Investigation of Taxol biosynthetic genes for the production of novel taxanes in heterologous plant systems

By Nathalie Narraidoo, BSc

Thesis submitted to The University of Nottingham for the degree of

Doctor of Philosophy, August 2011

Abstract

The diterpenoid paclitaxel (Taxol^{TM}) is one of the most effective anticancer drugs, used against a wide range of cancers. It is produced as a secondary metabolite in the vascular cambial region of the bark of *Taxus brevifolia* from which it was first extracted in 1971. Taxol also accumulates in low concentrations in several other *Taxus* species. As the demands for Taxol greatly exceeded its supply, alternative routes for producing the drug and its related taxanes were developed. Taxol is presently manufactured by semi-synthesis from its precursors baccatin III and 10-deacetylbaccatin III found in *Taxus* needles. The biosynthesis of Taxol mostly occurs via the 2-*C*-methyl D-erythritol 4-phospahte pathway and requires at least 19 enzymatic steps from the precursor geranylgeranyl diphosphate. This study set out to heterologously express the early genes of the Taxol biosynthetic pathway in *Nicotiana tabacum* for the subsequent redirection of this precursor for the synthesis of novel taxanes.

The first five genes of the Taxol biosynthetic pathway, namely taxadien-5-α-hydroxylase, taxadien-5α-acetyltransferase, taxoids 10β-, 13α- and 7β-hydroxylase, were isolated from *Taxus baccata* mRNA. Individual transgenic tobacco lines were generated expressing each of the first three enzymes of the biosynthetic pathway. These lines were crossed with each other in order to obtain all three transgenes expressed together in individual transgenic lines. Progenies from the crosses, expressing the first three transgenes were analysed, however, GS-MS analysis failed to detect the compound taxadiene-5α-ol and its acetylated compound taxadiene-5α-yl acetate. The expression of the Taxol biosynthetic genes in transgenic tobacco plants were accompanied by phenotypic effects, including dwarfism and low fertility of the transgenic plants. To circumvent these sterility issues which made crossing of the plants difficult, a construct

was prepared carrying the first two genes of the Taxol biosynthetic pathway, to be transformed in *yellow flesh* tomato mutant at a later stage.

The localisation of taxadiene synthase, 5α-hydroxylase and taxadien-5α acetyltransferase was investigated by making translational fusions to fluorescent protein tags. Confocal microscopy was used to detect the fluorescent proteins GFP, YFP and CFP in *Arabiodopsis thaliana* roots and tobacco leaf and root cells. Taxadiene synthase was found to be localised to the plastids, taxadien-5-α-hydroxylase spatially positioned on the plastid envelope and the endoplasmic reticulum membrane and taxadien-5α acetyltransferase was localised to the endoplasmic reticulum.

Acknowledgment

First and foremