-15
SI00233609	pectin methylesterase [Nicotiana tabacum]	CAB95025	3.43E-68
SI00233610	Unknown	No Hit	null
SI00233611	Arabidopsis thaliana clone RAFL15-15-P17 (R20579)	BT004240	1.66E-43
SI00233612	subtilisin-like serine protease, putative; protein id: At3g14067.1 [Arabidopsis thaliana]	NP_566473	6.09E-71
SI00233613	Petunia x hybrida putative metallothionein-like protein (Mt2) mRNA, complete cds	AF201384	1.98E-11
SI00233614	Sesame mRNA for stearoyl-acyl carrier protein desaturase, complete cds, clone CDES01	D42086	1.00E-109
SI00233615	sugar transporter protein [Arabidopsis thaliana] >gi|15230212|ref|NP_188513.1| sugar transport, putative; protein id: At3g18830.1	BAB01812	1.30E-80

	[Arabidopsis thaliana]		
SI00233616	Arabidopsis thaliana chromosome 2 CHR2v07142002 genomic sequence	NM_127601	2.66E-79
SI00233617	unknown protein [Arabidopsis thaliana]	AAG51484	1.28E-80
SI00233618	Arabidopsis thaliana chromosome 1 CHR1v07142002 genomic sequence	NM_103828	1.89E-15
SI00233619	M.domestica ribulose-1,5-bisphosphate carboxylase/oxygenase activase mRNA	Z21794	7.73E-45
SI00233620	Medicago truncatula clone mth2-2b2, complete sequence	AC123572	1.21E-53
SI00233621	Arabidopsis thaliana chromosome 3 CHR3v07142002 genomic sequence	NM_113449	3.28E-32
SI00233622	Arabidopsis thaliana chromosome 4 CHR4v07142002 genomic sequence	NM_119749	4.75E-93
SI00233623	unknown [Arabidopsis thaliana]	AAM61357	1.08E-41
SI00233624	Unknown	No Hit	null
SI00233625	cytochrome P450 [Citrus sinensis]	AAL24049	2.22E-64
SI00233626	Arabidopsis thaliana Unknown protein (At5g61450) mRNA, complete cds	BT001154	2.00E-18
SI00233627	Putative kinesin-related protein [Oryza sativa (japonica cultivar-group)]	AAN62776	7.74E-18
SI00233628	Unknown	No Hit	null
SI00233629	Unknown	No Hit	null
SI00233630	Arabidopsis thaliana clone C105312 putative SOH1 protein (At5g19910) mRNA, complete cds	BT002744	4.57E-22
SI00233631	N.tabacum mRNA for precursor of photosysten II 22 kDa protein	X84225	8.33E-24
SI00233632	emb|CAB69839.1~gene_id:MXE10.4~similar to unknown protein [Arabidopsis thaliana] >gi|15240704|ref|NP_196881.1| putative protein; protein id: At5g13770.1 [Arabidopsis thaliana]	BAB10598	1.20E-38
SI00233633	Vernicia fordii omega-3 fatty acid desaturase precursor, mRNA, partial cds	AF061027	1.59E-77
SI00233634	Arabidopsis thaliana chromosome 5 CHR5v07142002 genomic sequence	NM_124911	1.51E-11
SI00233635	putative ribosomal protein L10 [Arabidopsis thaliana] >gi|17473553|gb|AAL38253.1| putative ribosomal protein L10 [Arabidopsis thaliana] >gi|28173519|sp|Q8VZB9|R10A_ARATH 60S ribosomal protein L10a-1	AAM47861	8.32E-93
SI00233636	plasma membrane H(+)-ATPase [Vicia faba, Otafuku, abaxial epidermis, guard cells protoplasts, mRNA, 3319 nt]	S79323	1.56E-111
SI00233637	Arabidopsis thaliana chromosome 1 CHR1v07142002 genomic sequence	NM_103100	2.32E-13
SI00233638	S.tuberosum dpeP mRNA for 4-alpha-glucanotransferase	X68664	1.25E-16
SI00233639	Oxygen-evolving enhancer protein 3 precursor-like protein [Arabidopsis thaliana] >gi|7267278|emb|CAB81060.1| Oxygen-evolving enhancer protein 3 precursor-like protein [Arabidopsis thaliana] >gi|18206249|sp|Q41932|PSQ2_ARATH Oxygen-evolving enhancer protein 3-2, chloroplast precursor (OEE3) (16 kDa	AAM65554	1.98E-44
SI00233640	Arabidopsis thaliana unknown protein (At5g59080) mRNA, complete cds	AY063936	1.34E-13
SI00233641	Arabidopsis thaliana putative protein (At3g44190) mRNA, complete cds	BT002026	1.73E-14
SI00233642	Unknown	No Hit	null
SI00233643	Unknown	No Hit	null
SI00233644	NADP specific isocitrate dehydrogenase [Daucus carota]	BAA34112	9.07E-127
SI00233645	putative DNA binding protein [Arabidopsis thaliana] >gi|13624639|emb|CAC36939.1| putative DNA binding protein [Arabidopsis thaliana] >gi|24211633|sp|Q9ZPY0|DAG2_ARATH DOF zinc finger protein DAG2 (Dof affecting germination 2)	CAC36940	3.91E-34
SI00233646	Unknown	No Hit	null

SI00233647	Arabidopsis thaliana chromosome 3 CHR3v07142002 genomic sequence	NM_111123	1.03E-20
SI00233648	Arabidopsis thaliana chromosome 4 CHR4v07142002 genomic sequence	NM_117349	2.96E-26
SI00233649	Camellia sinensis violaxanthin de-epoxidase (vde) mRNA, complete cds	AF462269	3.84E-35
SI00233650	Arabidopsis thaliana clone RAFL15-47-F06 (R20904) unknown protein (At2g32910) mRNA, complete cds	BT004303	5.12E-32
SI00233651	Unknown	No Hit	null
SI00233652	Unknown	No Hit	null
SI00233653	ribulose 1,5-bisphosphate carboxylase (EC 4.1.1.39) (aa 1-482) [Pisum sativum]	CAA27483	4.94E-124
SI00233654	A.majus gene for protoporphyrin IX: Mg Chelatase, subunit	X73144	1.36E-33
SI00233655	Arabidopsis thaliana chromosome 5 CHR5v07142002 genomic sequence	NM_121368	1.37E-62
SI00233656	Elaeis guineensis mRNA for metallothionein-like protein	AJ236913	3.26E-12
SI00233657	Arabidopsis thaliana unknown protein (At1g22200) mRNA, complete cds	AY113908	1.87E-15
SI00233658	Quinone oxidoreductase-like protein [Arabidopsis thaliana]	AAM62737	1.95E-62
SI00233659	Arabidopsis thaliana At4g15930 gene, complete cds	BT004785	5.04E-27
SI00233660	Arabidopsis thaliana chromosome 1 CHR1v07142002 genomic sequence	NM_101067	7.57E-35
SI00233661	B.juncea mRNA for chlorophyll a/b-binding protein	X95727	8.16E-58
SI00233662	Vigna radiata chlorophyll a/b binding protein CP29 (CipCp29) mRNA, complete cds; nuclear gene for chloroplast product	AF139466	9.01E-30
SI00233663	Petunia hybrida thiol protease homolog P21 (PeTh3) mRNA, complete cds	U31094	3.46E-26
SI00233664	putative translation initiation factor SUI1 [Oryza sativa (japonica cultivar-group)]	BAB89060	4.30E-42
SI00233665	Solanum tuberosum mRNA for putative internal rotenone-insensitive NADH dehydrogenase (nda1 gene)	AJ245861	9.00E-30
SI00233666	Nicotiana tabacum mRNA for chloroplast FtsH protease, complete cds	AB017480	2.64E-79
SI00233667	Haynaldia villosa clone kong32 mRNA	AF498269	1.92E-11
SI00233668	Unknown	No Hit	null
SI00233669	Zea mays PCO086463 mRNA sequence	AY104736	1.94E-18
SI00233670	N.tabacum mRNA for chloroplast Rieske FeS precursor protein 1	X66009	1.96E-61
SI00233671	unknown [Arabidopsis thaliana]	AAM61228	8.13E-69
SI00233672	A.hypochondriacus Lhcb2*Ah1 mRNA	X74732	2.28E-70
SI00233673	NTGP4 [Nicotiana tabacum]	AAD09518	3.83E-50
SI00233674	Medicago truncatula mRNA for 1-deoxy-D-xylulose 5-phosphate synthase 1 (dxs1 gene)	AJ430047	2.29E-67
SI00233675	Fagus crenata mRNA for chlorophyll a/b-binding protein, complete cds	AB006081	4.02E-75
SI00233676	Unknown	No Hit	null
SI00233677	Tabacco Cab16 mRNA for major chlorophyll a/b binding protein	X52741	1.12E-87
SI00233678	Tabacco Cab16 mRNA for major chlorophyll a/b binding protein	X52741	1.67E-80
SI00233679	putative L5 ribosomal protein [Arabidopsis thaliana] >gi|15810311|gb|AAL07043.1| putative L5 ribosomal protein [Arabidopsis thaliana] >gi|15234136|ref|NP_192040.1| putative L5 ribosomal protein; protein id: At4g01310.1, supported by cDNA: 22919., supported by cDNA: gi_15810310, supported by cDNA: g	AAM45056	2.02E-81
SI00233680	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	4.55E-27

SI00233681	Arabidopsis thaliana chromosome 3 CHR3v07142002 genomic sequence	NM_112631	7.73E-21
SI00233682	Pisum sativum chloroplast EF-G (fusA) mRNA, partial cds	L16508	9.23E-73
SI00233683	Arabidopsis thaliana At5g58590 mRNA sequence	AY050947	6.74E-12
SI00233684	Zea mays PCO105026 mRNA sequence	AY104143	1.72E-18
SI00233685	Fagus crenata mRNA for chlorophyll a/b-binding protein, complete cds	AB006081	1.65E-77
SI00233686	Arabidopsis thaliana chromosome 3 CHR3v07142002 genomic sequence	NM_111905	9.16E-33
SI00233687	Arabidopsis thaliana bHLH transcription factor, putative (At1g51070) mRNA, complete cds	BT002433	5.04E-68
SI00233688	Arabidopsis thaliana chromosome 1 CHR1v07142002 genomic sequence	NM_101575	1.05E-22
SI00233689	unknown [Arabidopsis thaliana] >gi|9758961|dbj|BAB09404.1| gene_id:MYH9.4~unknown protein [Arabidopsis thaliana] >gi|27808606|gb|AAO24583.1| At5g09830 [Arabidopsis thaliana] >gi|18416103|ref|NP_568217.1| expressed protein; protein id: At5g09830.1, supported by cDNA: 37422. [Arabidopsis thaliana]	AAM65194	1.95E-28
SI00233690	Unknown	No Hit	null
SI00233691	Arabidopsis thaliana unknown protein (At5g05200) mRNA, complete cds	AY063020	6.59E-43
SI00233692	Hypothetical protein [Arabidopsis thaliana] >gi|25405624|pir||A96572 hypothetical protein F8L10.1 [imported] - Arabidopsis thaliana	AAF87857	4.84E-73
SI00233693	Lettuce mRNA for light-harvesting chlorophyll a/b-binding protein (LHCP), complete cds	D14002	0
SI00233694	unknown protein [Arabidopsis thaliana] >gi|18401163|ref|NP_564489.1| expressed protein; protein id: At1g44000.1, supported by cDNA: gi_15028026, supported by cDNA: gi_20259312 [Arabidopsis thaliana] >gi|20259313|gb|AAM14392.1| unknown protein [Arabidopsis thaliana]	AAK76544	7.03E-59
SI00233695	Capsicum annuum mRNA for partial glyceraldehyde-3-phosphate dehydrogenase, subunit GapA2	AJ246009	1.13E-84
SI00233696	Lotus japonicus genomic DNA, chromosome 4, clone:LjT10J15, TM0007, complete sequence	AP004473	4.35E-13
SI00233697	Oryza sativa genomic DNA, chromosome 4, BAC clone: OSJNBa0041A02, complete sequence	AL606638	2.85E-11
SI00233698	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	1.14E-31
SI00233699	Unknown	No Hit	null
SI00233700	Arabidopsis thaliana chromosome 5 CHR5v07142002 genomic sequence	NM_124471	5.27E-59
SI00233701	Arabidopsis thaliana Unknown protein mRNA, complete cds	BT002510	0
SI00233702	osmotic stress-activated protein kinase [Nicotiana tabacum]	AAL89456	1.62E-19
SI00233703	P.sativum Cab-8 gene for photosystemm II chlorophyll a/b binding protein	X56538	0
SI00233704	unknown protein [Arabidopsis thaliana] >gi|21592833|gb|AAM64783.1| unknown [Arabidopsis thaliana] >gi|26453214|dbj|BAC43681.1| unknown protein [Arabidopsis thaliana] >gi|20260264|gb|AAM13030.1| unknown protein [Arabidopsis thaliana] >gi|12323736|gb|AAG51832.1|AC016163_21 unknown protein; 39989-3874	AAM91334	2.80E-77
SI00233705	gb|AAF67764.1~gene_id:MSH12.7~similar to unknown protein [Arabidopsis thaliana] >gi|22022566|gb|AAM83240.1| AT5g13610/MSH12_7 [Arabidopsis thaliana] >gi|15240669|ref|NP_196865.1| putative protein; protein id: At5g13610.1, supported by cDNA: 26016. [Arabidopsis thaliana] >gi|23308429|gb|AAN18184.1|	BAB08687	7.03E-62
SI00233706	Arabidopsis thaliana chromosome 5 CHR5v07142002 genomic sequence	NM_123907	3.63E-35
SI00233707	Arabidopsis thaliana chromosome 3 CHR3v07142002 genomic sequence	NM_112302	3.39E-100
SI00233708	putative translation factor [Pinus pinaster]	CAC84489	1.98E-53
SI00233709	Unknown	No Hit	null
SI00233710	manganese-binding protein PsbY precursor, photosystem II-associated - spinach >gi|6093830|sp|P80470|PSBY_SPIOL Photosystem II core complex proteins psbY, chloroplast precursor (L-arginine metabolising enzyme) (L-AME) [Contains: Photosystem II protein	T08902	3.40E-26

	psbY-1 (psbY-A1); Photosystem II protein psbY-2		
SI00233711	Unknown	No Hit	null
SI00233712	S.oleracea atpG mRNA for chloroplast CF(o)II ATP synthase subunit 9	X71397	8.20E-11
SI00233713	M.sativa mRNA for G protein beta subunit-like protein	Y08678	1.78E-46
SI00233714	putative giberellin beta-hydroxylase; protein id: At2g36690.1 [Arabidopsis thaliana] >gi|25285698|pir||E84783 probable giberellin beta-hydroxylase [imported] - Arabidopsis thaliana >gi|4415914|gb|AAD20145.1| putative giberellin beta-hydroxylase [Arabidopsis thaliana]	NP_181207	4.44E-25
SI00233715	F.anomala of gcsH gene encoding H-protein	Z37524	4.07E-40
SI00233716	P.silvestris mRNA for type II chlorophyll a/b binding protein of LHC I	X58516	5.41E-20
SI00233717	G.hirsutum cab gene for chlorophyll ab binding protein	X54090	1.74E-55
SI00233718	Spinach mRNA for O-acetylserine(thiol)-lyase, complete cds	D10476	2.14E-70
SI00233719	Putative F-box protein [Oryza sativa (japonica cultivar-group)]	AAN05494	1.40E-63
SI00233720	Arabidopsis thaliana chromosome 3 CHR3v07142002 genomic sequence	NM_115553	6.35E-77
SI00233721	Arabidopsis thaliana chromosome 4 CHR4v07142002 genomic sequence	NM_116335	2.99E-14
SI00233722	CONSTANS B-box zinc finger family protein; protein id: At5g24930.1, supported by cDNA: gi_15450658 [Arabidopsis thaliana]	NP_197875	1.42E-50
SI00233723	At2g20920/F5H14.11 [Arabidopsis thaliana] >gi|25412002|pir||H84594 hypothetical protein At2g20920 [imported] - Arabidopsis thaliana >gi|21700805|gb|AAM70526.1| At2g20920/F5H14.11 [Arabidopsis thaliana] >gi|4454458|gb|AAD20905.1| unknown protein [Arabidopsis thaliana] >gi|15226418|ref|NP_179688.1| u	AAL47447	9.42E-73
SI00233724	Arabidopsis thaliana chromosome 1 CHR1v07142002 genomic sequence	NM_101161	7.37E-86
SI00233725	L.esculentum mRNA for ammonium transporter	X95098	2.06E-21
SI00233726	Arabidopsis thaliana chromosome 1 CHR1v07142002 genomic sequence	NM_103431	7.87E-18
SI00233727	Expressed protein [Arabidopsis thaliana] >gi|20147419|gb|AAM10419.1| At2g46220/T3F17.13 [Arabidopsis thaliana] >gi|13605752|gb|AAK32869.1|AF361857_1 At2g46220/T3F17.13 [Arabidopsis thaliana] >gi|16226514|gb|AAL16188.1|AF428419_1 At2g46220/T3F17.13 [Arabidopsis thaliana] >gi|15215835|gb|AAK91462.1|	AAC62884	3.00E-69
SI00233728	Arabidopsis thaliana putative 23 kDa polypeptide of oxygen-evolving complex (OEC) (At1g06680) mRNA, complete cds	AY096487	1.26E-19
SI00233729	Vigna radiata chlorophyll a/b binding protein CP29 (CipCp29) mRNA, complete cds; nuclear gene for chloroplast product	AF139466	5.85E-34
SI00233730	photosystem I subunit PSI-E [Nicotiana sylvestris] >gi|7443147|pir||T16963 photosystem I chain PSI-E, isoform b - wood tobacco >gi|2499967|sp|Q41229|PSE2_NICSY Photosystem I reaction center subunit IV B, chloroplast precursor (PSI-E B)	AAB31705	4.35E-30
SI00233731	light-harvesting chlorophyll a/b-binding protein (LHCP) precursor [Lactuca sativa]	BAA03104	1.08E-31
SI00233732	Arabidopsis thaliana chromosome 4 CHR4v07142002 genomic sequence	NM_117349	2.93E-26
SI00233733	cytochrome P450 71A2 - eggplant >gi|441185|dbj|BAA03635.1| Cytochrome P-450EG4 [Solanum melongena] >gi|408140|emb|CAA50645.1| P450 hydroxylase [Solanum melongena] >gi|584861|sp|P37118|C712_SOLME Cytochrome P450 71A2 (CYPLXXIA2) (P-450EG4)	S36806	4.22E-31
SI00233734	Bruguiera gymnorhiza HPR mRNA for hydroxypyruvate reductase, complete cds	AB060810	1.51E-33
SI00233735	Nicotiana tabacum beta-carbonic anhydrase (CA) mRNA, complete cds; nuclear gene for chloroplast product	AF454759	7.50E-15
SI00233736	Pisum sativum chloroplast EF-G (fusA) mRNA, partial cds	L16508	1.00E-72
SI00233737	Capsicum annuum mRNA for partial glyceraldehyde-3-phosphate dehydrogenase, subunit GapB	AJ246010	5.01E-93
SI00233738	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 58	AL161558	2.76E-29
SI00233739	Arabidopsis thaliana chromosome 5 CHR5v07142002 genomic sequence	NM_123592	5.55E-25

SI00233740	Prunus persica omega-3 desaturase mRNA, partial cds; nuclear gene for chloroplast product	AF517831	1.07E-65
SI00233741	Crataegus crus-galli putative plastidic glutamine synthetase (GS2) mRNA, complete cds; nuclear gene for chloroplast product	AY162465	1.00E-75
SI00233742	Unknown	No Hit	null
SI00233743	Tomato CAB-8 gene (constucted from mRNA and DNA) for type III chlorophyll a/b binding polypeptide of photosystem I	X15258	1.05E-44
SI00233744	Gossypium hirsutum GhCAP mRNA for adenylyl cyclase associated protein, complete cds	AB014884	8.25E-21
SI00233745	Petunia x hybrida putative metallothionein-like protein (Mt2) mRNA, complete cds	AF201384	2.09E-11
SI00233746	Arabidopsis thaliana putative cdc20 protein (CDC20.2) mRNA, complete cds	AF029263	3.28E-26
SI00233747	outer membrane lipoprotein-like [Arabidopsis thaliana]	AAM62904	2.10E-75
SI00233748	Arabidopsis thaliana putative beta-ketoacyl-CoA synthase (At1g04220) mRNA, partial cds	AY074518	1.91E-18
SI00233749	Spinacea oleracea mRNA for 20 kDa protein of CP24	Z25886	8.17E-58
SI00233750	contains similarity to transcription regulator~gene_id:MRG7.19 [Arabidopsis thaliana]	BAB09481	1.04E-40
SI00233751	Unknown	No Hit	null
SI00233752	serine carboxypeptidase, putative; protein id: At1g33540.1 [Arabidopsis thaliana] >gi|12322376|gb|AAG51208.1|AC051630_5 serine carboxypeptidase, putative; 88458-86107 [Arabidopsis thaliana] >gi|25289776|pir||C86459 probable serine carboxypeptidase, 88458-86107 [imported] - Arabidopsis thaliana	NP_174619	4.45E-59
SI00233753	Papaver somniferum major latex protein (MLP146) gene, complete cds	L06467	6.15E-15
SI00233754	Flaveria pringlei mRNA for glycine hydroxymethyltransferase	Z25859	2.10E-64
SI00233755	Unknown	No Hit	null
SI00233756	coclaurine N-methyltransferase [Coptis japonica]	BAB71802	2.56E-30
SI00233757	Arabidopsis thaliana chromosome 4 CHR4v07142002 genomic sequence	NM_117266	1.20E-13
SI00233758	Mesembryanthemum crystallinum ribulose-1,5- bisphosphate carboxylase/oxygenase small subunit (rbcS-4) mRNA, complete cds	M38318	9.49E-48
SI00233759	Arabidopsis thaliana chromosome 3 CHR3v07142002 genomic sequence	NM_112201	9.61E-11
SI00233760	Arabidopsis thaliana thylakoid lumenal 17.4 kD protein, chloroplast precursor (P17.4) (At5g53490) mRNA, complete cds	BT003410	2.76E-17
SI00233761	Nicotiana tabacum mRNA for chloroplast FtsH protease, complete cds	AB017480	4.58E-13
SI00233762	Arabidopsis thaliana chromosome 3 CHR3v07142002 genomic sequence	NM_115553	2.12E-61
SI00233763	Zea mays PCO091925 mRNA sequence	AY106973	1.51E-12
SI00233764	Arabidopsis thaliana unknown protein (At2g25740) mRNA, complete cds	BT002106	9.02E-33
SI00233765	T17H3.9 [Arabidopsis thaliana]	AAD45997	1.47E-70
SI00233766	Arabidopsis thaliana ribosomal protein L17-like protein (At3g54210) mRNA, complete cds	BT000413	2.82E-29
SI00233767	Nicotiana tabacum beta-carbonic anhydrase (CA) mRNA, complete cds; nuclear gene for chloroplast product	AF454759	1.86E-15
SI00233768	Arabidopsis thaliana RAFL04-14-M02 mRNA sequence	AY091157	4.74E-32
SI00233769	Arabidopsis thaliana At4g35360/F23E12_80 mRNA, complete cds	AY143933	5.71E-31
SI00233770	Lettuce mRNA for light-harvesting chlorophyll a/b-binding protein (LHCP), complete cds	D14002	4.02E-75
SI00233771	Arabidopsis thaliana chromosome 5 CHR5v07142002 genomic sequence	NM_121608	7.93E-21
SI00233772	Mesembryanthemum crystallinum (clone: rbcs-1) rubisco small subunit gene, complete cds	L10212	1.15E-31
SI00233773	Medicago sativa glyoxysomal malate dehydrogenase precursor (gmdh) mRNA, complete cds	AF020270	1.34E-62

SI00233774	Mesembryanthemum crystallinum cinnamyl-alcohol dehydrogenase Eli3 mRNA, complete cds	U79770	7.97E-21
SI00233775	putative pyrophosphate-fructose-6-phosphate 1-phosphotransferase [Arabidopsis thaliana] >gi|22325825|ref|NP_179834.2| putative pyrophosphate--fructose-6-phosphate 1-phosphotransferase; protein id: At2g22480.1, supported by cDNA: gi_17979344, supported by cDNA: gi_20466012 [Arabidopsis thaliana] >gi	AAL49898	4.11E-79
SI00233776	Populus tremuloides cinnamyl alcohol dehydrogenase mRNA, complete cds	AF217957	2.09E-24
SI00233777	Oryza sativa chromosome 1 clone OSJNBa0048I01, complete sequence	AF229199	2.80E-17
SI00233778	Nicotiana sylvestris Lhcb1*8 gene for light harvesting chlorophyll a/b-binding protein, complete cds	AB012640	3.53E-106
SI00233779	copper amine oxidase-like protein, incomplete - Arabidopsis thaliana >gi|7267928|emb|CAB78270.1| copper amine oxidase like protein (fragment1) [Arabidopsis thaliana] >gi|5281038|emb|CAB45974.1| copper amine oxidase like protein (fragment1) [Arabidopsis thaliana] >gi|15234488|ref|NP_192964.1| copper	T48137	1.39E-35
SI00233780	Arabidopsis thaliana RAFL04-14-M02 mRNA sequence	AY091157	6.15E-47
SI00233781	Spinacia oleracea 24 kDa RNA binding protein mRNA, nuclear gene encoding chloplast protein, partial cds	U34742	1.84E-15
SI00233782	Arabidopsis thaliana chromosome 5 CHR5v07142002 genomic sequence	NM_124908	6.17E-111
SI00233783	unknown protein [Arabidopsis thaliana] >gi|22327037|ref|NP_197857.2| putative protein; protein id: At5g24690.1, supported by cDNA: gi_13877882, supported by cDNA: gi_15912312 [Arabidopsis thaliana] >gi|15912313|gb|AAL08290.1| AT5g24690/MXC17_8 [Arabidopsis thaliana] >gi|22136914|gb|AAM91801.1| unkn	AAK44019	7.40E-21
SI00233784	Fagus crenata mRNA for chlorophyll a/b-binding protein, complete cds	AB006081	3.13E-94
SI00233785	N.tabacum mRNA for chloroplast Rieske FeS precursor protein 1	X66009	1.98E-61
SI00233786	Arabidopsis thaliana chromosome 1 BAC T22E19 genomic sequence, complete sequence	AC016447	4.57E-41
SI00233787	Ricinus communis mRNA for glycine-rich RNA-binding protein (grp1 gene)	AJ245939	4.34E-19
SI00233788	Auxin-binding protein ABP19a precursor >gi|4098517|gb|AAD00295.1| auxin-binding protein ABP19 [Prunus persica]	Q9ZRA4	9.08E-73
SI00233789	Unknown	No Hit	null

Appendix 2

qRT-PCR results for the ten differential *top1* genes.

Appendix 2.1 -116D7
Appendix 2.2 -91G3
Appendix 2.3 -27D12
Appendix 2.4 -157C4
Appendix 2.5 -91B6
Appendix 2.6 -54F9
Appendix 2.7 -No results could be obtained for 52F2.
Appendix 2.8 -221C9
Appendix 2.9 -151H8
Appendix 2.10-111C7
Appendix 2.11-Salutaridine acetyltransferase (Control)

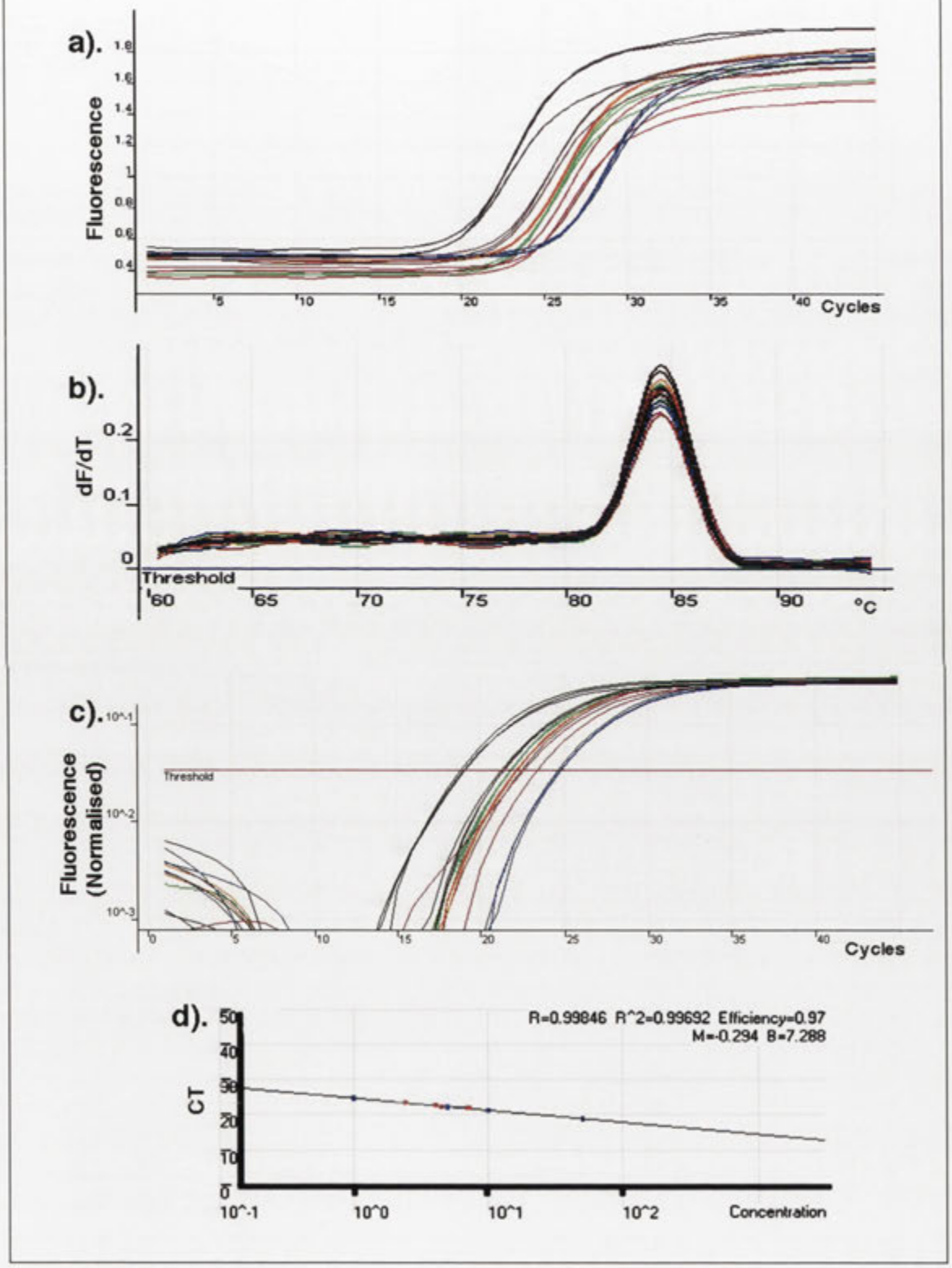
Brief description of the four charts:
a) Illustrates the fluorescence of the PCR product versus the number of PCR cycles. Each sample was repeated in triplicate. cDNA samples consisted of 5 different concentrations. A water control or zero concentration, cDNA (control or *top1*) at 1 ng/µL, cDNA (control or *top1*) at 5 ng/µL, cDNA (control or *top1*) at 10 ng/µL, and cDNA (control or *top1*) at 50 ng/µL.

b). This graph illustrates the melting temperature of the product in each reaction. This parameter gives confidence in the product being amplified from the cDNA and not primer dimers or unspecified primer binding.

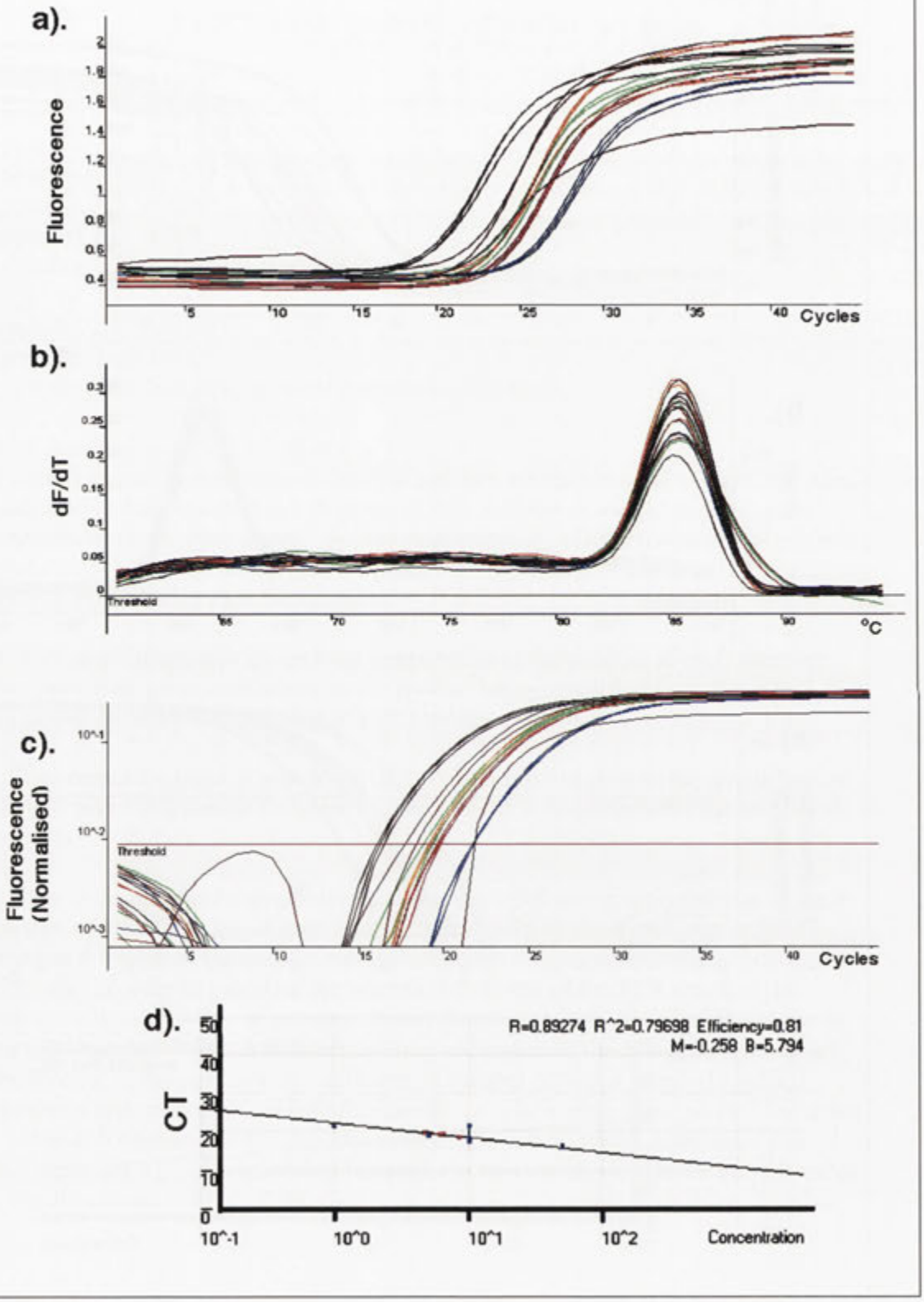
c). Represents the point at which the qRT-PCR threshold crosses the amplification curve to give the Ct value (in order to calculate the concentration of the amplified product).

d). This is the standard curve obtained using the 5 different concentrations of each reaction and the Ct value as calculated automatically by the Rotorgene software (Version 5.0 Corbett Research Corbett Research:1/14 Hilly St Mortlake NSW 2137 Australia) in order to calculate the reaction efficiency of the PCR reaction (ie theoretically the number of samples should double for each cycle and the closer to one that this value is, represents a doubling of product at the designated Ct value). The R-value is the correlation coefficient (a number between 1 and –1), which represents how well a line of best fit describes a relationship between two variables. A value of 0 implies that the data does not fit this relationship. A value of 1 or –1 does represent a good relationship between the two variables. Thus a good R-value is usually around 0.99 or –0.99.

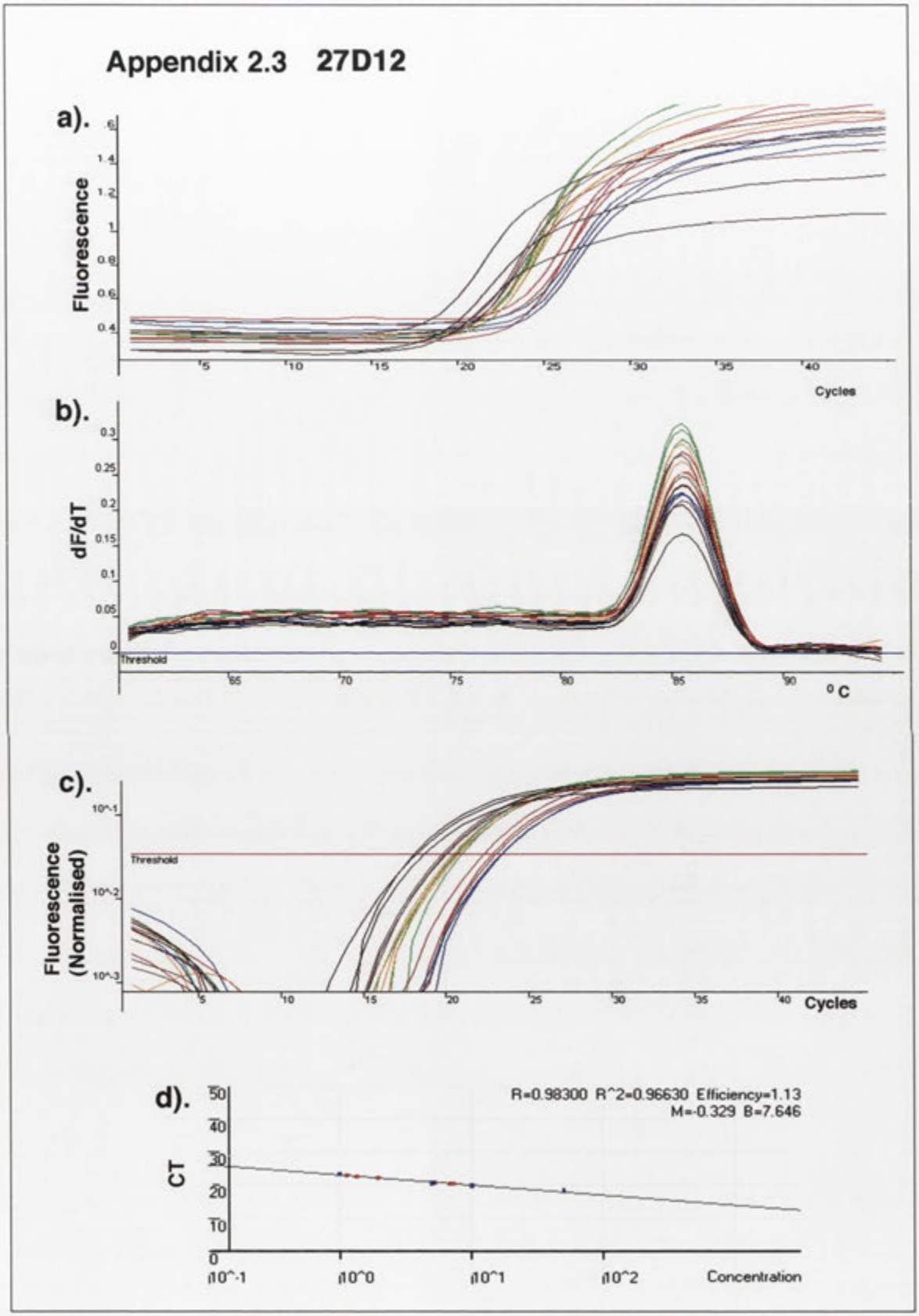
Appendix 2.1 116D7



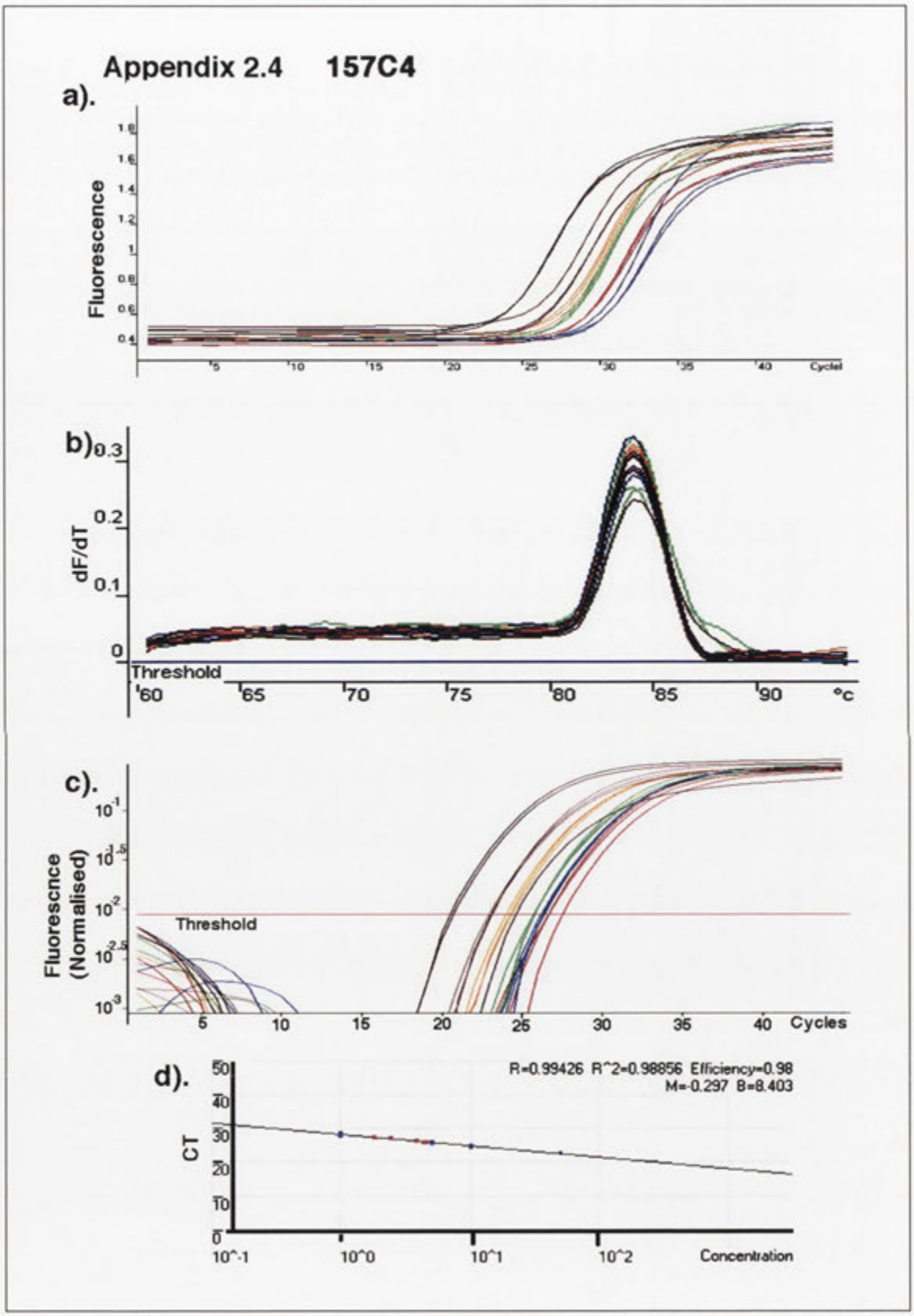
Appendix 2.2 91G3



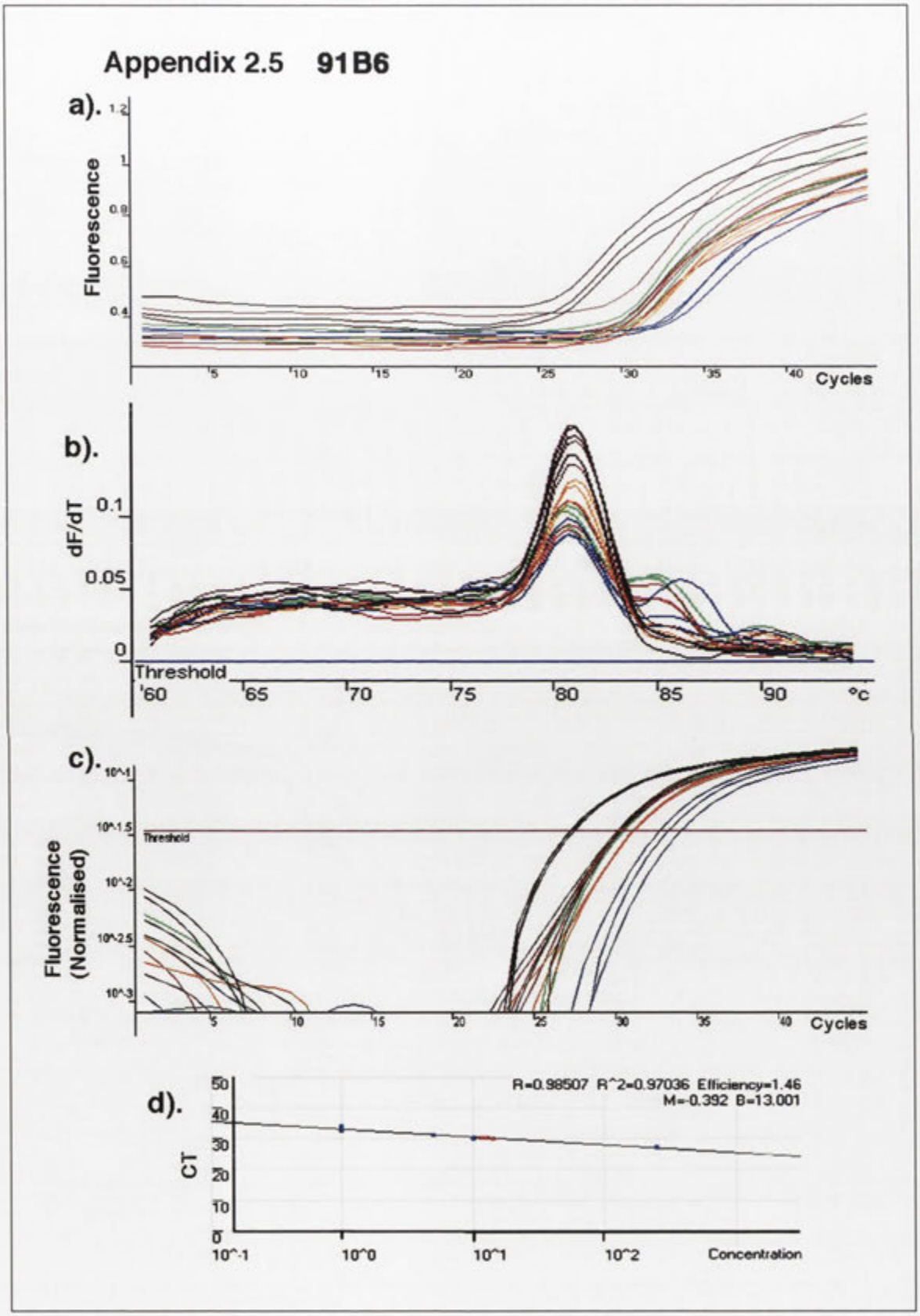
Appendix 2.3 27D12



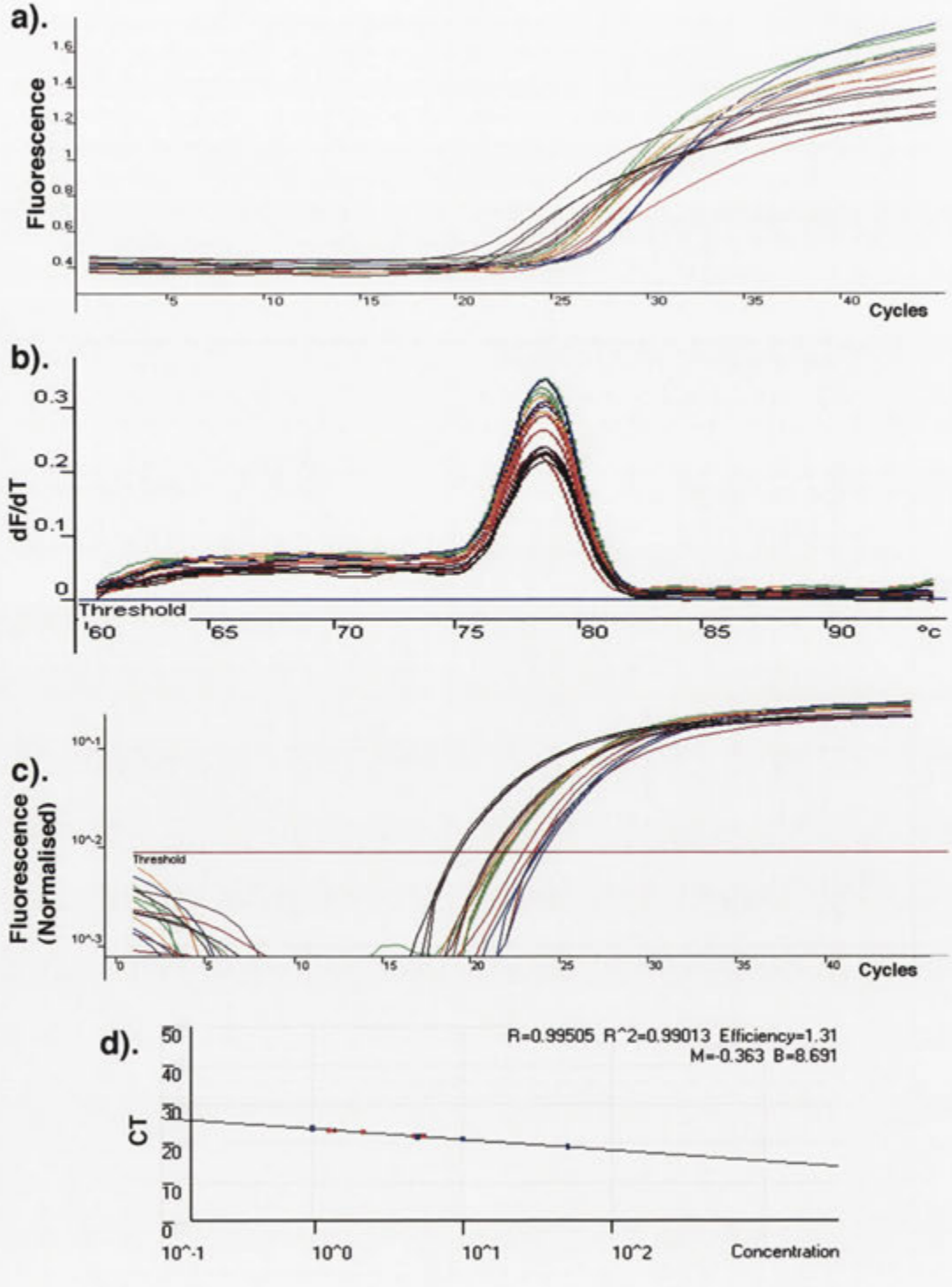
Appendix 2.4 157C4

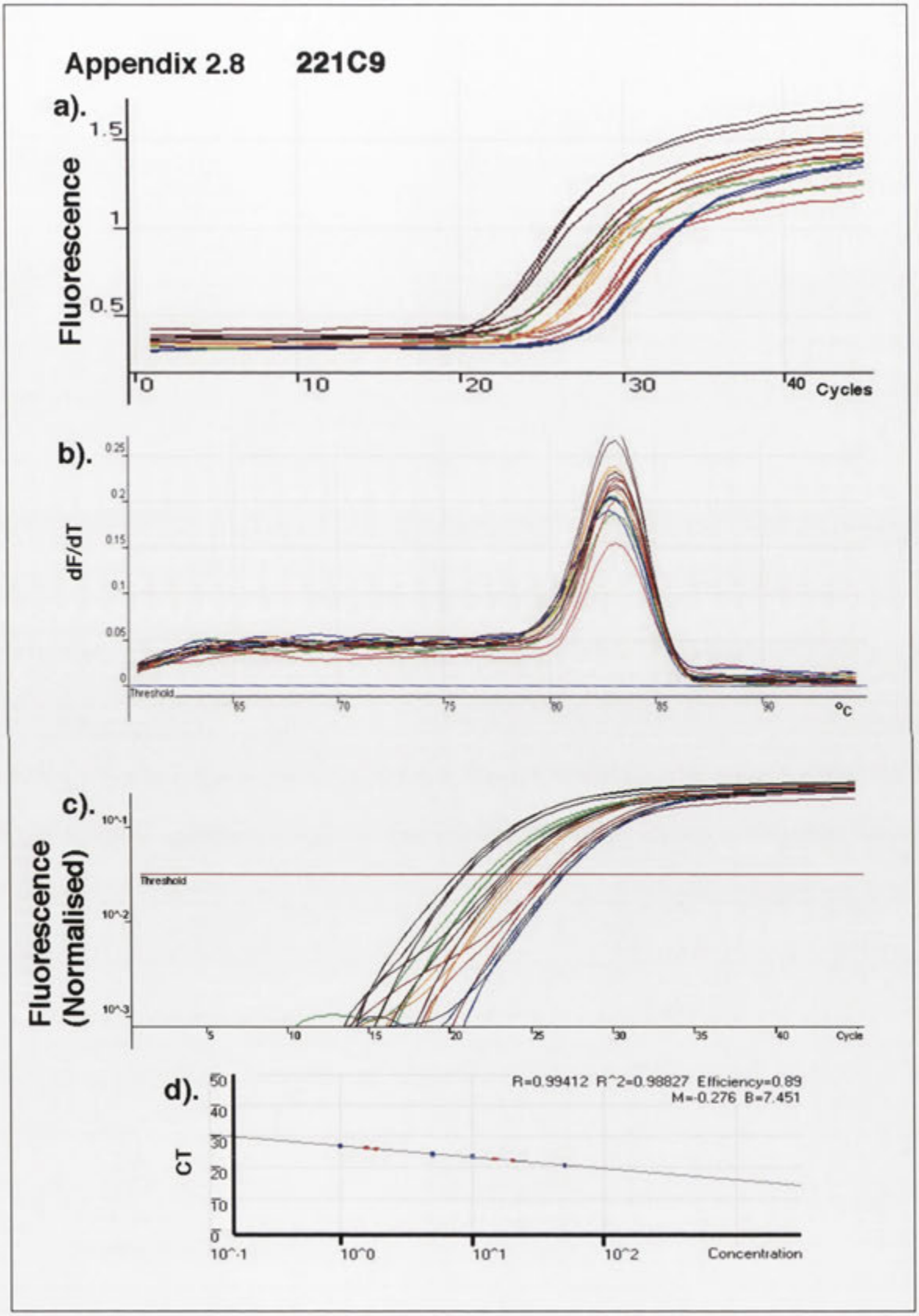


Appendix 2.5 91B6

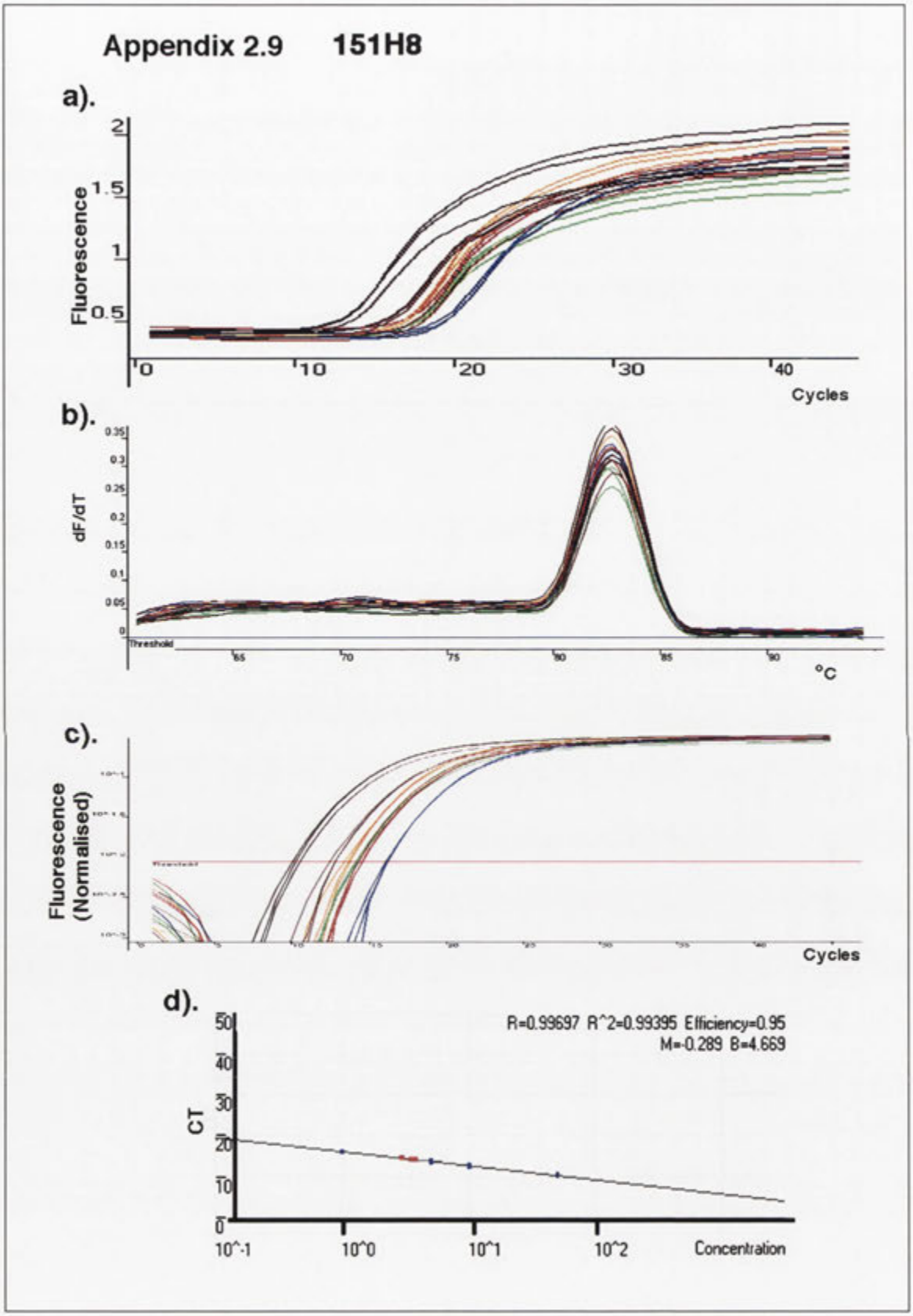


Appendix 2.6 54F9

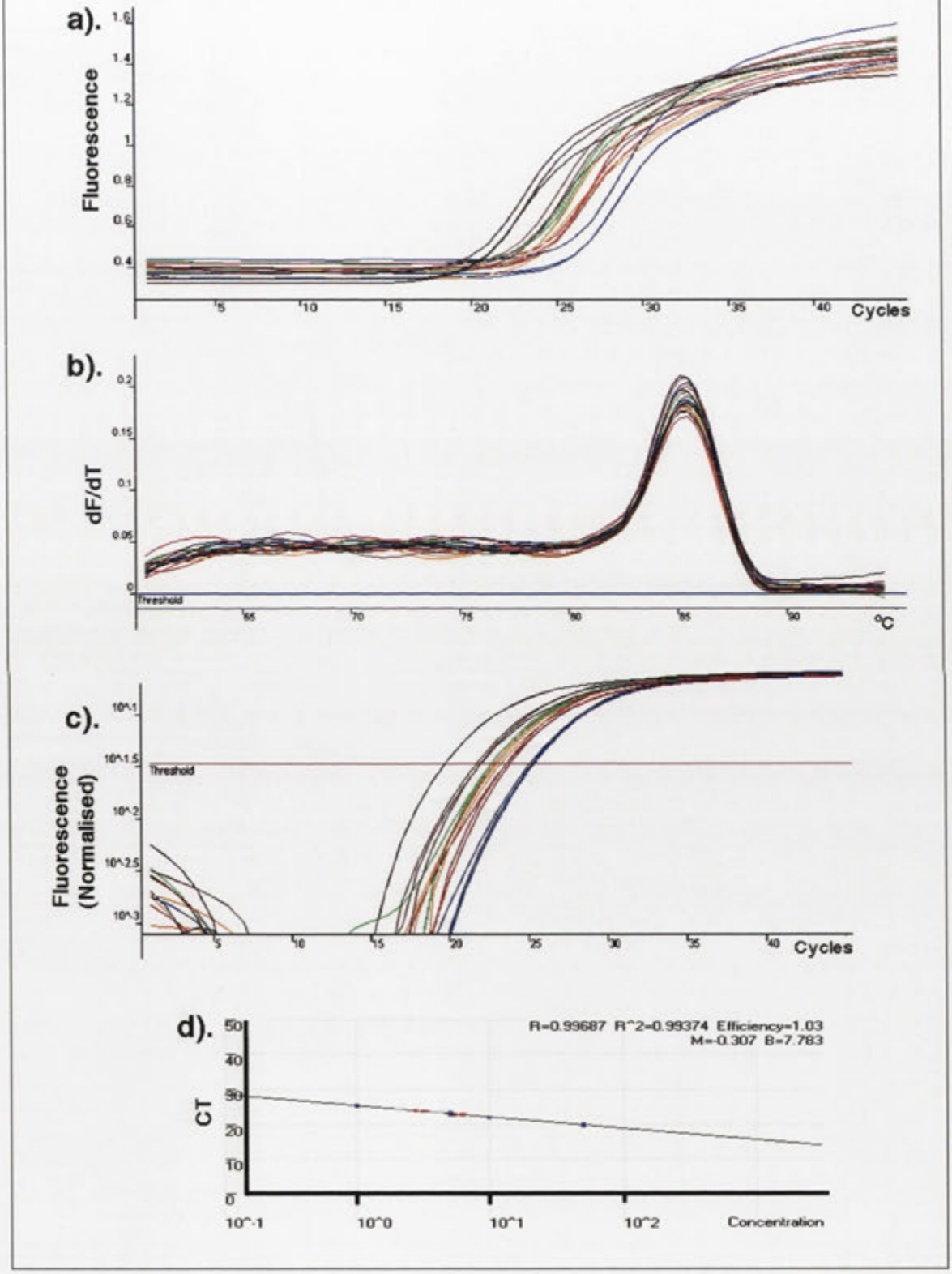



Appendix 2.8 221C9
a).
Fluorescence
1.5
1
0.5
0
10
20
30
40
Cycles
b).
dF/dT
0.25
0.2
0.15
0.1
0.05
0
Threshold
65
70
75
80
85
90
°C
c).
Fluorescence (Normalised)
10^-1
10^-2
10^-3
Threshold
5
10
15
20
25
30
35
40
Cycle
d).
R=0.99412 R^2=0.98827 Efficiency=0.89
M=-0.276 B=7.451
CT
50
40
30
20
10
0
10^-1
10^0
10^1
10^2
Concentration

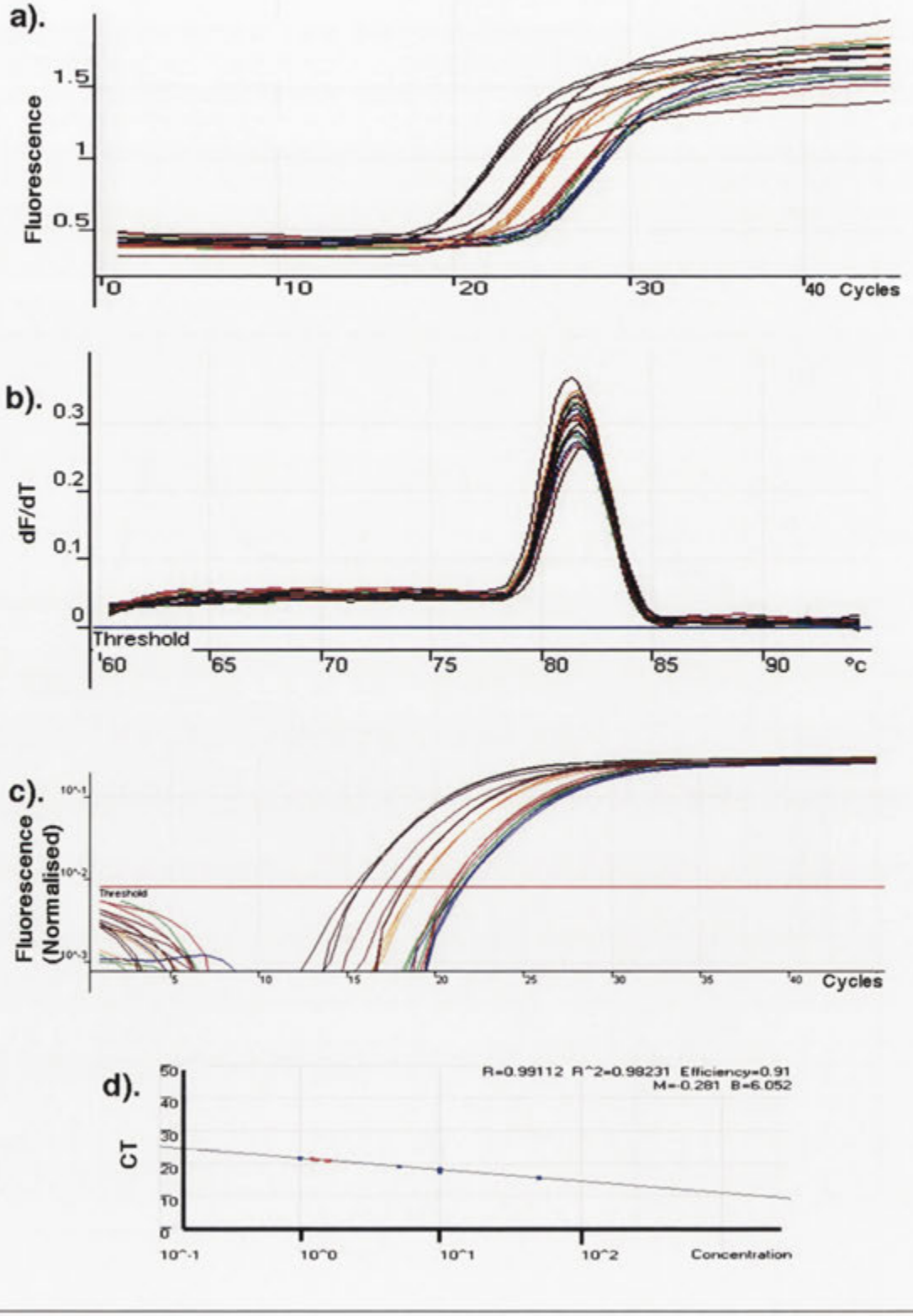
Appendix 2.9 151H8



Appendix 2.10 111C7



Appendix 2.11 Salutaridine Acetyl Transferase



Appendix 3

The differentially expressed genes from the low morphinan cultivar *P.somniferum* cv. Marianne compared to the high morphine cultivar *P. somniferum* C052.

Clone Name	Microarray ratio down regulated in the low morphine *P. somniferum* cv Marianne	Clone Name	Microarray ratio upregulated in the low morphine *P. somniferum* cv Marianne
154A5	2.78	194E9	1.38
86G4	2.75	221G8	1.47
139D3	2.42	116F7	1.56
143A4	2.47	157G11	1.73
54G7	2.15	25D10	1.68
54E4	2.74	25E11	1.39
21A7	2.70	83B1	1.74
144G8	2.61	162H7	1.48
83A3	2.59	152H6	1.74
115C7	2.39	13G6	1.57
146E5	2.38	189A9	1.48
158A8	2.21	178D6	1.74
57F5	1.95	158B10	1.47
210C5	2.19	105A10	1.55
149G4	1.97	48H4	1.48
93G12	2.50	105H7	1.48
82D6	2.47	143C6	1.72
12F3	2.15	63G12	1.68
121A4	2.48	133H8	1.65
150B1	2.09	21A1	1.73
5D2	2.32	104B7	1.41
21E9	2.18	35B6	1.52
29H2	1.83	180E10	1.49
98E11	2.26	76E10	1.39
174D11	1.82	209G10	1.75
108B2	2.12	209C9	1.51
84A9	2.39	157D3	1.40
23A10	2.58	16G12	1.45
194G12	2.01	138E12	1.48
33H3	1.98	110G2	1.49
22E7	1.95	209B12	1.40
183H10	1.99	199D8	1.49
19B8	1.94	110D8	1.49

179E7	1.90	140B7	1.49
175E5	1.89	10E8	1.77
20A7	2.53	53E5	1.50
145D12	1.72	196G8	1.40
112B10	1.90	85A11	1.46
151C1	1.89	14B12	1.59
99G8	1.89	218C9	1.41
133H10	1.87	202D4	1.57
105A3	1.89	60G4	1.58
111H8	1.86	90F1	1.60
24E1	1.83	30C1	1.51
20C4	1.84	48G10	1.58
59C6	2.30	108G3	1.46
183C8	1.98	29F10	1.43
193C9	1.85	196G3	1.50
180D9	2.17	133F11	1.51
36C12	1.82	206H6	1.51
25D3	2.00	132F2	1.51
84H6	1.83	162F5	1.72
110D7	2.33	67A3	1.52
193B8	1.82	80E3	1.52
191D9	2.31	174F6	1.42
202E5	2.30	29H1	1.42
14G8	1.79	179E10	1.54
83C2	2.22	175A7	1.43
77C5	1.80	162C10	1.67
206E2	2.28	54C12	1.43
79D8	1.77	23G7	1.43
146H5	1.77	87C9	1.74
115C5	1.78	14A4	1.82
223D1	2.24	81H3	1.48
27G5	1.75	15H10	1.54
133A5	1.76	209G8	1.57
103E12	1.76	152E4	1.54
76F5	1.74	185B5	1.70
221C9	1.73	117D8	1.44
146G5	1.73	133A2	1.49
139A8	1.89	46H2	1.55
174C1	2.18	144B11	1.55
213E3	1.63	201E7	1.45
60G8	1.72	147E5	1.47
218H9	1.58	30A3	1.49
114G12	1.87	38A7	1.82
20E7	2.00	78B11	1.66
99B10	2.14	83H7	1.54

37A8	1.86	158D3	1.56
40E12	1.68	95B7	1.50
34B1	2.12	98F2	1.53
46D3	1.56	136G9	1.56
151F2	2.03	136D2	1.56
202A7	2.11	81B5	1.69
38C5	1.68	27D12	1.56
88B4	1.68	1A6	1.64
145G4	1.68	52A5	1.45
17D9	1.68	12E5	1.56
154G6	1.68	139F8	1.48
58B10	1.83	117A4	1.59
192B7	1.82	139C9	1.59
150C1	2.09	110D10	1.59
221G3	2.09	218E7	1.67
78A3	1.67	178F9	1.58
79A9	1.88	158H9	1.47
34D10	1.81	179D4	1.71
83C10	2.06	63F1	1.51
79B10	2.08	172F12	1.59
22H11	1.98	144E12	1.59
110A4	1.62	174F7	1.63
30F1	1.85	51E1	1.74
111C7	1.67	10H5	1.60
71D3	2.07	112F1	1.48
184C1	1.70	43B3	1.68
206F8	2.06	194G3	1.50
112D11	2.03	105H5	1.60
169G12	1.79	63C9	1.61
52D6	1.70	143H12	1.77
133C8	2.04	142H3	1.78
98F3	1.66	46F9	1.50
16C4	2.04	2F8	1.66
5G1	1.52	23B10	1.50
78B8	1.78	156G2	1.50
23F8	1.64	175D7	1.50
160A8	1.65	46G12	1.72
4C4	2.01	218H5	1.65
143A3	2.01	195E6	1.63
155C7	1.65	158A11	1.55
2H3	2.01	206C2	1.66
83B9	1.63	46D6	1.65
7H12	1.59	98C12	1.52
52A12	2.00	113C5	1.66
4F5	1.99	175D1	1.66

112C1	1.63	32A4	1.66
32F3	1.50	91F5	1.66
91F10	1.75	117H4	1.60
27E3	1.97	82G7	1.65
23A3	1.55	19H11	1.65
24A1	1.73	29F11	1.67
60H7	1.83	20A10	1.58
104H9	1.49	202H12	1.88
174G8	1.60	210A11	1.55
180C4	1.58	89F6	1.54
58D9	1.96	132F8	1.68
88D6	1.69	132F4	1.68
160H10	1.56	39F8	1.68
133G3	1.74	152D10	1.67
147C1	1.59	74A2	1.59
58H10	1.61	59A7	1.63
29E4	1.74	161B1	1.57
54F1	1.94	117G4	1.85
33A9	1.71	83F7	1.56
7G9	1.49	39A7	1.56
21E12	1.63	132G12	1.70
51E6	1.86	78G7	1.71
57B12	1.93	168A6	1.71
200E8	1.63	76F9	1.78
38A4	1.59	159B12	1.70
88E7	1.87	151F8	1.72
47C4	1.68	108B11	1.77
82A1	1.69	22D7	1.72
22G4	1.47	40E11	1.57
116F10	1.58	46C8	1.73
111A8	1.57	197E2	1.73
55H2	1.89	52B4	2.17
194E6	1.57	121H5	1.74
223E3	1.46	30G3	1.59
19A11	1.56	89F7	1.60
206E4	1.46	146A5	1.67
133H10	1.56	52A4	1.77
109C9	1.62	135G5	1.70
14A6	1.88	1D6	1.64
137B9	1.64	16B3	1.66
58H4	1.55	136C12	1.70
100E11	1.56	67E6	1.63
108C2	1.56	224B5	1.79
209F1	1.61	147H1	1.72
210B12	1.74	193E12	1.84

78C4	1.57	112H3	1.89
135C1	1.48	39D8	1.77
27G7	1.66	94D10	1.65
28E8	1.51	94D8	1.78
97A10	1.55	181C1	1.69
185F10	1.74	91F2	1.80
140G4	1.61	48B9	1.92
60D5	1.44	133F11	1.87
52F2	1.54	183F11	1.91
221F3	1.48	20E6	1.77
27E4	1.44	97D11	1.78
11B3	1.58	103E6	1.85
140F1	1.44	30F3	2.14
146B9	1.85	21B4	2.04
79A11	1.55	200B9	1.94
96H2	1.70	147C5	1.85
82A7	1.44	224F3	1.95
5D5	1.53	99B9	2.11
108E2	1.75	18A10	2.00
180D2	1.64	71B1	1.97
186G4	1.83	90B12	2.27
34D11	1.61	180H4	3.54
29A3	1.52	180E1	4.95
224A10	1.82		
28G4	1.82		
31A8	1.56		
194H3	1.82		
117D11	1.46		
155A3	1.63		
151H3	1.51		
89A5	1.51		
85C12	1.48		
82A9	1.52		
84B9	1.50		
84F8	1.55		
179E6	1.80		
21C4	1.65		
146B1	1.79		
121H3	1.51		
63D11	1.51		
4A9	1.79		
190D5	1.59		
112D12	1.74		
37B4	1.41		
88E8	1.59		

209D4	1.79		
57H11	1.78		
78F6	1.75		
183B10	1.51		
35H2	1.63		
137C3	1.51		
26G7	1.50		
34F7	1.61		
107C12	1.49		
84B12	1.52		
26C2	1.40		
11F5	1.77		
210F1	1.50		
159C8	1.49		
17A9	1.68		
139E6	1.76		
108A11	1.49		
162F3	1.75		
178D5	1.57		
210H11	1.40		
140F4	1.49		
162B8	1.76		
168H6	1.64		
109B11	1.42		
27B2	1.66		
221B4	1.75		
133D5	1.48		
160C1	1.75		
206C9	1.61		
58H7	1.49		
25A11	1.39		
137G1	1.46		
157C4	1.57		
102H3	1.47		
202E7	1.59		
213A4	1.41		
139C1	1.47		
78D10	1.39		
186C3	1.40		
221F9	1.56		
57B1	1.39		
121F8	1.68		
189D1	1.49		
96D3	1.47		
213G9	1.47		

4E2	1.73		
27D12	1.73		
26H10	1.47		
2F6	1.43		
135F8	1.65		
210D9	1.50		
63C1	1.42		
151E6	1.72		
144H8	1.44		
104F10	1.46		
83H6	1.72		

Appendix 4

The differentially expressed genes from the low morphine accumulating *P. somniferum* mutant D205.

Clone Name	Microarray ratio down regulated in the mutant D120	Clone Name	Microarray ratio up regulated in the mutant D120
27D12	7.52	94E4	2.01
28G4	7.25	81H8	2.03
91B6	7.20	76H3	2.03
116D7	6.48	86A12	2.03
111C7	6.36	225C5	2.15
91G3	5.86	199A8	2.20
71G5	5.69	29H11	2.24
156F9	5.64	137A3	2.29
157C4	5.50		
108H11	5.45		
33H5	5.42		
52F2	4.82		
83G7	4.70		
C E12	4.42		
153C2	4.39		
112B7	4.32		
151H8	3.94		
20A7	3.62		
221C9	3.57		
91E9	3.40		
83F5	3.34		
145E11	3.15		
46A6	3.03		
148A9	2.98		
194D4	2.96		
54F9	2.80		
33F11	2.78		
147A12	2.59		
139A8	2.52		
209F3	2.49		
102D3	2.43		
186F8	2.43		
16E8	2.43		
137B7	2.39		
137A5	2.36		

196B12	2.34
40F6	2.34
144F10	2.33
18D8	2.31
90B12	2.25
32D4	2.24
209D12	2.23
39G1	2.17
198D10	2.16
121D2	2.15
158G12	2.15
25D3	2.15
100C6	2.15
29F4	2.14
83D11	2.14
51D10	2.13
14C6	2.13
37F4	2.12
17A7	2.11
79B8	2.09
52C5	2.09
27C12	2.09
46F7	2.09
58B3	2.08
105D7	2.08
43D7	2.07
57E4	2.07
76C11	2.07
97F4	2.06
48F11	2.05
196E12	2.04
30B6	2.04
22D6	2.04
79B7	2.04
7E1	2.04
98B6	2.03
19D4	2.03
60B2	2.03
82B7	2.03
102F3	2.02
55E11	2.02
25F7	2.02
159C8	2.01
114D3	2.01

20F6	2.00
156G7	2.00
149H2	2.00

Appendix 5

The differentially expressed genes from the high total alkaloid accumulating *P. somniferum* mutant D242.

Clone Name	Microarray ratio down regulated in the mutant D242	Clone Name	Microarray ratio up regulated in the mutant D242
27D12	9.98	18D11	2.02
91G3	8.68	38F9	2.03
111C7	7.66	77D6	2.06
151H8	7.64	15E9	2.06
52F2	6.84	26D12	2.07
91B6	6.24	57F3	2.09
108H11	6.10	18G11	2.11
71G5	5.74	193D12	2.12
156F9	5.65	77B10	2.12
157C4	5.38	108D11	2.19
221C9	5.32	98B10	2.29
91E9	5.19	86A12	2.66
28G4	5.12	209B8	2.72
54F9	4.92	25B4	2.77
186F8	4.60		
147A12	4.57		
116D7	4.34		
153C2	4.28		
33H5	4.21		
11B6	4.15		
83F5	4.02		
83G7	3.91		
112B7	3.90		
144F10	3.27		
148A9	2.94		
83E5	2.87		
162G6	2.69		
189B10	2.65		
137A7	2.57		
15C7	2.56		
20A7	2.55		
81B12	2.53		
145E11	2.53		
198D10	2.49		
191D7	2.49		
117B11	2.48		

117D11	2.48
116G4	2.45
202D1	2.44
51D8	2.43
110D4	2.42
169E12	2.38
160H9	2.38
43A12	2.38
75C2	2.37
26B11	2.34
23A10	2.32
194B5	2.32
33F11	2.31
83A3	2.30
196B12	2.28
98C6	2.27
12F3	2.26
30D5	2.25
54E4	2.25
29E10	2.22
39D5	2.22
34A11	2.19
196H1	2.19
160B3	2.17
111F12	2.16
21A7	2.15
132G1	2.15
146D2	2.14
26A1	2.14
158B3	2.12
161D7	2.12
139A8	2.11
111C8	2.11
3D10	2.11
36A7	2.10
87B5	2.10
30B11	2.09
117G3	2.09
37F9	2.07
22E7	2.06
121A4	2.06
18H8	2.05
169G12	2.04
138D10	2.04

19C9	2.04
10D7	2.02
194D4	2.02
63E8	2.01
106F8	2.01
93H2	2.00

Appendix 6

The differentially expressed genes from the low morphinan and high papaverine accumulating *P. somniferum* mutant D120.

Clone Name	Microarray ratio down regulated in the mutant D120	Clone Name	Microarray ratio upregulated in the mutant D120
91G3	14.06	25H4	2.02
27D12	11.55	152F3	2.02
28G4	11.27	31C8	2.02
91B6	10.28	24F2	2.02
80F11	9.78	78F7	2.02
111C7	8.79	29E3	2.03
54F9	8.64	97H11	2.03
221C9	7.57	26D4	2.04
151H8	7.31	132D2	2.04
52F2	7.22	181A3	2.04
91E9	6.94	115G4	2.04
83F5	6.92	2H2	2.05
157C4	6.76	160B6	2.06
83G7	6.60	110B8	2.06
116D7	6.27	159D4	2.06
83E5	4.63	186G12	2.06
153C2	4.51	157E12	2.07
112B7	4.51	71G2	2.07
198D10	4.46	213E7	2.09
108H11	4.01	159D10	2.09
63D4	3.93	25D4	2.09
C E12	3.65	77H11	2.09
18H8	3.43	38F5	2.10
3D10	3.25	88E4	2.10
63G12	2.92	15H5	2.10
196E10	2.83	140D2	2.11
104G9	2.73	199D3	2.11
150H6	2.70	201B4	2.12
30F11	2.65	46H4	2.13
132F8	2.65	156D7	2.13
132F4	2.59	29B4	2.14
116E12	2.57	38G4	2.14
89F7	2.55	16G7	2.15
34G11	2.49	78C8	2.15
206C2	2.47	162E8	2.16
95B7	2.46	88E5	2.16

48D12	2.45	51F8	2.17
158H9	2.44	161D11	2.17
103E6	2.44	38E5	2.18
155F3	2.44	221H9	2.18
178C2	2.41	200D3	2.18
22D7	2.41	27D1	2.23
151F8	2.39	94B4	2.23
147A12	2.38	159F3	2.23
89F6	2.36	74F7	2.26
76F9	2.33	83F8	2.27
175D1	2.32	46F1	2.28
218H11	2.31	11H8	2.29
21A1	2.27	26D12	2.33
172F12	2.27	115D12	2.33
117G4	2.26	10B5	2.35
16B3	2.25	57B7	2.38
190A7	2.24	140B1	2.38
157B9	2.24	116A12	2.38
116G9	2.23	106F7	2.40
20E6	2.23	58A11	2.44
1D6	2.23	102B12	2.44
94D10	2.22	15B2	2.47
74A2	2.22	59B8	2.47
206H6	2.22	157A11	2.49
17G5	2.22	175F5	2.51
19H11	2.22	29D2	2.55
117D11	2.21	148E5	2.57
135G5	2.21	105B7	2.58
133F11	2.20	87F9	2.60
179D4	2.20	31B5	2.60
99B9	2.19	37H2	2.60
30B9	2.19	143E3	2.63
113D7	2.18	221E4	2.64
97D11	2.17	121E10	2.68
46H2	2.15	146C4	2.70
39F8	2.14	26B10	2.76
133F11	2.14	174F10	2.80
10H5	2.14	38E8	2.81
102D3	2.14	137A3	2.89
158F6	2.13	15E9	2.90
48H4	2.13	192E10	2.96
151G10	2.13	132A11	2.99
183F11	2.13	105H7	3.00
160B3	2.13	193D12	3.00

112H3	2.12	22E1	3.01
83F7	2.12	58A4	3.03
63C9	2.12	209H2	3.14
147H1	2.12	18D11	3.16
137G3	2.11	86A12	3.21
90B12	2.10	7A9	3.37
91F1	2.10	76H3	3.43
39D8	2.09	77F10	3.52
53G10	2.09	71E4	3.54
37F9	2.09	39H6	3.65
162H7	2.09	25B4	3.94
111C8	2.08	152F12	4.04
174A1	2.08	18G11	4.07
51E1	2.06	209B8	4.18
153F8	2.06	29H11	4.26
12G11	2.05	38F9	4.53
53E5	2.03		
80E3	2.03		
29D1	2.01		
46C8	2.01		
105H5	2.01		
79B5	2.00		
15C7	2.00		
30C1	2.00		

Appendix 7

Alignment of the 5 cDNAs that align to the 26S ribosomal RNA DNA.

```
                      1                                                  50
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11        (1)
« 46A6          (1) TTTTTTTTTTTTTTTTTATTTTTTACGAATCGGTGCGACATGGGGCTGAAT
 Consensus      (1) TTTTTTTTTTTTTTTTTATTTTTTACGAATCGGTGCGACATGGGGCTGAAT
                (2)
                     51                                                 100
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11        (1)
« 46A6         (51) CTCAGTGGATCGTGGCAGCAAGGCCACTCTGCCACTTACAATACCCCGTC
 Consensus     (51) CTCAGTGGATCGTGGCAGCAAGGCCACTCTGCCACTTACAATACCCCGTC
               (52)
                    101                                                 150
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11        (1)
« 46A6        (101) GCGTATTTAAGTCGTCTGCAAAGGATTCTACCCGCCACCCGGTGGTAATA
 Consensus    (101) GCGTATTTAAGTCGTCTGCAAAGGATTCTACCCGCCACCCGGTGGTAATA
              (102)
                    151                                                 200
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11        (1)
« 46A6        (151) GTAYTNTWAGGGCGGCCCGCGCAGCTCGTCCGCTGCGAGGGCTACGCCCA
 Consensus    (151) GTANTNTNAGGGCGGCCCGCGCAGCTCGTCCGCTGCGAGGGCTACGCCCA
              (152)
                    201                                                 250
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11        (1)
« 46A6        (201) CGGCACGTGCCTCTGGGGACCCGAAGGTCCCTACTGCAGGTCGGCAATCG
 Consensus    (201) CGGCACGTGCCTCTGGGGACCCGAAGGTCCCTACTGCAGGTCGGCAATCG
              (202)
                    251                                                 300
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11        (1)
« 46A6        (251) GGCGACGGGCGCATGTGCCGCTTCTTTAGCCTGGATTCTGACTTAGAGGC
 Consensus    (251) GGCGACGGGCGCATGTGCCGCTTCTTTAGCCTGGATTCTGACTTAGAGGC
              (252)
                    301                                                 350
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11        (1)
« 46A6        (301) GTTCAGTCATAATCCTGCACACGGTAGCTTCGCGCCACTGGCTTTTCAAC
 Consensus    (301) GTTCAGTCATAATCCTGCACACGGTAGCTTCGCGCCACTGGCTTTTCAAC
              (302)
                    351                                                 400
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11        (1)
« 46A6        (351) CAAGCGCGATGACCAATTGTGTGAATCAACGGTTCCTCTCGTACTAGGTT
 Consensus    (351) CAAGCGCGATGACCAATTGTGTGAATCAACGGTTCCTCTCGTACTAGGTT
              (352)
```

```
                    401                                                 450
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11        (1)
« 46A6        (401) GAATTACTATCGCGGCACTGTCATCAGTAGGGTAAAACTAACCTGTCTCA
 Consensus    (401) GAATTACTATCGCGGCACTGTCATCAGTAGGGTAAAACTAACCTGTCTCA
              (402)
                    451                                                 500
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11        (1)                GGCAG-TGGGGTTGGGGGGCGCAGC-ATATGCAGCT
« 46A6        (451) CGACGGTCTAATCC--CAGCTCACGTT------C-C--CTAT-TG--G-T
 Consensus    (451) CGACGGTCTAATCCGGCAGCTSRSGTTGGGGGGCGCAGCTATATGCAGCT
              (452)
                    501                                                 550
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11       (37) AGTTTACACATTATACACTTAAAAATTT-T-ATATTTACCTTAGAGCTTT
« 46A6        (501) GGGTGA-ACA--AT-C-C--AACACTTGGTGA-ATT--C-T--G--CTTC
 Consensus    (501) RGKTKACACATTATACACTTAAMAMTTKGTGATATTTACCTTAGAGCTTY
              (502)
                    551                                                 600
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11       (87) A-AATCTCTG-TAGGTAGTTTGTCC-A-ATT-ATGTCA-CACCACAGAAG
« 46A6        (551) ACAAT----GATAGGAAGA--G-CCGACATCGAAGG-ATCA--A-A-AAG
 Consensus    (551) ACAATCTCTGATAGGWAGWTTGTCCGACATYGAWGKCATCACCACAGAAG
              (552)
                    601                                                 650
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11      (137) TAAGGTTC-CT-TCA-CAAA-GA-T-CC-CAAGCTAGCAGTTTTCCCAGT
« 46A6        (601) CAACGT-CGCTATGAACGCTTGGCTGCCACAAGC---CAGTTATCCCTGT
 Consensus    (601) YAASGTTCGCTATSAACRMWTG  TGCCACAAGCTAGCAGTTWTCCCWGT
              (602)
                    651                                                 700
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11      (187) CACG-A-CGTTGTAAAAC-GACGGCCAGTGCCTAGCTT-ATAATAC-GA-
« 46A6        (651) G--GTAACTTT-T----CTGACA-CC--T-C-TAGCTTCA-AATTCCGAA
 Consensus    (651) SACGTAACKTTGTAAAACTGACRGCCAGTGCCTAGCTTCATAATWCCGAA
              (652)
                    701                                                 750
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11      (237) -CTCACTATAGG---G--A--CCACT-TTGTACAAGAAAGCTGGGTACGC
« 46A6        (701) GGTC--TAAAGGATCGATAGGCCACGCTT-T-CACG---GTTCG-TATTC
 Consensus    (701) GSTCACTAWAGGATCGATAGGCCACKCTTGTACAMGAAAGYTSGGTAYKC
              (702)
                    751                                                 800
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11      (287) GTAAGCTTGGGCCCCTCGAGGGATACTCTAGAGCG-GCCG---CCCTTTT
« 46A6        (751) GTA--CT-GGAAA--TC-AG--A-A-TCAA-A-CGAGCTTTTACCCTTTT
 Consensus    (751) GTAAGCTTGGRMMCCTCGAGGGATACTCWAGAGCGAGCYKTTACCCTTTT
              (752)
                    801                                                 850
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11      (337) GTTCCACACGAGATTTCTGTTCTCGTTGAGCTCATCTTAGGACACCTGCG
« 46A6        (801) GTTCCACACGAGATTTCTGTTCTCGTTGAGCTCATCTTAGGACACCTGCG
 Consensus    (801) GTTCCACACGAGATTTCTGTTCTCGTTGAGCTCATCTTAGGACACCTGCG
```

```
                (802)

                  851                                                 900
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11      (387) TTATCTTTTAACAGATGTGCCGCCCCAGCCAAACTCCCCACCTGACAATG
« 46A6        (851) TTATCTTTTAACAGATGTGCCGCCCCAGCCAAACTCCCCACCTGACAATG
 Consensus    (851) TTATCTTTTAACAGATGTGCCGCCCCAGCCAAACTCCCCACCTGACAATG
              (852)
                  901                                                 950
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11      (437) TCTTCCGCCCGGATCGGCATAGCAAGCTAGCCTTGGATCTAAAAAGAGGG
« 46A6        (901) TCTTCCGCCCGGATCGGCATAGCAAGCTAGCCTTGGATCTAAAAAGAGGG
 Consensus    (901) TCTTCCGCCCGGATCGGCATAGCAAGCTAGCCTTGGATCTAAAAAGAGGG
              (902)
                  951                                                1000
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11      (487) GCGATGCCCCGCTTCCGCCTCACGGAATAAGTAAAATAACGTTAAAAGTA
« 46A6        (951) GCGATGCCCCGCTYCCGCCTCMCGGAATAAGTAAAATAACGTTAAAAGTA
 Consensus    (951) GCGATGCCCCGCTYCCGCCTCMCGGAATAAGTAAAATAACGTTAAAAGTA
              (952)
                 1001                                                1050
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11      (537) GTGGTATTTCACCTTCGCCGCCAAAGACGG-CTCCCACTTATCCTACACC
« 46A6       (1001) GTGGTATTTCACCTTCGCCGCCAAAGAAGGGCTCCCACTTATCCTACACC
 Consensus   (1001) GTGGTATTTCACCTTCGCCGCCAAAGAMGGGCTCCCACTTATCCTACACC
             (1002)
                 1051                                                1100
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11      (587) TCTCAAGTCATTTCACAAAGTCGGACTAGAGTCAAGCTCAACAGGGTCTT
« 46A6       (1051) TCTCAAGTCATTTCACAAAGTCGGACTAGAGTCAAGCTCAACAGGGTCTT
 Consensus   (1051) TCTCAAGTCATTTCACAAAGTCGGACTAGAGTCAAGCTCAACAGGGTCTT
             (1052)
                 1101                                                1150
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11      (637) CTTTCCCCGCTGATTCTGCCAAGCCCGTTCCCTTGGCTGTGGTTTCGCTG
« 46A6       (1101) CTTTCCCCGCTGATTCTGCCAAGCCCGTTCCCTTGGCTGTGGTTTCCCTG
 Consensus   (1101) CTTTCCCCGCTGATTCTGCCAAGCCCGTTCCCTTGGCTGTGGTTTCSCTG
             (1102)
                 1151                                                1200
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11      (687) GATAGTAGACAGGGACAGTGGGAATCTCGATAATCCATTCATGCGCGTCA
« 46A6       (1151) GATAGTAGACAGGGACAGDGGGAATCTCGTTAATCCATTCATGCGWGTCA
 Consensus   (1151) GATAGTAGACAGGGACAGDGGGAATCTCGWTAATCCATTCATGCGHGTCA
             (1152)
                 1201                                                1250
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11      (737) CTAATTAGATGACGAGG-CATTTG-GCTACCTTAAGAGAGTCATAGTTAC
« 46A6       (1201) CTASTTAGATGACGAGGACATTTGCGCTACYTAAAGAGAGTCATAGTTAC
 Consensus   (1201) CTAVTTAGATGACGAGGACATTTGCGCTACYTWAAGAGAGTCATAGTTAC
             (1202)
```

```
                    1251                                                1300
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11      (787) TCCC-GCC-GTTTACCCGCGCTTGGTTGAATTTCTTCACTTTGACATTCA
« 46A6       (1251) TCCCCGCCCGTTTACCCGCGCTTGGTTGAATTTCTTCACTTTGACATTCA
 Consensus   (1251) TCCCCGCCCGTTTACCCGCGCTTGGTTGAATTTCTTCACTTTGACATTCA
             (1252)
                    1301                                                1350
« 33H5          (1)
« 156F9         (1)
« 71G5          (1)
« 145E11      (837) GAGCACTGGGCAGAAATCACATTGCGTTAGCATCCGCAGGGACCATCGCA
« 46A6       (1301) GAGCACTGGGCAGAAATCACATTGCGTTAGCATCCGCAGGGACCATCGCA
 Consensus   (1301) GAGCACTGGGCAGAAATCACATTGCGTTAGCATCCGCAGGGACCATCGCA
             (1302)
                    1351                                                1400
« 33H5          (1)   TTTTTTTTTTTTTTAATTAAACAGTCG-GATTCCCC-TTGTCCGTAC-
« 156F9         (1)  TTTTTTTTTTTTTTTAATTAAACAGTCG-GATTCCCC-TTGTCCGTAC-
« 71G5          (1)  TTTTTTTTTTTTTTTAATTAAACAGTCG-GATTCCCCNTTGTCCGTAC-
« 145E11      (887) ATGCTTTGTTTT----AATTAAACAGTCG-GATTCCCC-TTGTCCGTAC-
« 46A6       (1351) ATGCTTTGTTTT----AASTAAACAGTGGAGATTCCCC-TTGTCCGTACA
 Consensus   (1351) ATTTTTTTTTTTTTTTAATTAAACAGTCGNGATTCCCCNTTGTCCGTACN
             (1352)
                    1401                                                1450
« 33H5         (49) CAGTTCTGAGTCGACTGTTCGACGCACGCGGAAGGCCCCGTGAGGAGCCG
« 156F9        (50) CAGTTCTGAGTCGACTGTTCGACGCACGCGGAAGGCCCCGTGAGGAGCCG
« 71G5         (50) CAGTTCTGAGTCGACTGTTCGACGCACGCGGAAGGCCCCGTGAGGAGCCG
« 145E11      (937) CAGTTCTGAGTCGACTGTTCGACGCACGCGGAAGGCCCCGTGAGGAGCCG
« 46A6       (1401) CAGTTCTGAGTCGACTGTTCGACGCACGCGGAAGGCCCCGTGAGGAGCCG
 Consensus   (1401) CAGTTCTGAGTCGACTGTTCGACGCACGCGGAAGGCCCCGTGAGGAGCCG
             (1402)
                    1451                                                1500
« 33H5         (99) TTCCCAATCCGTCCCCCGGCCGGCACGCGACGACCCGCTCTCGCCGCGTG
« 156F9       (100) TTCCCAATCCGTCCCCCGGCCGGCACGCGACGACCCGCTCTCGCCGCGTG
« 71G5        (100) TTCCCAATCCGTCCCCCGGCCGGCACGCGACGACCCGCTCTCGCCGCGTG
« 145E11      (987) TTCCCAATCCGTCCCCCGGCCGGCACGCGACGACCCGCTCTCGCCGCGTG
« 46A6       (1451) TTCCCAATCCGTCCCCCGGCGGGCACGWGACGACCCGCTCTCGCCGCGTG
 Consensus   (1451) TTCCCAATCCGTCCCCCGGCCGGCACGCGACGACCCGCTCTCGCCGCGTG
             (1452)
                    1501                                                1550
« 33H5        (149) AGCAGCTCGAGCAGT-TC-GCCGACAGCCGACGGGTTCGGGACTGGGACC
« 156F9       (150) AGCAGCTCGAGCAGT-TC-GCCGACAGCCGACGGGTTCGGGACTGGGACC
« 71G5        (150) AGCAGCTCGAGCAGT-TC-GCCGACAGCCGACGGGTTCGGGACTGGGACC
« 145E11     (1037) AGCAGCTCGAGCAGT-TC-GCCGACAGCCGACGGGTTCGGGACTGGGACC
« 46A6       (1501) AGCAGCTCGAGCA-WCTCC  AGACAGCCGACGGGTTCGGGACTGGGACC
 Consensus   (1501) AGCAGCTCGAGCAGTNTCNGCCGACAGCCGACGGGTTCGGGACTGGGACC
             (1502)
                    1551                                                1600
« 33H5        (199) TCGTGCCCAGCCCTCAGAGCCAATCCTTTTCCCGAGGTTACGGATCCATT
« 156F9       (200) TCGTGCCCAGCCCTCAGAGCCAATCCTTTTCCCGAGGTTACGGATCCATT
« 71G5        (200) TCGTGCCCAGCCCTCAGAGCCAATCCTTTTCCCGAGGTTACGGATCCATT
« 145E11     (1087) TCGTGCCCAGCCCTCAGAGCCAATCCTTTTCCCGAGGTTACGGATCCATT
« 46A6       (1551) TCGTGCCCAGCCCTCAGAGCCAATCCTTTTCCCGAGGTTACGGATCCATT
 Consensus   (1551) TCGTGCCCAGCCCTCAGAGCCAATCCTTTTCCCGAGGTTACGGATCCATT
             (1552)
                    1601                                                1650
« 33H5        (249) TTG-CCGACTTCCCTTGCCTACATTGTTCCA-TCGACCAGAGGCTGTTCA
« 156F9       (250) TTGGCCGACTTCCCTTGCCTACATTGTTCCAATCG  CAGAGGCTGTTCM
« 71G5        (250) TTG-CCGACTTCCCTTGCCTACATTGTTCCA-TCGACCAGAGGCTGTTCA
« 145E11     (1137) TTG-CCGACTTCCCTTGCCTACATTGTTCCA-TCGACCAGAGGCTGTTCA
« 46A6       (1601) TTG-CCGACTTCCCTTGCCTACATTGTTCCA-TCGACCAGAGGCTGTTCC
 Consensus   (1601) TTGNCCGACTTCCCTTGCCTACATTGTTCCANTCGACCAGAGGCTGTTCA
             (1602)
                    1651                                                1700
« 33H5        (299) CCTTGGA-GACCTGATGCGGTTATGAGTACGACCGGGCGTGGATGGCACT
« 156F9       (300) CCYTGGAAGACCTGATGCGGTTATGAGTACGACCGGGWGTGGATGGCACT
« 71G5        (300) CCTTGGA-GACCTGATGCGGTTATGAGTACGACCGGGCGTGGATGGCACT
« 145E11     (1187) CCTTGGA-GACCTGATGCGGTTATGAGTACGACCGGGCGTGGATGGCACT
« 46A6       (1651) CCTTGGA-GACCTGATGCGGTTATGAGTACGMCCGGGCGTGGAGGGCACT
 Consensus   (1651) CCTTGGANGACCTGATGCGGTTATGAGTACGACCGGGCGTGGATGGCACT
```

```
                (1652)

                        1701                                                1750
« 33H5           (349) CGGTCCTCCGGATTTTCAAGGGCC-GCCGGGGGCGCACCGGACACCACGC
« 156F9          (350) CGGTCCTCCGGATTTTCAAGGGCCCGCCGGGGGCGCACCGGACACCACGW
« 71G5           (350) CGGTCCTCCGGATTTTCAAGGACC-GCGGGGGGCGCACCGGACACCACGG
« 145E11        (1237) CGGTCCTCCGGATTTTCAAGGGCC-GCCGGGGGCGCACCGGACACCACGC
« 46A6          (1701) CGGTCCTCCWWMTTTTCAARGGCC-CCCGGGGGCGCCCCGGACACCMCGC
 Consensus      (1701) CGGTCCTCCGGATTTTCAAGGGCCNGCCGGGGGCGCACCGGACACCACGC
                (1702)
                        1751                                                1800
« 33H5           (399) GACGTGCGGT-GCTCTTCCGGCCGCTGGACCCTACCTCCGGCTGAACCGT
« 156F9          (400) GACGTGCGGT-GYTYTTCCGGCCGWTGGACCYTACCTCCGGCTGAACCGT
« 71G5           (400) GACGTGCGGT-GCTCTTCCGGCCGCTGGACCCTACCTCCGGCTGAACCGT
« 145E11        (1287) GACGTGCGGT-GCTCTTCCGGCCGCTGGACCCTACCTCCGGCTGAACCGT
« 46A6          (1751) GACGCGMGGTAGCTCTTCCGGCCGCTGGACCCTACCTCCGGCTGAMCCGT
 Consensus      (1751) GACGTGCGGTNGCTCTTCCGGCCGCTGGACCCTACCTCCGGCTGAACCGT
                (1752)
                        1801                                                1850
« 33H5           (449) TTCCAGGGTGGGCAGGCCGTTAAACAGAAAAGATAACTCTTTC-CGAGGC
« 156F9          (450) TTCCAGGGTGGGCAGGCCGTTAAMCAGAAAAGATAACTCTTTC-CGAGGC
« 71G5           (450) TTCCAGGGTGGGCAGGCCGTTAAACAGAAAAGATAACTCTTTC-CGAGGC
« 145E11        (1337) TTCCAGGGTGGGCAGGCCGTTAAACAGAAAAGATAACTCTTTC-CGAGGC
« 46A6          (1801) TTCCAGGGTGGCCAG  CG-----C-G----GA---CGCGTGGGCGA--C
 Consensus      (1801) TTCCAGGGTGGGCAGGCCGTTAAACAGAAAAGATAACTCTTTCNCGAGGC
                (1802)
                        1851                                                1900
« 33H5           (499) CC-C-CG-CCGACGTCTCCGGACT-CCCTAACGTTGCCGTCA-GCCGCCA
« 156F9          (500) CC-C-CG-CCGACGTCTCCGGACTTCCCTAACGTTGCCGTCAAGCCGCCA
« 71G5           (500) CC-C-CG-CCGACGTCTCCGGACT-CCCTAACGTTGCCGTCA-GCCGCCA
« 145E11        (1387) CC-C-CG-CCGACGTCTCCGGACT-CCCTAACGTTGCCGTCA-GCCGCCA
« 46A6          (1851) CCACGCGTCCG
 Consensus      (1851) CCNCNCGNCCGACGTCTCCGGACTNCCCTAACGTTGCCGTCANGCCGCCA
                (1852)
                        1901                                                1950
« 33H5           (549) CGT-CCCGGTTCA-GGAATTTTAA-CCCGATT-CCCTTTCGAAGTTCGCG
« 156F9          (550) CGT-CCCGGTTCAAGGAATTTTAAACCCGATTTCCCTTTCGAAGTTCGCG
« 71G5           (550) CGT-CCCGGTTCA-GGAATTTTAA-CCCGATT-CCCTTTCGAAGTTCGCG
« 145E11        (1437) CGTNCCCGGTTCA-GGAATTTTAA-CCCGATT-CCCTTTCGAAGTTCGCG
« 46A6         (•1861)
 Consensus      (1901) CGTNCCCGGTTCANGGAATTTTAANCCCGATTNCCCTTTCGAAGTTCGCG
                (1902)
                        1951                                                2000
« 33H5           (599) TGTTGCACGCTCTCTGACGGGCTTCCCCC-GTCTCTTAGGATCGACTAAC
« 156F9          (600) TGTTGCACGCTCTCTGACGGGCTTCCCCCCGTCTCTTAGGATCGACTAAC
« 71G5           (600) TGTTGCACGCTCTCTGACGGGCTTCCCCC-GTCTCTTAGGATCGACTAAC
« 145E11        (1487) TGTTGCACGCTCTCTGACGGGCTTCCCCC-GTCTCTTAGGATCGACTAAC
« 46A6         (•1861)
 Consensus      (1951) TGTTGCACGCTCTCTGACGGGCTTCCCCCNGTCTCTTAGGATCGACTAAC
                (1952)
                        2001                                                2050
« 33H5           (649) CCATGTGCAAGTGCCGTTCACATGGAA-CCTTTCCCCTCTTCGGCCTTCA
« 156F9          (650) CCATGTGCAAGTGCCGTTCACATGGAAACCTTTCCCYTCTTCGGCCTTCA
« 71G5           (650) CCATGTGCAAGTGCCGTTCACATGGAA-CCTTTCCCCTCTTCGGCCTTCA
« 145E11        (1537) CCATGTGCAAGTGCCGTTCACATGGAA-CCTTTCCCCTCTTCGGCCTTCA
« 46A6         (•1861)
 Consensus      (2001) CCATGTGCAAGTGCCGTTCACATGGAANCCTTTCCCCTCTTCGGCCTTCA
                (2002)
                        2051                                                2100
« 33H5           (699) AAGTTCTCATTTGAATATTTGCTACTACCACCAAGATCTGCACCGACGGA
« 156F9          (700) AAGTTCTCATTTGAATATTTGCTACTACCACCAAGATCTGCACCGACGGA
« 71G5           (700) AAGTTCTCATTTGAATATTTGCTACTACCACCAAGATCTGCACCGACGGA
« 145E11        (1587) AAGTTCTCATTTGAATATTTGCTACTACCACCAAGATCTGCACCGACGGA
« 46A6         (•1861)
 Consensus      (2051) AAGTTCTCATTTGAATATTTGCTACTACCACCAAGATCTGCACCGACGGA
                (2052)
```

```
                    2101                                                2150
« 33H5       (749) GGCTCCGCCCGGCCTCACGGCCTAGGTTTCGCAGCCACCGCCGCGCCCTC
« 156F9      (750) GGCTCCGCCCGGCCTCACGGCCTAG-TTTCGCAGCCACCGCCGCGCCCTC
« 71G5       (750) GGCTCCGCCCGGCCTCACGGCCTAGGTTTCGCAGCCACCGCCGCGCCCTC
« 145E11    (1637) GGCTCCGCCCGGCCTCACGGCCTAGGTTTCGCAGCCACCGCCGCGCCCTC
« 46A6     (•1861)
 Consensus  (2101) GGCTCCGCCCGGCCTCACGGCCTAGGTTTCGCAGCCACCGCCGCGCCCTC
            (2102)
                    2151                                                2200
« 33H5       (799) CTACTCATCGGGG-CCTAACGATTGCCCCGACGGCCGGGTGTGGGTCGCG
« 156F9      (800) CTACTCATCGGGGGCCTAACGATTGCCCCRACGGCCGGGTGTGGGTCGCG
« 71G5       (800) CTACTCATCGGGG-CCTAACGATTGCCCCGACGGCCGGGTGTGGGTCGCG
« 145E11    (1687) CTACTCATCGGGG-CCTAACGATTGCCCCGACGGCCGGGTGTGGGTCGCG
« 46A6     (•1861)
 Consensus  (2151) CTACTCATCGGGGNCCTAACGATTGCCCCGACGGCCGGGTGTGGGTCGCG
            (2152)
                    2201                                                2250
« 33H5       (849) CGCTTCAGCGCCATCCATTTT-CGGGGCTAGTTGATTCGGCAGGTGAGTT
« 156F9      (850) CGCTTCAGCGCCATCCATTTTTCGGGGCTAGTTGATTCGGCAGGKGAGTT
« 71G5       (850) CGCTTCAGCGCCATCCATTTT-CGGGGCTAGTTGATTCGGCAGGTGAGTT
« 145E11    (1737) GGCTTCAGCGCCATCCATTTT-CGGGGCTAGTTGATTCGGCAGGTGAGTT
« 46A6     (•1861)
 Consensus  (2201) CGCTTCAGCGCCATCCATTTTNCGGGGCTAGTTGATTCGGCAGGTGAGTT
            (2202)
                    2251                                                2300
« 33H5       (899) GTTACACACTCCTTAGCGGATTTCGACTTCCATGACCACCG-TCCTGCTG
« 156F9      (900) GTTACMCACTCCTTA  GGATTTCGACTTCMATGAMCACCGGTCCTGWTG
« 71G5       (900) GTTACACACTCCTTAGCGGATTTCGACTTCCATGACCACCG-TCCTGCTG
« 145E11    (1787) GTTACACACTCCTTAGCGGATTTCGACTTCCATGACCACCG-TCCTGCTG
« 46A6     (•1861)
 Consensus  (2251) GTTACACACTCCTTAGCGGATTTCGACTTCCATGACCACCGNTCCTGCTG
            (2252)
                    2301                                                2350
« 33H5       (949) TCTTAATCGACCAACACCCTTTGTGGGTTCTAGGTTAGCGCGCAGTTGGG
« 156F9      (950) TCTTAATCGACMAACACCCTTTGTGGGTTCTARGTTA  GCGCAGTTGGG
« 71G5       (950) TCTTAATCGACCAACACCCTTTGTGGGTTCTAGGTTAGCGCGCAGTTGGG
« 145E11    (1837) TCTTAATCGACCAACACCCTTTGTGGGTTCTAGGTTAGCGCGCAGTTGGG
« 46A6     (•1861)
 Consensus  (2301) TCTTAATCGACCAACACCCTTTGTGGGTTCTAGGTTAGCGCGCAGTTGGG
            (2302)
                    2351                                                2400
« 33H5       (999) CACCGTAACCCGGCTTCCGGTTCATCCCGCAT-CGCCAGTTCTGCTTACC
« 156F9     (1000) CACCGTAACCCGGCTTCCGGTTCATCCCGCATTCGCCAGTTCTGCTTACM
« 71G5      (1000) CACCGTAACCCGGCTTCCGGTTCATCCCGCAT-CGCCAGTTCTGCTTACC
« 145E11    (1887) CACCGTAACCCGGCTTCCGGTTCATCCCGCAT-CGCCAGTTCTGCTTACC
« 46A6     (•1861)
 Consensus  (2351) CACCGTAACCCGGCTTCCGGTTCATCCCGCATNCGCCAGTTCTGCTTACC
            (2352)
                    2401                                                2450
« 33H       (1049) AAAAATGGCCCACTTGGAGCTCTCGATTCCGTGGCGCGGCT-CATTGGAG
« 156F9     (1050) AAAAATGGCCCACTTGGAGCTCTCGASTCCGKGGCGCGGCTTYSTTGGAG
« 71G5      (1050) AAAAATGGCCCACTTGGAGCTCTCGATTCCGTGGCGCGGCT-CATTGGAG
« 145E11    (1937) AAAAATGGCCCACTTGGAGCTCTCGATTCCGTGGCGCGGCT-CATTGGAG
« 46A6     (•1861)
 Consensus  (2401) AAAAATGGCCCACTTGGAGCTCTCGATTCCGTGGCGCGGCTNCATTGGAG
            (2402)
                    2451                                                2500
« 33H5      (1099) CAGCCGCGCCGTCCTACCTATTTAAAGTTTGAGAATAGGTCGAGGGCGTT
« 156F9     (1100) CAGCCGCGCCGTCCTACCTATTTAAAGTTTGAGAATAGGTCGAGGGCGTT
« 71G5      (1100) CAGCCGCGCCGTCCTACCTATTTAAAGTTTGAGAATAGGTCGAGGGCGTT
« 145E11    (1987) CAGCCGCGCCGTCCTACCTATTTAAAGTTTGAGAATAGGTCGAGGGCGTT
« 46A6     (•1861)
 Consensus  (2451) CAGCCGCGCCGTCCTACCTATTTAAAGTTTGAGAATAGGTCGAGGGCGTT
            (2452)
                    2501                                                2550
« 33H5      (1149) GCGCCCCCGATGCCTCTAATCATTGGCTTTACCCGATAGAACTCGCCCGC
« 156F9     (1150) GCGCCCCCGATGCCTCTAATCATTGGCTTTACCCGATAGAACTCGCCCGC
« 71G5      (1150) GCGCCCCCGATGCCTCTAATCATTGGCTTTACCCGATAGAACTCGCCCGC
« 145E11    (2037) GCGCCCCCGATGCCTCTAATCATNGGCTTTACCCGATAGAACTCGCCCGC
« 46A6     (•1861)
 Consensus  (2501) GCGCCCCCGATGCCTCTAATCATTGGCTTTACCCGATAGAACTCGCCCGC
```

```
                (2502)



                        2551                                                2600
« 33H5          (1199) GGGCTCCAGCTATCCTGAGCCCACGCGTCCGCGGACGCGTGGG
« 156F9         (1200) GGGCTCCAGCTATCCTG
« 71G5          (1200) GGGCTCCAGCTATCCTGAGCCCACGCGTCCGCGGACGCGTGGG
« 145E11        (2087) GGGCTCCAGCTATCCTGAGGGAAA-CTTC-G-G-A-G-G-G--AACCAGC
« 46A6         (•1861)
 Consensus      (2551) GGGCTCCAGCTATCCTGAGCCCACGCGTCCGCGGACGCGTGGGAACCAGC
                (2552)
                        2601                                                2650
« 33H5         (•1241)
« 156F9        (•1216)
« 71G5         (•1242)
« 145E11        (2137) TACTAGACGGTTCGATTAGTCTTTCGCCCCTATACCCAAGTCAGACGAAC
« 46A6         (•1861)
 Consensus      (2601) TACTAGACGGTTCGATTAGTCTTTCGCCCCTATACCCAAGTCAGACGAAC
                (2602)
                        2651                                                2700
« 33H5         (•1241)
« 156F9        (•1216)
« 71G5         (•1242)
« 145E11        (2187) GATTTGCACGTCAGTATCGCTGCGAGCCTCCACCAGAGTTTCCTCTGGCT
« 46A6         (•1861)
 Consensus      (2651) GATTTGCACGTCAGTATCGCTGCGAGCCTCCACCAGAGTTTCCTCTGGCT
                (2652)
                        2701                                                2750
« 33H5         (•1241)
« 156F9        (•1216)
« 71G5         (•1242)
« 145E11        (2237) TCGCCCCGCTCAGGCATAGTTCACCATCTTTCGGGTCCCGACAGGTATGC
« 46A6         (•1861)
 Consensus      (2701) TCGCCCCGCTCAGGCATAGTTCACCATCTTTCGGGTCCCGACAGGTATGC
                (2702)
                        2751                                    2792
« 33H5         (•1241)
« 156F9        (•1216)
« 71G5         (•1242)
« 145E11        (2287) TNTCACTCGAACCCTTCACAGAAGATCTAGGTCGGTCGGCGG
« 46A6         (•1861)
 Consensus      (2751) TNTCACTCGAACCCTTCACAGAAGATCTAGGTCGGTCGGCGG
                (2752)
```

This page was intentionally left blank.

Morphine-pathway block in *top1* poppies

This opium poppy mutant provides chemical precursors for non-addictive analgesics.

The opium poppy is a source of the pharmaceuticals codeine, morphine and their derived analgesics. Here we describe the initial characterization of the poppy mutant known as *top1* (for 'thebaine oripavine poppy 1'), which accumulates the morphine and codeine precursors thebaine and oripavine and does not complete their biosynthesis into morphine and codeine. The original discovery of *top1* stimulated a re-engineering of the opioid industry in the island state of Tasmania, which grows over 40% of the world's licit opiates, in order to produce thebaine and oripavine efficiently from morphine-free poppy crops to provide precursors for highly effective analgesics and for treatment of opioid addiction.

The morphinan branch of the normal poppy pathway, from salutaridine to morphine, involves two demethylation steps and proceeds by two alternative routes (Fig. 1)[1,2]. Although the enzymatic synthesis of morphine has been almost completely elucidated[3–5], knowledge about the genes involved in the morphinan branch, including their coordination and regulation, is sparse.

We treated seeds of a commercial poppy cultivar (*Papaver somniferum*) with a mutagen and then screened the progeny plants.

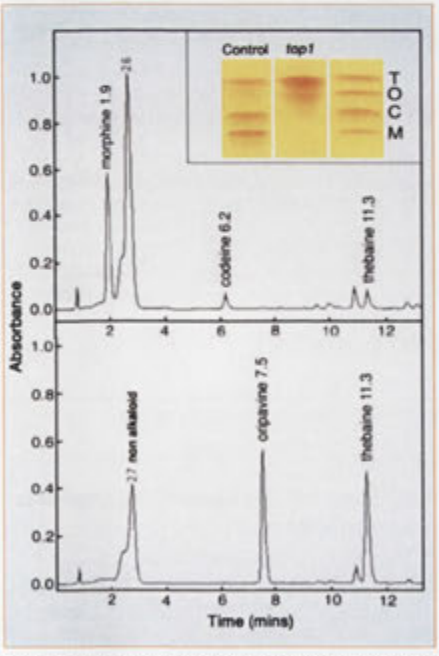


Figure 2 High-pressure liquid chromatography traces of alkaloid profiles from the control parent poppy (upper) and the mutant *top1* (lower). The *top1* trace lacks codeine and morphine peaks. Inset, thin-layer chromatogram of latex from the control and the mutant *top1* (standards in third lane: T, thebaine; O, oripavine; C, codeine; M, morphine). Thebaine and oripavine accumulate as the dominant alkaloids in the mutant at the expense of morphine and codeine, which are evident in the parental phenotype.

The mutant *top1* was found to accumulate thebaine and oripavine but not morphine or codeine (Fig. 2; for details of methods, see supplementary information). The segregation of 375 F_2 individuals was mendelian and was consistent with the involvement of a single gene.

The only visible phenotypic change in the mutant is that its latex is pigmented rather than the normal white. The latex colour change co-segregates with the thebaine phenotype as a reliable visual marker. Feeding of radioactive intermediates (see supplementary information) confirmed that there is a block in both arms of the bifurcated pathway at thebaine and oripavine in *top1*. This block may be due to a defect in the enzyme thebaine demethylase, which is likely to be responsible for the 6-*O*-demethylation of both thebaine and oripavine.

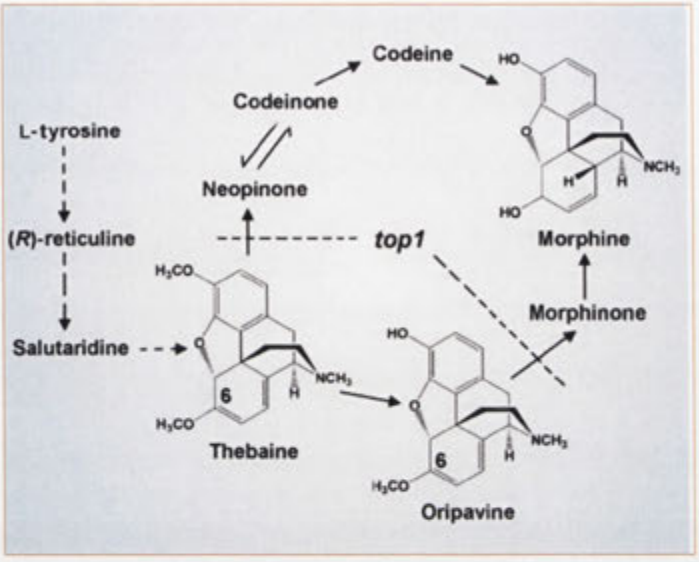


Figure 1 An abbreviated biosynthesis pathway, leading from tyrosine to morphine and showing the two alternative routes from thebaine to morphine. (*S*)-reticuline is a major branchpoint precursor of a number of other alkaloids found in *Papaver somniferum* and in other plant species. The blocks in the pathway in the mutant *top1* are indicated by dotted lines through the enzymatic steps.

We developed a 17,000-gene poppy microarray and used it to explore global changes in gene expression in the mutant. Ten genes were identified as being significantly differentially underexpressed in *top1*, a result verified by quantitative real-time polymerase chain reaction; seven of the clones were highly homologous to known genes of other species (for details, see supplementary information). These include a component of the signal-recognition particle that mediates protein trafficking, a flippase ATP-dependent transmembrane transporter, and a homologue to ftsH protein, a transmembrane ATP-dependent metalloprotease. Three other clones encoded proteins with similarity to known enzymes: phosphoenolpyruvate carboxykinase, aspartate aminotransferase and aldose 1-epimerase.

A comparison of latex proteins revealed many obvious changes in *top1* in both the serum and membrane fractions (results not shown). It seems that the *top1* mutation triggers a series of changes in gene transcripts, proteins and secondary products. None of the candidate genes from the microarray study has yet been proved to be a demethylase or otherwise responsible for the *top1* phenotype. Although the apparently coordinated changes in these transcripts and proteins are not understood, a window is opening into the complex metabolic network associated with the morphine pathway and new avenues for modifying the pathway are emerging.

There are several possible explanations for the *top1* mutation phenotype: the gene encoding the 6-*O*-demethylase that acts on thebaine and oripavine might be affected, as could a gene that regulates its function or expression (such as a transcriptional regulator or a protein in a multiprotein complex). Alternatively, there may be an alteration in a structural or transport component that prevents the substrates thebaine and oripavine from arriving at the subcellular compartment where *O*-demethylation occurs[6–8].

The *top1* mutant is not only contributing to the genetic dissection of the morphine pathway, it has been swiftly deployed for an agriculturally viable supply of thebaine and is a crop that carries little risk of diversion for illicit purposes. Its manufactured derivatives include oxycodone, buprenorphine, naloxone and naltrexone. Buprenorphine is a mixed agonist–antagonist and a potent analgesic with a favourable side-effect profile. It is playing an expanding role in both severe pain control and in the treatment of addiction.

Anthony G. Millgate*†, Barry J. Pogson†, Iain W. Wilson*, Toni M. Kutchan‡, Meinhart H. Zenk§, Wayne L. Gerlach‖, Anthony J. Fist¶, Philip J. Larkin*

*CSIRO Plant Industry, Black Mountain, Canberra, ACT 2601, Australia
e-mail: philip.larkin@csiro.au
†School of Biochemistry and Molecular Biology, The Australian National University, Canberra, ACT 0200, Australia
‡Leibniz Institute of Plant Biochemistry, Department of Natural Product Biotechnology, 06120 Halle (Saale), Germany
§Biozentrum Universität Halle, 06120 Halle (Saale), Germany
||Johnson & Johnson, Locked Bag 4555, Strawberry Hills, New South Wales 2012, Australia
¶Tasmanian Alkaloids, Westbury, Tasmania, 7303 Australia

1. Parker, H. L., Blaschke, G. & Rapoport, H. *Am. Chem. Soc.* **94,** 1276–1282 (1972).
2. Brochmann-Hanssen, E. *Planta Medica* **50,** 343–345 (1984).
3. Kutchan, T. M. *Gene* **179,** 73–81 (1996).
4. Zenk, M.H. in *Organic Reactivity: Physical and Biological Aspects* (eds Maskill, H., Golding, B. T. & Griffin, R. J.) 89–109 (Royal Society of Chemistry, Cambridge, UK, 1995).
5. Kutchan, T. M., Frick, S. & Weid, M. in *Advances in Plant Biochemistry and Molecular Biology* (eds Bohnert, H. J. & Nguyen, H. T.) (Elsevier Science, Oxford, in the press).
6. Bird, D. A., Franceschi, V. R. & Facchini, P. J. *Plant Cell* **15,** 2626–2635 (2003).
7. De-Luca, V. & St Pierre, B. *Trends Plant Sci.* **5,** 168–173 (2000).
8. Bock, A., Wanner, G. & Zenk, M. H. *Planta* **216,** 57–63 (2002).

Supplementary information accompanies this communication on *Nature*'s website.
Competing financial interests: declared none.

Ecology

Widespread colonization by polar hypoliths

High-latitude polar deserts are among the most extreme environments on Earth. Here we describe a large and previously unappreciated habitat for photosynthetic life under opaque rocks in the Arctic and Antarctic polar deserts. This habitat is created by the periglacial movement of the rocks, which allows some light to reach their underside. The productivity of this ecosystem is at least as great as that of above-ground biomass and potentially doubles previous productivity estimates for the polar desert ecozone.

The underside of rocks in climatically extreme deserts acts as a refugium for photosynthetic microorganisms (defined as 'hypoliths') and their community (the 'hypolithon')[1]. Here, the organisms are protected from harsh ultraviolet radiation[2] and wind scouring, and trapped moisture can provide them with a source of liquid water[3]. Colonization also usually requires the rocks to be sufficiently translucent to allow for the penetration of light — all hypoliths reported so far have been found under quartz, which is one of the most common translucent rocks[4,5].

We examined 850 randomly selected opaque dolomitic rocks, without regard to the local patterns of periglacial rock sorting, on Cornwallis Island and Devon Island in the Canadian high Arctic. These are regions typical of extreme polar desert, where vegetation cover was measured to be 1.2% or less[6]. On Devon Island, 95% of rocks were found to be colonized, and on Cornwallis Island, 94% were colonized. The photosynthetic organisms form a well defined green band on the underside of the rocks. The mean (±s.d.) band width of these communities was 3.1 ± 1.9 cm and 3.0 ± 1.6 cm on Devon Island and Cornwallis, respectively.

The Arctic hypoliths are dominated by cyanobacteria. Species found include *Gloeocapsa* cf. atrata Kützing, *Gloeocapsa* cf. punctata Nägeli, *Gloeocapsa* cf. kuetzingiana Nägeli and *Chroococcidiopsis*-like cells; unicellular algal chlorophytes were also present[7].

We investigated the colonization of well developed polygons, which are just one manifestation of a diversity of linear and circular features caused by the long-term action of periglacial processes[8]. In the Arctic, we found colonization on 68% of rocks within polygons, with a mean (±s.d.) band width of 0.7 ± 0.8 cm, where the rocks are surrounded by fine soil sorted into their centres. At the edges, where the cracks around the rocks are larger, we found 100% colonization with a mean (±s.d.) band width of 3.6 ± 1.4 cm (Fig. 1).

We studied similar polygonal terrains at Mars Oasis on Alexander Island in the Antarctic Peninsula, a location where hypolithic colonization occurs. The percentage colonization was 5% within polygons and 100% on the edges of polygons, with mean (±s.d.) band widths of 0.7 ± 0.1 and 2.1 ± 0.3 cm, respectively. We propose that rock sorting by periglacial action, including that during freeze–thaw cycles, improves light penetration around the edges of rocks, one factor that might account for the widespread colonization of the underside of opaque rocks.

Figure 1 Colonization by polar hypolithon. **a,** Hypolithic colonization is enhanced by rock sorting which, among other factors, increases light penetration to the underside of rocks, shown here at the edges of polygonal terrain (white lines indicate polygon edges). **b,** Opaque rocks are colonized on their underside by well defined bands of photosynthetic communities, seen here on an excavated rock.

We estimated the productivity of the Arctic communities by monitoring the uptake of ^{14}C-labelled sodium bicarbonate (for methods, see supplementary information). Assuming no carbon uptake from other sources, a conservative estimate of mean (±s.d.) productivity is about 0.8 ± 0.3 g m^{-2} yr^{-1}. Given that the estimated mean (±s.d.) productivity from plants, lichens and bryophytes on Devon Island is 1.0 ± 0.4 g m^{-2} yr^{-1} (ref. 6), the polar hypolithon may be just as important in sequestering carbon, at least doubling previous estimates of the productivity of the rocky polar desert.

Charles S. Cockell*, M. Dale Stokes†
*British Antarctic Survey, Natural Environment Research Council, High Cross, Cambridge CB3 0ET, UK
e-mail: csco@bas.ac.uk
†Marine Physical Laboratory, Scripps Institution of Oceanography, La Jolla, California 92037-0238, USA

1. Golubic, S., Friedmann, I. & Schneider, J. *J. Sediment. Petrol.* **51,** 0475–0478 (1981).
2. Cockell, C. S., Rettberg, P., Horneck, G., Scherer, K. & Stokes, D. M. *Polar Biol.* **26,** 62–69 (2003).
3. Smith, M. C., Bowman, J. P., Scott, F. J. & Line, M. A. *Antarctic Sci.* **12,** 177–184 (2000).
4. Broady, P. A. *Brit. Phycol. J.* **16,** 231–240 (1981).
5. Schlesinger, W. H. *et al. Ecology* **84,** 3222–3231 (2003).
6. Bliss, L. C., Svoboda, J. & Bliss, D. I. *Holarctic Ecol.* **7,** 305–324 (1984).
7. Cockell, C. S. *et al. Meteor. Planet. Sci.* **37,** 1287–1298 (2002).
8. Kessler, M. A. & Werner, B. T. *Science* **299,** 380–383 (2003).

Supplementary information accompanies this communication on *Nature*'s website.
Competing financial interests: declared none.

brief communications arising online
www.nature.com/bca

Corrigendum: Does gut hormone PYY_{3-36} decrease food intake in rodents?

M. Tschöp, T. R. Castañeda, H. G. Joost, C. Thöne-Reineke, S. Ortmann, S. Klaus, M. M. Hagan, P. C. Chandler, K. D. Oswald, S. C. Benoit, R. J. Seeley, K. P. Kinzig, T. H. Moran, A. G. Beck-Sickinger, N. Koglin, R. J. Rodgers, J. E. Blundell, Y. Ishii, A. H. Beattie, P. Holch, D. B. Allison, K. Raun, K. Madsen, B. S. Wulff, C. E. Stidsen, M. Birringer, O. J. Kreuzer, M. Schindler, K. Arndt, K. Rudolf, M. Mark, X. Y. Deng, D. C. Withcomb, H. Halem, J. Taylor, J. Dong, R. Datta, M. Culler, S. Craney, D. Flora, D. Smiley, M. L. Heiman
Nature **430,** 165 (2004); doi:10.1038/nature02665 (Corrigendum doi:10.1038/nature03019)

Supplementary Methods

Sequencing of Differentially-Expressed Genes

Primers to the vector at the 5' and 3' regions of the inserted cDNA were used to get sequence data of the differentially-expressed genes. Sequences were analysed on BioManager by ANGIS (http://www.angis.org.au). against the NCBI database (GenBank main, SwissProt and SpTr EMBL databases).

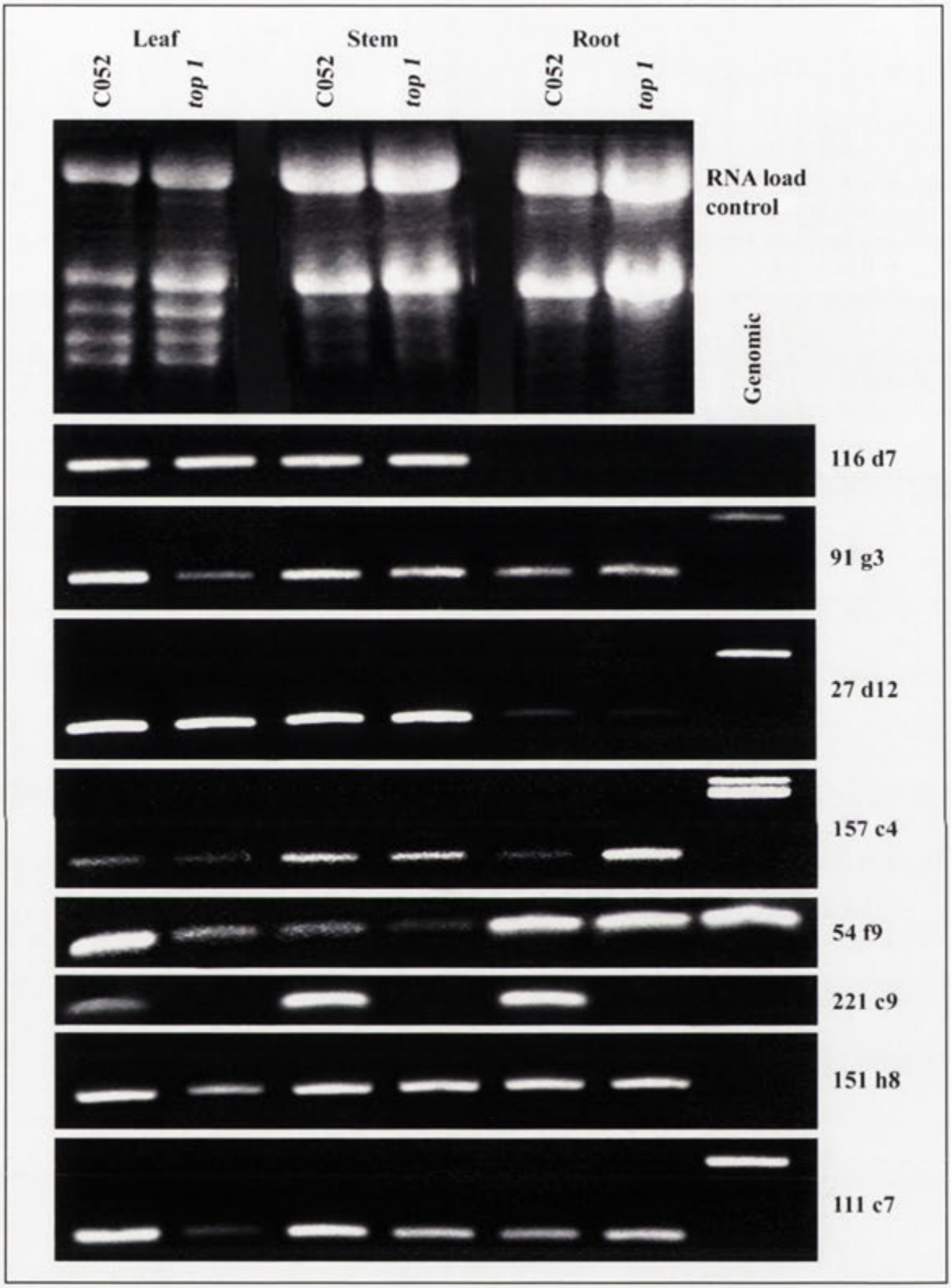
2D protein Analysis.

Latex was collected into water from young green capsules of *top1* and CO52, pooled from 8 plants and matched for developmental stage. Samples were separated into serum and 1000*g* pellet fractions. The serum proteins were acetone precipitated and dried. The 1000g pellet was washed, solubilised with 1% Triton X100 and sonicated, acetone precipitated and dried. At the Australian Proteome Analysis Facility, 100µg samples were separated first with isoelectric focusing pH 3-10 (87 kV hours), and then in the second dimension on 8-18% polyacrylamide gradients (3 h at 3mA, 14 h at 15 mA). Proteins were visualised with SYPRO Ruby fluorescent stain.

Real-Time Quantitative Reverse Transcriptase-Mediated PCR and Qualitative RT-PCR

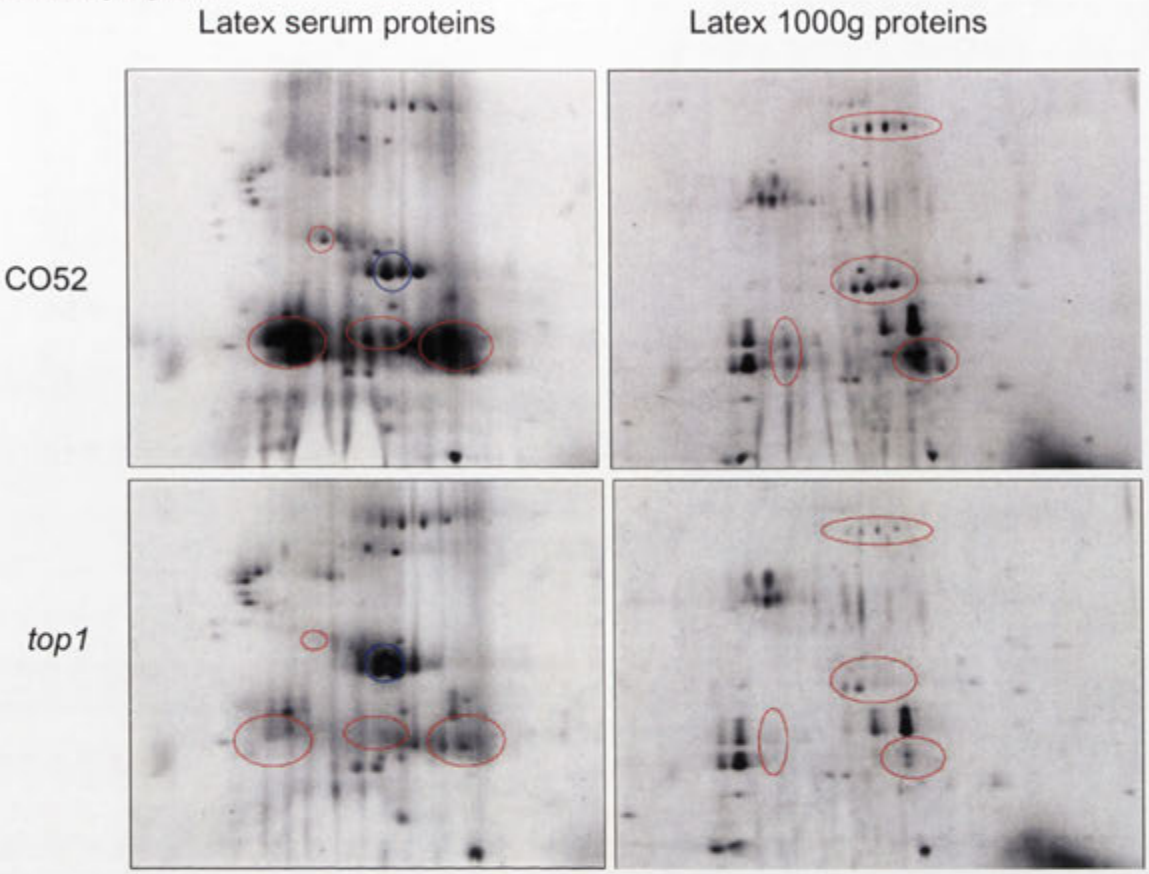
Gene-specific primer sets were designed using the Primer3 software package (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi) and made by Sigma (Sigma-Aldrich, Sydney, Australia). These primers were designed to have a T_m of 59-63 °C (with the optimal being =60 °C), yield a PCR product around 225-275 bp, GC content from 40-60%, self-complementarity <4, and 3' self-complementarity <3. Quantitative Real Time PCR (qRT-PCR) reactions were carried out as per [12]. Ratio values were determined by comparing gene expression in C052 to *top1*. RNA was extracted from de-veined leaf two days after petal drop. Three technical replicates of each sample were tested.

For Reverse transcriptase PCR analysis, RNA was isolated at two days post petal drop from the leaf, stem and roots. 20 ng of cDNA was used per PCR reaction using the primers (each at 6 pmoles) designed for qRT-PCR. Qiagen Hotstar Taq Polymerase mix was used in each 20 µL reaction. An annealing temperature of 50° C was used in the PCR reaction. The number of cycles was adjusted to discern differences in expression levels. Two biological and two technical replicates were run for each sample and set of primers. RNA was quantified by comparison of the ribosomal loading on an RNA gel.



Supplementary Figure 1: Gels of RT-PCR of the nine genes to compare transcript levels between the mutant *top 1* and the control across leaf, stem and root. The RNA gel at the top shows the even loading of the RNA for the RT-PCRs

Supplementary Figure 2



Supplementary Figure 2. Two dimensional protein separations using pH 3-10 isoelectric focussing, followed by 8-18%T polyacrylamide in the second dimension. Proteins were visualised with SYPRO ruby fluorescent stain.

A. 100µg protein from the serum fraction of top1 and CO52

B. 100µg protein from the 1000g pellet fraction of top1 and CO52

In each case some of the areas or difference are circled.

Supplementary Table 2
Microarray expression ratios of control verses mutant poppies and *E.californica*.

Clone name	A	B	C	D	E	*E.californica*
116d7	5.11	nd	6.27	4.34	6.48	6.25
91g3	8.97	nd	14.06	8.68	5.86	6.25
27d12	10.17	1	11.55	9.98	7.52	7.69
157c4	6.77	nd	6.76	5.38	5.5	9.09
91b6	6.2	1.1	10.28	6.24	7.2	2.70
221c9	6.56	nd	7.57	5.32	3.57	4.54
54f9	5.84	nd	8.64	4.92	2.8	4.76
52f2	6.34	nd	7.22	6.84	4.82	6.25
151h8	8.81	nd	7.31	7.64	3.94	5.26
111c7	7.69	0.94	8.79	7.66	6.36	7.14

Supplementary Table 2. Microarray expression ratios of control verses mutant poppies and *E. californica*. The values in the Table are the spatially normalised back-transformed data on the log (base 2) scale giving the expression ratio of the clones in the mutants or *E. californica* (see methods) compared to the control *P. somniferum*. A positive value reveals higher expression in the control (reference sample) than in the mutants or *E. californica* (test sample).

A is a low morphine high thebaine mutant.

B is a reticuline accumulating and morphine free.

C is a mutant with high papaverine low morphine

D is a high oripavine and thebaine mutant, but also with morphine and codeine.

E is a low morphine mutant.

Supplementary Table 1. F2 segregation data for the *top1* phenotype

	Cross 1[a]	Cross 2[b]	Total
Morphine Phenotype	189	83	272
Thebaine Phenotype	81	22	103

Supplementary Table 1. The results of HPLC analysis of the latex of random F2 segregates from 2 independent crosses, both involving a homozygous mutant *top1* and a homozygous morphine accumulating parent line. It was sometimes possible to see a partial thebaine phenotype in heterozygotes because top1 was incompletely dominant rather than recessive. However scoring heterozygotes was unreliable and not attempted in the data shown.

[a,b] Two separate crosses between *top1* and normal morphine lines were made and advanced to F2.

As it was not always possible to distinguish heterozygotes from homozygous wild type, these two categories are combined for the analysis. Chi squared for equality of segregation ratio in the 2 crosses was 3.106, P between 5 and 10%, therefore acceptable that they represent the same genetic ratio.

Chi squared for the totals against a 3:1 ratio with Yates' correction, was 0.44 (P= 0.5068) and therefore conforms to a 3:1 ratio.

nature biotechnology

RNAi-mediated replacement of morphine with the nonnarcotic alkaloid reticuline in opium poppy

Robert S Allen[1], Anthony G Millgate[1], Julie A Chitty[1], Jennifer Thisleton[1], James A C Miller[2], Anthony J Fist[2], Wayne L Gerlach[3] & Philip J Larkin[1]

We report on the silencing of codeinone reductase (COR) in the opium poppy, *Papaver somniferum*, using a chimeric hairpin RNA construct designed to silence all members of the multigene COR family through RNA interference (RNAi). After gene silencing, the precursor alkaloid (*S*)-reticuline—seven enzymatic steps upstream of codeinone—accumulated in transgenic plants at the expense of morphine, codeine, oripavine and thebaine. Methylated derivatives of reticuline also accumulated. Analysis verified loss of *Cor* gene transcript, appearance of 22-mer degradation products and reduction of enzyme activity. The surprising accumulation of (*S*)-reticuline suggests a feedback mechanism preventing intermediates from general benzylisoquinoline synthesis entering the morphine-specific branch. However transcript levels for seven other enzymes in the pathway, both before and after (*S*)-reticuline, were unaffected. This is the first report of gene silencing in transgenic opium poppy and of metabolic engineering to cause the high-yield accumulation of the nonnarcotic alkaloid reticuline.

Opium poppy is one of the oldest cultivated plants, and its analgesic use can be traced to the beginnings of civilization[1,2]. Opium, which is the dried latex of poppy, contains alkaloids presumably involved in ecological defense, with codeine and morphine being two of the most abundant[3]. In certain strains the morphinan intermediates oripavine[4] and thebaine[5] also accumulate to similar levels. Presently, codeine and morphine remain two of the most important and effective analgesics used in medicine worldwide. With five centers of chirality, the structures of morphinan alkaloids present a complexity that renders commercial synthesis uneconomic[6] and cultivation of poppy continues to be the most effective means to produce opiate analgesics[7].

Several other intermediates of the morphine biosynthetic pathway have medicinal importance, particularly in the synthesis of drugs derived from these compounds. Powerful analgesics, such as buprenorphine and oxycodone[8], are manufactured using thebaine as a substrate[9].(*R*)- or (*S*)-reticuline, six and eight steps respectively before morphine, are potential substrates for the manufacture of various compounds that have shown antimalarial[10–13] or anticancer activities[14,15]. Reticuline itself stimulates hair growth[16]. These are examples of morphine pathway intermediates that have attracted interest, and whose availability is limited. It would be advantageous to have cultivars capable of accumulating desired alkaloids that normally exist only as transitory compounds. We demonstrate here for the first time that hairpin RNA silencing is an effective means to produce such modified poppies.

The enzymology of the entire morphine pathway is largely elucidated[17,18] (**Fig. 1**); cDNAs for six of the enzymes involved in forming (*S*)-reticuline from tyrosine have been cloned[18–20], and two genes from the morphine-specific branch are now cloned[21,22]. A beginning has been made in determining the cellular location of morphinan alkaloid biosynthesis[23,24]. Despite these important advances, virtually nothing is known of what regulates the accumulation of morphine and its intermediates in poppy latex. We anticipate that perturbation of specific enzymatic steps through gene silencing will provide information about pathway controls.

The cloning of COR, the penultimate enzyme of morphine biosynthesis, represents important progress[21]. Expression profiles indicated COR was a constitutive enzyme, and it was not induced with elicitor treatment or wounding, in contrast to other morphine biosynthetic genes such as salutaradinol acetyl transferase (SAT)[19]. We wanted to explore the regulatory and biochemical consequences of silencing genes of morphine biosynthesis; *Cor* was the most terminal gene available and was therefore chosen for this study.

The efficacy of gene silencing in plants using inverted-repeat transgene constructs encoding hairpin RNA (hpRNA) has been demonstrated[25–27]. This technology permits one, some or all members of a multigene family to be silenced by targeting sequences that are unique or shared.

COR is encoded by a multigene family[21]. We used RNAi to silence all members and show that silenced *P. somniferum* lines accumulate (*S*)-reticuline, which occurs seven enzymatic steps before COR. The implications of this surprising result of metabolic engineering are discussed together with new insights into morphinan alkaloid regulation. The capacity to engineer crops to accumulate specific

[1]CSIRO Plant Industry, PO Box 1600, Canberra, Australian Capital Territory 2601, Australia. [2]Tasmanian Alkaloids Pty Ltd, PO Box 130, Westbury, Tasmania 7303, Australia. [3]Johnson & Johnson Research Pty Ltd, Locked Bag 4555, Strawberry Hills, New South Wales 2012, Australia. Correspondence should be addressed to P.J.L. (philip.larkin@csiro.au).

Published online 14 November 2004; doi:10.1038/nbt1033

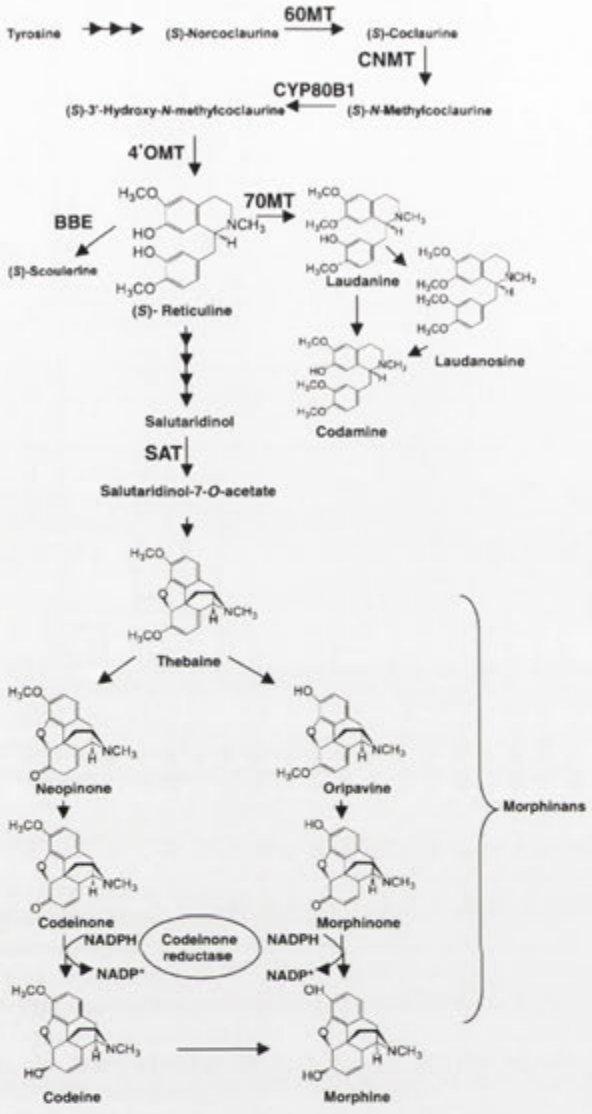


Figure 1 A summary of the biosynthetic pathway leading to morphine, showing also the methylated derivatives of reticuline; enzymes indicated are those for which cDNAs have been previously cloned and that are investigated in this study. Enzyme abbreviations: 6OMT, (*R,S*)-norcoclaurine-6-*O*-methyltransferase; CNMT, (*S*)-coclaurine *N*-methyltransferase; CYP80B1, (*S*)-*N*-methylcoclaurine-3′-hydroxylase; 4′OMT, (*S*)-3′-hydroxy-*N*-methylcoclaurine-4′-*O*-methyltransferase; 7OMT, (R,S)-reticuline 7-*O*-methyltransferase; BBE, berberine bridge enzyme; SAT, salutaridinol 7-*O*-acetyltransferase.

intermediates of pharmaceutical significance may enable deployment of displacement poppy crops, which cannot readily be diverted to illicit uses.

RESULTS

A chimeric hpRNA construct targets the *Cor* gene family

An hpRNA binary vector was designed to silence all members of the *Cor* gene family. The family contains seven members, *Cor1.1–1.6* (AF108432–108437) and *Cor2* (AF108438)[21]. Regions of cDNA from both *Cor1.1* and *Cor2* were incorporated into an hpRNA-based vector driven by the S4S4 promoter from subterranean-clover stunt virus. The region chosen to target *Cor1.1–1.6* transcripts contained many

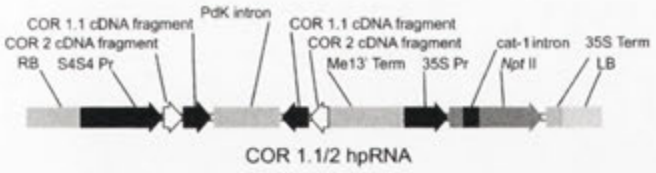


Figure 2 The T-DNA from the transformation vector COR 1.1/2 hpRNA designed to produce hpRNA and initiate silencing of all members of the *Cor* family in *P. somniferum*. Shown are the left border (LB) and right border (RB), S4S4 promoter, *Cor* cDNA fragments, pyruvate orthophosphate dikinase (PdK) intron, malic enzyme (Me1) terminator and selectable marker *nptII* gene (35S promoter and terminator).

22-bp sequences of 100% homology across all six genes. In addition a sequence specific for *Cor2* was included to ensure silencing of this distinctive member of the family. The vector also contained the *nptII* gene with the 35S promoter (**Fig. 2**).

Analyses of the altered alkaloid profiles for hpRNA lines

Thin layer chromatography (TLC) of latex collected from transgenic T_0 lines transformed with the vector COR 1.1/2 hpRNA revealed a drastic alteration of the alkaloid profile (**Fig. 3**). All 18 transgenics consistently displayed a 'reticuline' phenotype, consisting of reticuline and its methylated derivatives codamine, laudanine and laudanosine. This was in stark contrast to the normal TLC pattern observed for nontransgenic controls, which consisted of morphine, codeine, oripavine and thebaine, with no detectable reticuline. Morphine was evident in most of the transgenics, although at substantially lower levels.

High-performance liquid chromatography (HPLC) and mass spectrometry (MS) analysis confirmed the identity of the alkaloids associated with the reticuline phenotype (**Fig. 4**) and enabled quantification of the various alkaloids (**Fig. 5a**). For example, transgenic lines 220-14-2, 215-25-1b, and 215-16-1 accumulated very little morphine, with over 98% of total alkaloid comprised of reticuline and its methylated derivatives, while families 220-2-1 and 220-13-2 still accumulated morphine at up to 71% of total alkaloid. However, all transgenic lines consistently accumulated reticuline and its methylated derivatives to at least 25% and up to 100% of total alkaloids. In contrast, neither the Tasmanian parental line CO58-34 nor any of many commercial poppy cultivars examined, accumulate reticuline beyond 2% of total alkaloids.

Most of the 18 transgenics segregated in the T_1 generation in a manner consistent with a single transgenic locus and there was no obvious association between the amount of reticuline and the number of transgenic loci. For example 220-2-1 and 215-16-1 both had two or three transgenic loci, yet the first had the lowest level of reticuline (25%) and the second had the highest (100%). Conversely, 220-13-2 and 215-25-1b both had a single transgenic locus, yet differed in reticuline content from second lowest (40%) to highest (100%). It was evident from the 18 independent primary transgenic lines that reticuline and its derivatives accumulate in direct proportion to the decrease in morphine (**Fig. 5b**, linear correlation, $r^2 = 0.9386$).

To determine whether the reticuline, codamine, laudanine and laudanosine observed in the hpRNA transgenics were of the (*R*) or (*S*) form, the compounds were separated by HPLC from transgenic latex, fully methylated to laudanosine *in vitro*, and loaded onto chiral columns. Optically pure standards of (*R*)- and (*S*)-laudanosine were available. The reticuline and methylated derivatives present in the transgenic latex samples were predominantly of the (*S*)-isoform.

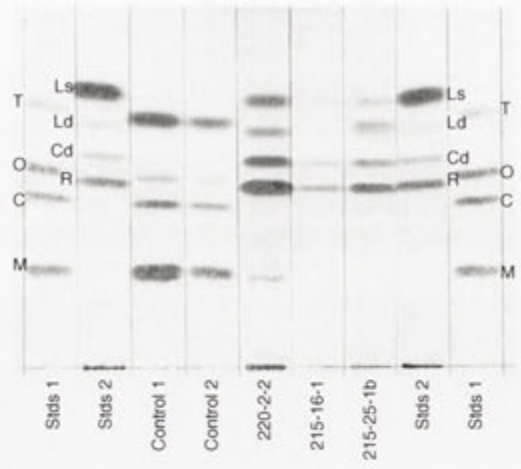


Figure 3 Thin layer chromatography of poppy latex alkaloids. Two sets of standards are shown: stds 1 contains morphine (M), codeine (C), oripavine (O) and thebaine (T); stds 2 contains (*S*)-reticuline (R), codamine (Cd), laudanine (Ld) and laudanosine (Ls); latex from two C058-34 control poppies (control 1 and control 2); latex from three COR hpRNA T_0 transgenics (220-2-2, 215-16-1, 215-251b).

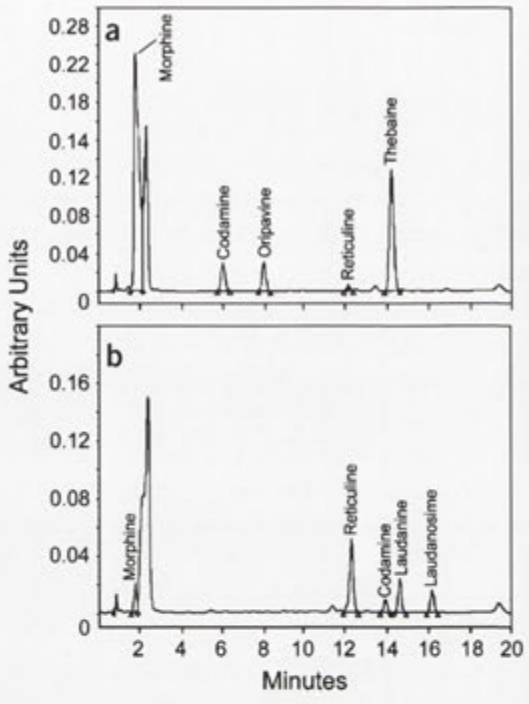


Figure 4 HPLC chromatograms. (**a**) Control nontransgenic latex. (**b**) COR hpRNA transgenic latex. The identity of the major alkaloids was determined by comparison to retention time and spectra of standards. A nonalkaloid major peak is shown eluting at approximately 2.3 min.

COR enzyme activity is lower in COR hpRNA lines

Transgenic poppies carrying the COR hpRNA construct were analyzed for COR activity in leaf extracts. For the forward reaction, the loss of NADPH from the reaction mixture was measured to determine the specific activity of COR. The specific activity for the forward reaction, using codeinone and NADPH as substrates, was considerably lower compared to the control; control poppy leaf had activity of 439 ± 41 pkat mg^{-1} protein, compared to 72 ± 51 pkat mg^{-1} protein for leaf extracts of transgenic plant 220-2-2. Measurement of the backward reaction using $NADP^+$ and codeine as substrates showed the control plant had 62 pkat mg^{-1} protein and transgenic plant 215-14-2 had zero activity.

Cor message is diminished in *Cor* hpRNA transgenics

Poppy lines transformed with the COR hpRNA (hpCOR) were tested for the presence of *Cor* message, and a family of eight T_1 segregants from a high reticuline–accumulating parent (220-2-2) was examined in detail. RT-PCR was initially used to distinguish transgenic T_1 segregants by detection of *nptII* message (**Fig. 6a**). A multiplex RT-PCR assay was developed to detect both *Cor* and actin message simultaneously for this set of segregants. Although actin message was relatively uniform in all T_1 segregants screened, *Cor* message was only visible for nontransgenic segregants (**Fig. 6b**). Using primers specific for *Cor2* message, we showed that *Cor2* transcript was also much lower in the transgenic segregants of 220-2-2 and other families (**Supplementary Fig. 1** online). Northern blotting supported the multiplex RT-PCR analysis. The T_1 segregants had equal amounts of RNA as evidenced by the ribosomal RNA bands; the *Cor* probe strongly hybridized with nontransgenic segregants (dark bands), but not with transgenic segregants (very faint bands) (**Fig. 6c,d**). RNA was separated electrophoretically, blotted to membrane and hybridized to detect the 22 mers generated by Dicer protein. Using a probe derived from the cDNAs in the silencing construct, 22-bp fragments corresponding to *Cor* were detected in the transgenic silenced lines, but no 22-mer fragments were detectable in nontransgenic segregants (**Fig. 6e**). TLC of latex collected from this segregating family confirmed the correspondence between presence of the transgene; loss of COR transcript; appearance of 22-bp RNA degradation products; loss of morphine, codeine, oripavine and thebaine; and appearance of reticuline, codamine, laudanine and laudanosine (**Fig. 6**).

No transcriptional effect on other genes in the pathway

To shed further light on the nature of the altered pathway regulation occurring in the hpCOR plants, transcript levels of other pathway genes were examined, seven in all. Besides *Cor*, only one other gene of the morphine-specific branch after (*S*)-reticuline is cloned, salutaridinol acetyl transferase (SAT)[22]. Relative to controls and nontransgenic segregants, *Sat* transcript was unaffected in the hpCOR transgenics (**Supplementary Fig. 1** online).

Berberine bridge enzyme (BBE) initiates the alternative benzophenanthridine pathway, acting directly on (*S*)-reticuline, forming scoulerine. The gene encoding BBE has been cloned[28,29]. RT-PCR demonstrated that transcript levels of *BBE* are not affected in hpCOR plants relative to controls and nontransgenic segregants (**Supplementary Fig. 1** online).

(*R,S*)-reticuline-7-*O*-methyltransferase (7OMT) converts (*S*)-reticuline into laudanine; a cDNA has been cloned[20]. Because the reticuline-accumulating hpCOR plants also contain elevated levels of laudanine, codamine and laudanosine (**Fig. 1**), we were interested to see if there was upregulation of 7OMT. RT-PCR demonstrated there was no feed-forward transcriptional activation of this step relative to controls (**Supplementary Fig. 1** online).

We also used RT-PCR to examine the transcript levels of the four enzymatic steps before (*S*)-reticuline (**Fig. 1**). These cloned steps were norcoclaurine 6-*O*-methyltransferase (6OMT)[19,20], coclaurine-*N*-methyltransferase (CNMT)[19], (*S*)-*N*-methylcoclaurine

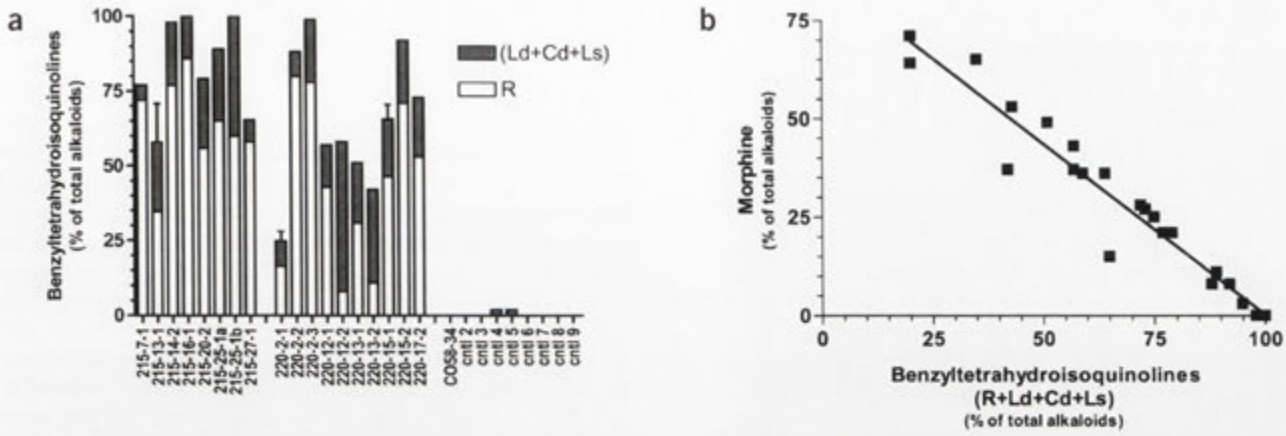


Figure 5 HPLC quantification of the alkaloid contents of 18 T_0 hpRNA transgenics and a set of nine independent, nontransgenic control CO58-34 plants (cntls 1–9). (**a**) The reticuline (R) and codamine + laudanine + laudanosine (Cd +Ld +Ls) contents shown as a percentage of the total alkaloids. (**b**) The relationship between morphine (M) and benzyl tetrahydroisoquinolines (R+Ld+Cd+Ls) shown as a percentage of the total alkaloids in the 18 T_0 hpRNA transgenics.

3′-hydroxylase (CYP80B1)[30] and 3′-hydroxyl-*N*-methylcoclaurine-4′-*O*-methyltransferase (4′OMT)[19,20]. Because the yield of reticuline in hpCOR plants was very high, it was anticipated that earlier steps in the pathway would be unaffected. RT-PCR of *6OMT, CNMT, CYP80B1* and *4′OMT* confirmed that transcript levels for these genes were the same as in controls and nontransgenic segregants (**Supplementary Fig. 1** online).

Stable inheritance of the transgenic phenotype

T_1 individuals from 18 T_0 parental lines were examined for segregation of the transgenes and reticuline phenotype. Segregation of the *nptII* gene was determined with 40–50 individuals for most lines and TLC confirmed the reticuline phenotype was stably inherited, and strictly correlated with the presence of *nptII*. As exemplified with the 220-2-2 family (**Fig. 6f**), the reticuline phenotype was also strictly correlated with the reduction of *Cor* gene transcript; *Cor* message was uniformly diminished and reticuline levels were uniformly high. Most of these families segregated consistently with having either one or two transgenic loci. The alkaloids were quantified by HPLC in over 100 T_1 individuals in a number of lines. Generally the highly silenced T_0 lines with nearly 100% reticuline and methylated derivatives produced T_1 progeny in which all the transgenic segregants were also highly silenced with high reticuline yields.

DISCUSSION

Our results provide an example of an hpRNA-induced major modification of the morphine pathway in opium poppy. Silencing of COR has led to the accumulation of (*S*)-reticuline at the expense of morphine, codeine, thebaine and oripavine. Furthermore, the clear trade-off between reticuline and morphine observed in lines where COR is silenced opens new insights into the regulation of morphine synthesis. The eight-enzyme branch leading to morphine can be substantially downregulated in response to the loss of the penultimate enzyme, COR. Somehow the perturbation of morphinan processing is relayed back to the metabolic junction point at (*S*)-reticuline. This intermediate is shared with other alkaloid pathways in *Papaver* and other genera, responsible for berberine, benzophenanthridines, phthalide isoquinolines and benzyl isoquinolines. We believe this gene silencing–induced feedback is the most dramatic example in secondary metabolism yet reported.

Screening of 18 primary transgenic (T_0) lines by TLC and HPLC revealed a drastic and consistent alkaloid profile alteration, with (*S*)-reticuline and its methylated derivatives laudanine, laudanosine and codamine accumulating at up to 100% of total latex alkaloids. Screening of T_1 individuals from all 18 T_0 transgenics confirmed that this phenotype segregated with the transgene; all nontransgenic segregants displaying a normal phenotype. Northern analysis and RT-PCR verified that *Cor* message of both classes (1 and 2) was substantially reduced in all transgenics, and enzyme assays confirmed a dramatic drop in COR activity. Clearly the hpRNA construct was successful in greatly reducing COR expression and downregulating the entire morphine-specific pathway.

Two previous studies used antisense to suppress berberine bridge enzyme in cell and root cultures of *Eschscholzia californica*, a non-morphine-producing species[31,32]. Decreased levels of benzophenanthridines and an increase in tyrosine pools were observed. However, unlike the present study, no reticuline accumulated in the cell or root cultures. *E. californica* produces sanguinerine but no morphinan alkaloids; clearly the alkaloid metabolic options in *E. californica* cell and root cultures are different and the pathway is regulated quite differently in *P. somniferum* plants.

Although the full chain of events, which has led to the reticuline phenotype, cannot yet be elucidated, we suggest three processes that could be responsible for the result observed. (i) Buildup of the substrates codeinone and neopinone may trigger a negative feedback on one or more earlier enzymes or transport steps of the morphine branch. (ii) The substrate feedback may act to inhibit transcription of genes encoding earlier enzymes or transporters. (iii) Loss of COR enzyme from a larger interdependent enzyme complex may disable the other enzyme reactions normally associated with the complex. These three possibilities are considered in turn.

Negative feedback on morphinan alkaloid synthesis may occur directly on enzyme activity. The substrates codeinone and morphinone may accumulate in the absence of COR and could directly inhibit the ability of reticuline oxidase to catalyze the oxidation of (*S*)-reticuline, which forms the 1,2-dehydroreticulinium ion as the entry point to the morphine pathway. The importance of

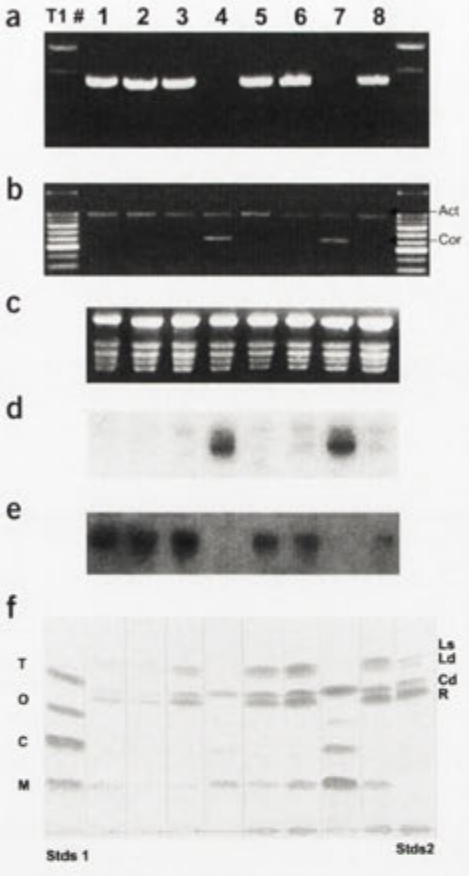


Figure 6 Analysis of T_1 family segregants derived from hpCOR parent 220-2-2. (**a**) *NptII* RT-PCR. (**b**) Multiplex RT-PCR of *Cor* message and actin message; bands corresponding to actin and *Cor* are labeled. (**c**) Ribosomal RNA. (**d**) Northern analysis of *Cor*. (**e**) Detection of 22-mer *Cor* fragments. (**f**) TLC analysis; stds 1 contains morphine (M), codeine (C), oripavine (O), thebaine (T). Stds 2 contains reticuline (R), codamine (Cd), laudanine (Ld) and laudanosine (Ls).

compartmentalization in alkaloid synthesis and accumulation has been indicated[23,24,33], and the possibility exists that the feedback inhibition operates through a perturbation of the normal transport and storage architecture. For example (*S*)-reticuline may be unable to move to the appropriate subcellular compartment for processing into morphinan alkaloids. It has been shown previously that (*S*)-reticuline is preferentially transported into vacuoles compared to the (*R*)-isomer[34]. Salutaridine synthase is responsible for the synthesis of salutaridine from (*R*)-reticuline, a defining chemical step in morphine synthesis[6]. If the silencing of COR results in the feedback inhibition of salutaridine synthase, then the inability of (*R*)-reticuline to be sequestered may account for the observed accumulation of (*S*)-reticuline.

Alternatively the inhibition may act at the transcriptional level, for example, by codeinone or morphinone acting directly or indirectly to block transcription of the reticuline oxidase gene. We demonstrated that the transcript levels of *6OMT*, *CNMT*, *CYP80B1*, *4'OMT* and *Sat* are all unaffected in the hpCOR plants. Furthermore, transcript levels for both enzymes that can act on (*S*)-reticuline to divert it from morphine synthesis, BBE and 7OMT, are unaffected. These results confirm that there is no general impact on transcription in the early part of the pathway, nor on the chief alternative branch pathways from (*S*)-reticuline. The result with *Sat* message indicates that the regulation leading to (*S*)-reticuline accumulation does not operate by coordinated transcriptional suppression of all the morphine branch genes.

However, it must be noted that the substrates codeinone and morphinone do not accumulate in wild-type or transgenic poppy and are undetectable by HPLC. Presumably these compounds have a fast turnover or are unable to be stored. If feedback inhibition by codeinone or morphinone is responsible for the reticuline phenotype, the pathway must be sensitive to small quantitative changes in these compounds. Further studies of feedback inhibition are hampered because the genes for reticuline oxidase and salutaridine synthase have not been cloned.

The third general hypothesis to explain the accumulation of (*S*)-reticuline concerns multienzyme complexes. In flavonoid biosynthesis in *Arabidopsis thaliana*, chalcone synthase, chalcone isomerase and dihydroflavonol-4-reductase are physically associated[35]. Similarly a multienzyme complex is implicated in polyamine synthesis[36]. It is postulated that such structures channel metabolites down a particular pathway, passing intermediates from one enzyme to the next. Such complexes, sometimes called metabolons, have been invoked to explain some features of *O*-methyltransferases in isoflavonoid biosynthesis[37] and their involvement has been postulated in benzylisoquinoline synthesis[38].

Two recent studies reported immunolocalization and *in situ* hybridization of alkaloid enzymes and transcripts in poppy[23,24]. These reports differed substantially on whether sieve elements were involved and on the location of COR. However, both implicated at least two cell types in morphine biosynthesis, laticifers and phloem parenchyma. If, as they indicated, there is spatial separation between accumulating products and enzymes, this greatly complicates a metabolon model. Morphine, codeine, oripavine, thebaine and, in the COR hpRNA transgenics, reticuline, all accumulate in the laticifer and yet most or all of the enzymes are in neighboring cells. Complex movement of intermediates is implicated between cell types and subcellular compartments. The present study has demonstrated that all tetrahydrobenzylisoquinoline synthetic enzymes upstream of (*S*)-reticuline formation are unaffected by the silencing of COR. However, physical associations between enzymes specific to the morphine branch of the pathway cannot yet be ruled out. Enzyme studies on hpCOR transgenics, or silencing of genes for other morphinan-specific enzymes may help resolve any interdependency between these enzymes.

Although the specific regulatory mechanisms responsible for the transgenic reticuline phenotype have not been resolved, the results presented here focus attention on (*S*)-reticuline as a central metabolite of importance in the regulation of morphine synthesis. Thebaine and oripavine can both accumulate to high levels in normal poppy; therefore, our initial expectation of silencing COR was to see the accumulation of these intermediates. The surprising accumulation of (*S*)-reticuline demonstrates that the highly specialized morphinan-alkaloid pathway can be coordinately regulated independently of the more general benzylisoquinoline pathway. Our results provide an example of the usefulness of hpRNA technology to dissect regulatory aspects of secondary metabolism. We have demonstrated the capability to concentrate novel alkaloids of potential commercial significance to levels comparable to the current high yields of morphine and codeine.

METHODS

Construction of *Cor* chimeric cDNA hpRNA silencing vector. pPLEX X002 vector[39,40] was digested with *Cla*I, blunted with *Pfu* polymerase and treated

with Calf Intestinal Alkaline Phosphatase. The cloning vector pBC KS$^+$ containing the orthophosphate dikinase (PdK) intron was digested with *Hin*cII and *Eco*RV to excise the intron. The intron fragment was ligated to the vector prepared above to produce pPLEX X002i. RT-PCR was then used to amplify a 336-bp fragment of *Cor* 1.1 (GenBank accession no. AF108432) from poppy RNA (cultivar C058-34). This fragment included a *Bam*HI site at the 5′ end, derived from the forward primer, necessary for successive cloning steps. Primers used were (F) (5′-**GGATCC**CATCACTTCCAAGCTCTGGTGC GCTGATGCTCAC-3′) and (reverse) (5′-GGGCTCATCTCCACTTGATTCACA ACTG-3′). The *Bam*HI site is indicated in bold. This fragment corresponded to *Cor* 1.1 cDNA from nucleotide 246 through 575. The *Cor* 1.1 fragment was ligated into pGEM-T. A positive clone was designated pGEM-T/Cor1.1. RT-PCR was used to amplify a 242-bp *Cor2* (AF108438) fragment from total poppy RNA (cultivar C058-34). This fragment included *Sac*II and *Eco*RV sites at the 5′end derived from the forward primer, and a *Bam*HI site at the 3′ end derived from the reverse primer, necessary for successive cloning steps. Primers used were (F) (5′-TCC**CCGCGGGATATC**ATGGACTCCGAGGTGCTGAACCAGAT TTCCCA-3′) and (reverse) (5′-CGC**GGATCC**TTAATGGAGACAAAAGGAT CACCGGTTGACAC-3′). The incorporated restriction enzyme sites are indicated in bold. This cDNA fragment corresponded to *Cor2* cDNA from nucleotide 688 through 923. pGEM-T/Cor 1.1(vector) was cut with *Sac*II and *Bam*HI, and the *Cor2* fragment was cut with *Sac*II and *Bam*HI and ligated to the vector. A positive clone was designated pGEM-T/ Cor1.1/2. The vector pPLEX X002i was digested with *Avr*II and *Hpa*I. pGEM-T/Cor1.1/2 was digested with *Eco*RV and *Spe*I. The Cor1.1/2 insert was ligated to cut X002i. A clone was selected that contained the correct size fragment and this was designated pPLEX X002i/Cor1.1/2 AS arm, which was digested with *Sna*BI and *Xho*I. pGEM-T/Cor1.1/2 was digested with *Eco*RV and *Sal*I. The Cor1.1/2 insert was ligated to cut the vector. A positive clone was identified and is designated COR 1.1/2 hpRNA.

Poppy hpCOR transformations. The construct COR 1.1/2 hpRNA was electroporated into *Agrobacterium tumefaciens* strain AGL1 (ref. 41) and a single colony was used to transform poppy cultivar CO58-34 as previously described[42]. Paromomycin at 25 mg/l was used as the selection agent.

Determination of transgenic status of T_1 progeny. Paromomycin leaf painting was used on young plants before rouging plants in pots to ensure both transgenic and nontransgenic segregants were grown for analysis. A solution of 2% paromomycin sulfate (Sigma) and 0.001% Tween 20 was applied with a cotton bud to the leaves. After 4–5 d, plants were scored as either tolerant or susceptible on the basis of necrosis at the site of painting. The test was repeated and the transgenic status was confirmed by the expression of the selectable marker gene *Npt*II using RT-PCR (see below).

RT-PCR analysis of hpRNA lines. Within a day of anthesis of the main flower, young upper leaves were harvested and frozen using liquid nitrogen. RNA was extracted using a Qiagen RNAeasy plant mini-kit, with Qiagen Rnase-free Dnase treatment. For RT-PCR, primers were designed that corresponded to *Cor* 1.1–1.4 cDNAs. The primer binding regions chosen had identical sequences for all four functional isoforms. Forward (F) and reverse (R) primers were (F): (5′-GAGAGTAATGGTGTACCTATGATCACTCTCAGTTC-3′) and (R): (5′-GG GCTCATCTCCACTTGATTCACAACTG-3′). Primers were also designed for the specific amplification of *Cor* 2 cDNA, (F): (5′- ATGGACTCCGAGGTGCT GAACCAGATTTCCCA-3′) and (R): (5′-TTAATGGAGACAAAAGGATCAC CGGTTGACAC-3′). Primers used for *npt*II were: (F): (5′-GCACAACAGA CAATCGGCTGCTC-3′) and (R): (5′- AGCACGAGGAAGCGGTCAGC-3′).

Primers were designed as follows to amplify (*S*)-*N*-methylcoclaurine 3′-hydroxylase (*CYP80B1*), salutaridinol acetyl transferase (*SAT*), berberine bridge enzyme (*BBE*), (*R,S*) norcoclaurine 6-*O*-methyltransferase (*6OMT*), (*R,S*) reticuline 7-*O*-methyltransferase (*7OMT*), (*S*)-coclaurine-*N*-methyltransferase (*CNMT*) and (*S*)-3′-hydoxy-*N*-methylcoclaurine-4′-methyltransferase (*4′OMT*).

CYP80B1. (F): (5′-GAGATTCAAGTCCTAAAGGTTTGCCACCAGGTC-3′) and (R): ('5-GTTTCATTACATTCAGACCACACAATAGAGTTCTCCAC-3′);

SAT. (F): (5′-CCATTATCAATCCTGTTAAACAGTTAAACAC-3′) and (R): (5′-GGATAAAAGTGAACTAGTGTTTTGGAAAGAGAACTC-3′);

BBE. (F): (5′-GTACGAGGTGGTGATGTTAATGATAATCTCCTCTC-3′) and (R): (5′-CACGACATTATCCGCAGCTAATCCGTACTTCCTC-3′);

6OMT. (F): (5′-GAGCTCAAATCATTCAATCATTCTTCTCATCAACAG-3′) and (R): (5′-GAATCAACCATTGATTTCTCCCAGCCTCTC-3′);

7OMT. (F): (5′-CTGCAGAAGAAAGGTTGAAAGGGCAAGCTG-3′) and (R): (5′-GATCCCAAATCTCACATCCATGAGCTCTCTC-3′);

CNMT. (F): (5′-TCTGTCTGAACTTCACTGGCAAGATCAATAGAC-3′) and (R): (5′-TCTCCATATACAAATCCAGCATCGCTATCTCAG-3′);

4′OMT. (F): (5′-CCGGTGATATGTTCATGTCTGTTCCAAGTG-3′) and (R): (5′-GGAAATAGGGAACATCTCCTCAAGTGAAGATGAC-3′).

Primers corresponding to *Ratus norvegicus* β-actin were (F): (5′-GGAGAA GATTTGGCATCACACTTTCTACAATGAG-3′) and (R): (5′-CTTCCTGATAT CCACAATCACACTTCATGATGG-3′). A Qiagen One-step RT-PCR kit was used for RT-PCR reactions. Reactions were performed as per the manufacturer's instructions. Cycling conditions were as follows: 1 cycle of 50 °C for 30 min; 27 cycles of 94 °C for 45 s, 65 °C for 45 s, 72 °C for 1 min; 1 cycle of 72 °C for 10 min. For the *NptII* RT-PCR reaction 30 denaturation/annealing/extension cycles were used. We combined 12 µl of each reaction with 3 µl of loading dye and ran it on a 1.3% agarose gel in Tris-acetate-EDTA running buffer stained with ethidium bromide. Gels were visualized using a Bio-Rad Gel-Doc system.

Northern analysis of transcripts. Denatured total RNA (15 µg) was run on 1.5% agarose gel containing formaldehyde and blotted for northern analysis. Hybridization was carried out at 65 °C in a SDS-phosphate buffer (0.5 M phosphate buffer pH 7.2, 7% SDS, 10 mM EDTA) using a ~570-bp cDNA fragment corresponding to *Cor*1.1–1.4 generated by the same primers used for RT-PCR analysis. The fragment was labeled using a Ready-To-Go DNA labeling kit (Amersham Biosciences). Membranes were washed twice at 65 °C with 2× SSC, 0.1% SDS for 15 min, then once with 0.1× SSC, 0.1% SDS. Blots were exposed on Kodak Biomax MS film with an intensifying screen at –70 °C for 12 h and developed with a Fuji CP 1000 developer.

Detection of 22-mer fragments. Denatured total RNA in 12-µl volumes was run on a 15% acrylamide gel containing 1 × Tris/borate/EDTA, 7 M urea, 2.6 mM ammonium persulfate, 3.3 M N,N,N′N′-tetramethyl-ethylenediamine. RNA was electroblotted to Hybond XL membrane. Prehybridization was carried out at 42 °C for 3 h in formamide buffer (0.125 mM sodium phosphate, 250 mM $NaCl_2$, 7% SDS, 50% formamide). A ~ 550 bp PCR product was amplified from COR 1.1/2 hpRNA using the same primers (*Cor2* F and *Cor1.1* R) used for construction of the hpRNA binary. This fragment was labeled with ^{32}P-dUTP using SP6 RNA polymerase to generate a riboprobe. The probe was then treated with carbonyl to create ~50-bp fragments by adding 300 µl of carbonate solution (80 mM $NaHCO_3$, 120 mM Na_2CO_3) to 20 µl of the purified probe and incubating at 60 °C for 3 h. After incubation 20 µl of 3M sodium acetate (pH 5) was added to the reaction and this was then added to the hybridization solution. Hybridization was carried out overnight at 42 °C. The membrane was washed twice for 30 min in 2× SSC/0.2% SDS at 42 °C and treated with 2 µg/ml RNase in 2× SSC. The membrane was exposed on Kodak Biomax MS film with an intensifying screen at –70 °C for 12 h and developed with a Fuji CP 1000 developer.

TLC analysis of T_0 and T_1 hpCOR plants. Small samples (50–100 µl) of latex were collected from cut green capsules on a spatula tip and suspended in 200 µl of the latex extraction buffer (8.65 mM SDS, 200 mM $NH_4H_2PO_4$, 20% ethanol, pH 4.7). Samples were vortexed then centrifuged for 4 min at 7,200*g*. We loaded 10 µl of supernatant onto the origin of Merck silica-gel-60 plastic-backed plates. Standards were supplied by Tasmanian Alkaloids; morphine, oripavine, thebaine, codeine, codamine, laudanine, laudanosine and reticuline were dissolved in ethanol and NH_3 (3:1) at 2–5 mg ml^{-1} and 10 µl was loaded for TLC. The plates were developed with 40% toluene, 40% acetone, 6% ethanol, 2% ammonia. Dragendorff reagent (25 mM $Bi(NO_3)_3$, 0.8 M KI, 14% glacial acetic acid) was sprayed onto the plates to visualize the alkaloids.

Enzyme assays. We froze in liquid N_2 3-g (fresh weight) stem or capsule tissue and ground it to powder in the presence of 300 mg insoluble polyclar

(polyvinyl polypyrolidone). As the sample thawed, 10 ml of buffer (100 mM Tris, 10% (vol/vol) glycerol, 1 mM DTT) was added and mixed. Protease inhibitors were added to the buffer immediately before the extraction (0.1 mM PMSF, 1 μg/ml leupeptin, 1 μg/ml pepstatin, 1 μg/ml E64 (*N*-(*trans*-epoxysuccinyl)-L-leucine 4-guanidinobutylamide)). The mixture was centrifuged at 10,000*g* for 10 min and the supernatant was used for assays. The protein concentration was determined using the Bradford method (Biorad protein assay reagent) and bovine serum albumin as standard.

We assayed the forward reaction, as described[43]. The assay buffer contained 75 μM codeinone, 150 μM NADPH, 100 mM sodium phosphate buffer, pH 7.0, and 1 mM DTT. Total protein per assay was 100 μg and assays were run for 60 min at 37 °C. Controls included boiled extracts (15 min boiling water bath) to inactivate the enzyme.

Reactions were stopped with 1 ml chloroform and 30 s of vortexing. The lower chloroform phase containing both codeine and codeinone was removed. The absorption of the aqueous phase was read at 340 nm to determine the rate of oxidation of NADPH.

The reverse reaction assay mixture comprised 0.4 mM codeine, 0.6 mM $NADP^+$, 200 mM glycine/NaOH and 1 mM DTT. 200 μg of protein (200 μl) was used per assay together with 800 μl of reaction mix in a 2.5 ml tube. We used an equal volume of boiled enzyme as a negative control. Reactions for both live and boiled extracts were performed in duplicate at 37 °C for 0, 5 and 20 min. Reactions were stopped with 1 ml of chloroform and 30 s of vortexing. The chloroform phase (containing codeine and codeinone) was removed. The extraction was repeated with a further 1 ml chloroform extraction. The absorbance of the aqueous phase was read at 340 nm to determine the rate of formation of NADPH.

HPLC and mass spectrometry of hpCOR latex. Latex was collected from mature capsules and suspended in the latex extraction buffer used in the TLC analysis described above. Samples were run on an Alltech platinum C18 53 × 7 mm rocket column at a flow rate of 1.5 ml/min and analyzed using a Waters 2487 Dual wavelength detector and a Millennium 32 version 3.05.01 data system with UV detection of peaks at 254 nm. Standards used were morphine, codeine, oripavine, salutaridine, reticuline, laudanine, thebaine, laudanosine, papaverine and noscapine. Chiral determinations were conducted by Alison Ung and Stephen Pyne at the University of Wollongong. For chiral determination of reticuline, laudanine and laudanosine, samples collected from the HPLC system were methylated (in the case of reticuline and laudanine) to form laudanosine, before separating with a normal phase Daicel chiral HPLC column (Chiral OD-H 25 cm (length) × 4.6 mm (internal diameter)), which used 5-μm particles of silica gel. The flow rate was 1 ml/min. Standards used were (S)-laudanosine and (R)-laudanosine (Tasmanian Alkaloids).

Note to be added in proof: The enzyme referred to in this paper as reticuline oxidase has now been partially purified and should be called 1,2-dehydroreticuline synthase[44].

Note: Supplementary information is available on the Nature Biotechnology website.

ACKNOWLEDGMENTS

We would like to thank Neil Smith and Peter Waterhouse for helpful discussions and 22-mer detection; Pat Hallam, Peta Dolan and Noel Davies for HPLC and LC/MS work; Stephen Pyne and Alison Ung for chiral determinations; Toni Kutchan, Meinhart Zenk and Susanne Frick for helpful discussions.

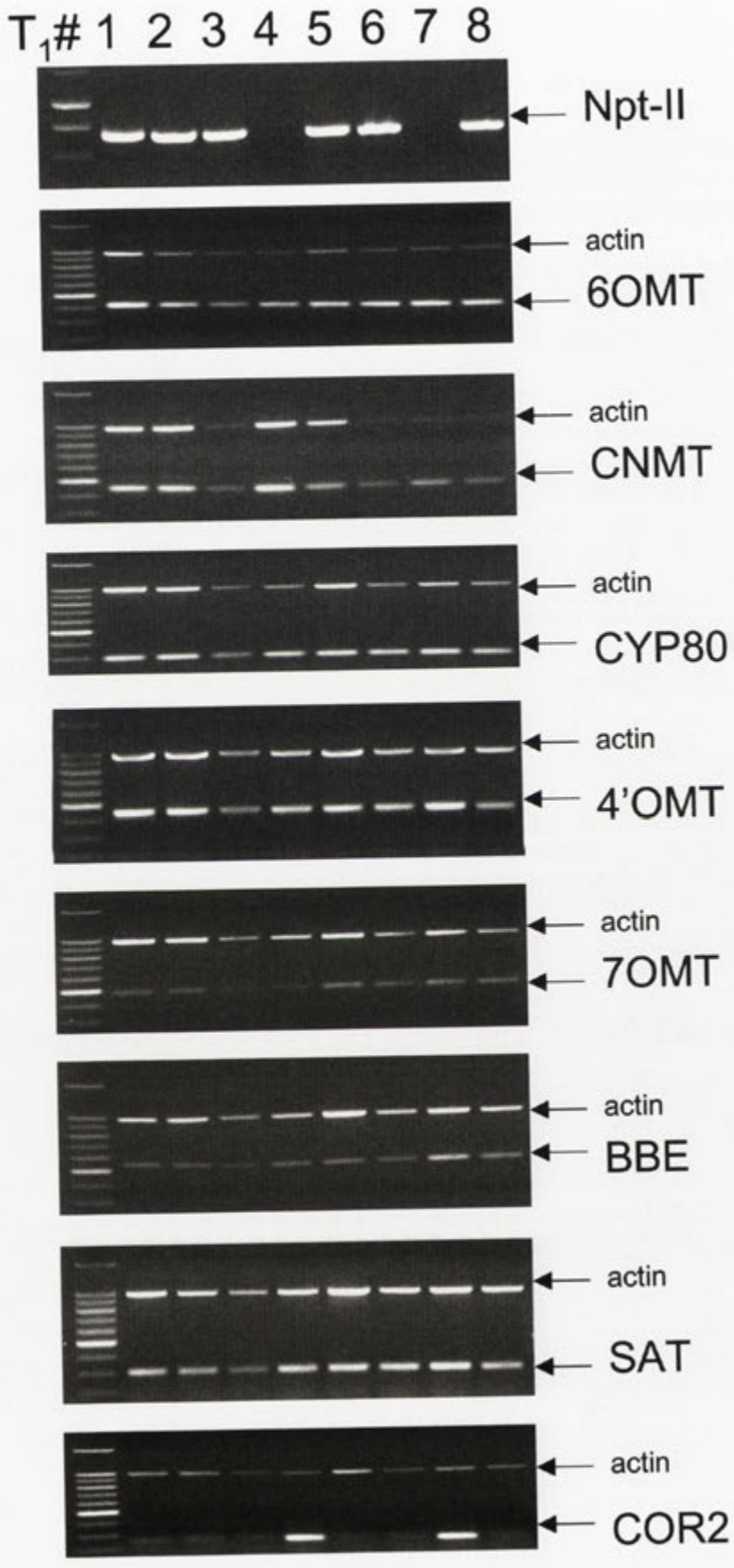
COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 12 April; accepted 22 September 2004
Published online at http://www.nature.com/naturebiotechnology/

1. Meijerink, W.J.H.J., Molina, P.E. & Abumrad, N.N. Mammalian opiate alkaloid synthesis: lessons derived from plant biochemistry. *Shock* **12**, 165–173 (1999).
2. Schiff, P.L.J. Opium and its alkaloids. *Amer. J. Pharm. Educ.* **66**, 186–194 (2002).
3. Caporale, L.H. Chemical Ecology—a view from the pharmaceutical industry. *Proc. Natl. Acad. Sci. USA* **92**, 75–82 (1995).
4. Nielsen, B., Roe, J. & Brochmann-Hanssen, E. Oripavine—a new opium alkaloid. *Planta Med.* **48**, 205–206 (1983).
5. Tetenyi, P. Opium poppy (*Papaver somniferum*): botany and horticulture. *Hortic. Rev.* **19**, 373–408 (1997).
6. Gerardy, R. & Zenk, M.H. Formation of salutaridine from (*R*)-reticuline by a membrane-bound cytochrome-P-450 enzyme from *Papaver somniferum*. *Phytochemistry* **32**, 79–86 (1993).
7. Laughlin, J.C., Chung, B. & Beattie, B.M. in *Poppy, The Genus Papaver* (ed. Bernath, J.) 249–277 (Hardwood Academic Publishers, The Netherlands, 1998).
8. Dewick, P.M. in *Medicinal Plant Products* (ed. Dewick, P.M.) 291–400 (John Wiley & Sons, New York, 2002).
9. Millgate, A.G. *et al.* Morphine-pathway block in top1 poppies. *Nature* **431**, 413–414 (2004).
10. Camacho, M.D. *et al. In vitro* activity of *Triclisia patens* and some bisbenzylisoquinoline alkaloids against *Leishmania donovani* and *Trypanosoma brucei brucei*. *Phytother. Res.* **16**, 432–436 (2002).
11. Rasoanaivo, P. *et al.* Alkaloids of *Hernandia voyronii*: chloroquine-potentiating activity and structure elucidation of herveline D. *Planta Med.* **64**, 58–62 (1998).
12. Tshibangu, J.N., Wright, A.D. & Konig, G.M. HPLC Isolation of the anti-plasmodially active bisbenzylisoquinone alkaloids present in roots of *Cissampelos mucronata*. *Phytochem. Analysis* **14**, 13–22 (2003).
13. Angerhofer, C.K., Guinaudeau, H., Wongpanich, V., Pezzuto, J.M. & Cordell, G.A. Antiplasmodial and cytotoxic activity of natural bisbenzylisoquinoline alkaloids. *J. Nat. Prod.* **62**, 59–66 (1999).
14. Chen, Q., Peng, W.L., Qi, S.J. & Xu, A.L. Apoptosis of human highly metastatic lung cancer cell line 95-D induced by acutiaporberine, a novel bisalkaloid derived from *Thalictrum acutifolium*. *Planta Med.* **68**, 550–553 (2002).
15. Seifert, F., Todorov, D.K., Hutter, K.J. & Zeller, W.J. Cell cycle effects of thaliblastine. *J. Cancer Res. Clin. Oncology* **122**, 707–710 (1996).
16. Nakaoji, K., Nayeshiro, H. & Tanahashi, T. Norreticuline and reticuline as possible new agents for hair growth acceleration. *Biol. Pharm. Bull.* **20**, 586–588 (1997).
17. Brochmann-Hanssen, E. A second pathway for the terminal steps in the biosynthesis of morphine. *Planta Med.* **50**, 343–345 (1984).
18. Kutchan, T.M. in *The Alkaloids* vol. 50 (ed. Cordell, G.) 253–316, (Academic Press, San Diego, 1998).
19. Facchini, P.J. & Park, S.U. Developmental and inducible accumulation of gene transcripts involved in alkaloid biosynthesis in opium poppy. *Phytochemistry* **64**, 177–186 (2003).
20. Ounaroon, A., Decker, G., Schmidt, J., Lottspeich, F. & Kutchan, T.M. (*R,S*)-Reticuline 7-*O*-methyltransferase and (*R,S*)-norcoclaurine 6-*O*-methyltransferase of *Papaver somniferum*—cDNA cloning and characterization of methyl transfer enzymes of alkaloid biosynthesis in opium poppy. *Plant J.* **36**, 808–819 (2003).
21. Unterlinner, B., Lenz, R. & Kutchan, T.M. Molecular cloning and functional expression of codeinone reductase: the penultimate enzyme in morphine biosynthesis in the opium poppy *Papaver somniferum*. *Plant J.* **18**, 465–475 (1999).
22. Grothe, T., Lenz, R. & Kutchan, T.M. Molecular characterization of the salutaridinol 7-*O*-acetyltransferase involved in morphine biosynthesis in opium poppy *Papaver somniferum*. *J. Biol. Chem.* **276**, 30717–30723 (2001).
23. Bird, D.A., Franceschi, V.R. & Facchini, P.J. A tale of three cell types: alkaloid biosynthesis is localized to sieve elements in opium poppy. *Plant Cell* **15**, 2626–2635 (2003).
24. Weid, M., Ziegler, J. & Kutchan, T.M. The roles of latex and the vascular bundle in morphine biosynthesis in the opium poppy. *Papaver somniferum*. *Proc. Natl. Acad. Sci. USA* **101**, 13957–13962 (2004).
25. Wesley, S.V. *et al.* Construct design for efficient, effective and high-throughput gene silencing in plants. *Plant J.* **27**, 581–590 (2001).
26. Wang, M.B. & Waterhouse, P.M. Application of gene silencing in plants. *Curr. Opin. Plant Biol.* **5**, 146–150 (2002).
27. Helliwell, C.A., Wesley, S.V., Wielopolska, A.J. & Waterhouse, P.M. High-throughput vectors for efficient gene silencing in plants. *Funct. Plant Biol.* **29**, 1217–1225 (2002).
28. Dittrich, H. & Kutchan, T.M. Molecular cloning, expression, and induction of berberine bridge enzyme, an enzyme essential to the formation of benzophenanthridine alkaloids in the response of plants to pathogenic attack. *Proc. Natl. Acad. Sci. USA* **88**, 9969–9973 (1991).
29. Bird, D.A. & Facchini, P.J. Berberine bridge enzyme, a key branch-point enzyme in benzylisoquinoline alkaloid biosynthesis, contains a vacuolar sorting determinant. *Planta* **213**, 888–897 (2001).
30. Pauli, H.H. & Kutchan, T.M. Molecular cloning and functional heterologous expression of two alleles encoding (*S*)-*N*-methylcoclaurine 3′-hydroxylase (CYP80B1), a new methyl jasmonate-inducible cytochrome P-450-dependent mono-oxygenase of benzylisoquinoline alkaloid biosynthesis. *Plant J.* **13**, 793–801 (1998).
31. Park, S.U., Yu, M. & Facchini, P.J. Antisense RNA-mediated suppression of benzophenanthridine alkaloid biosynthesis in transgenic cell cultures of California poppy. *Plant Physiol.* **128**, 696–706 (2002).
32. Park, S.U., Yu, M. & Facchini, P.J. Modulation of berberine bridge enzyme levels in transgenic root cultures of California poppy alters the accumulation of benzophenanthridine alkaloids. *Plant Mol. Biol.* **51**, 153–164 (2003).
33. Bock, A., Wanner, G. & Zenk, M.H. Immunocytological localization of two enzymes involved in berberine biosynthesis. *Planta* **216**, 57–63 (2002).

34. Deus-Neumann, B. & Zenk, M.H. Accumulation of alkaloids in plant vacuoles does not involve an ion-trap mechanism. *Planta* **167**, 44–53 (1986).
35. Burbulis, I.E. & Winkel-Shirley, B. Interactions among enzymes of the *Arabidopsis* flavonoid biosynthetic pathway. *Proc. Natl. Acad. Sci. USA* **96**, 12929–12934 (1999).
36. Panicot, M. *et al.* A polyamine metabolon involving aminopropyl transferase complexes in *Arabidopsis*. *Plant Cell* **14**, 2539–2551 (2002).
37. He, X.Z. & Dixon, R.A. Genetic manipulation of isoflavone 7-*O*-methyltransferase enhances biosynthesis of 4′-*O*-methylated isoflavonoid phytoalexins and disease resistance in alfalfa. *Plant Cell* **12**, 1689–1702 (2000).
38. Facchini, P.J., Bird, D.A., MacLeod, B.P., Park, S.-U. & Samanani, N. in *Integrative Phytochemistry: From Ethnobotany to Molecular Ecology* (ed. Romeo, J.T.) 143–180, (Pergamon, Amsterdam, 2003).
39. Schunmann, P.H.D. *et al.* A suite of novel promoters and terminators for plant biotechnology. *Funct. Plant Biol.* **30**, 443–452 (2003).
40. Schunmann, P.H.D., Surin, B. & Waterhouse, P.M. A suite of novel promoters and terminators for plant biotechnology—II. The pPLEX series for use in monocots. *Funct. Plant Biol.* **30**, 453–460 (2003).
41. Lazo, G.R., Stein, P.A. & Ludwig, R.A. A DNA transformation-competent *Arabidopsis* genomic library in *Agrobacterium*. *Biotechnology* **9**, 963–967 (1991).
42. Chitty, J.A., Allen, R.S., Fist, A.J. & Larkin, P.J. Genetic transformation in commercial Tasmanian cultivars of opium poppy, *Papaver somniferum*, and movement of transgenic pollen in the field. *Funct. Plant Biol.* **30**, 1045–1058 (2003).
43. Lenz, R. & Zenk, M.H. Purification and properties of codeinone reductase (NADPH) from *Papaver somniferum* cell cultures and differentiated plants. *Eur. J. Biochem.* **233**, 132–139 (1995).
44. Hirata, K., Poeaknapo, C., Schmidt, J. & Zenk, M.H. 1,2-dehydroreticuline synthase, the branch point enzyme opening the morphinan biosynthetic pathway. *Phytochemistry* **65**, 1039–1046 (2004).


T1# 1 2 3 4 5 6 7 8
Npt-II
actin
6OMT
actin
CNMT
actin
CYP80
actin
4'OMT
actin
7OMT
actin
BBE
actin
SAT
actin
COR2

Supplementary Figure 1 RT–PCR analysis of eight T_1 family segregants from hpCOR parent 220-2-2; these are the same individuals as shown in **Figure 6**. Top panel shows *NptII* transgene expression. The other panels show the corresponding RT–PCR results for the same individuals for other genes from the benzlisoquinoline pathway multiplexed in each case with actin. Gene abbreviations: (*R,S*)-norcocalurine-6-*O*-methyltransferase [6OMT]; (S)-coclaurine N-methyltrasnferase [CNMT]; (*S*)-*N*-methylcoclaurine-3'-hydroxylase [CYP80B1]; 3'-hydoxy-(*S*)-*N*-methylcoclaurine-4'-*O*-methyltransferase [4'OMT]; (*R,S*)-reticuline 7-*O*-methyltransferase [7OMT]; berberine bridge enzyme [BBE]; Salutaridinol 7-*O*-acetyltransferase [SAT]. The bottom panel shows the reduction in *Cor2* transcript for the same individuals.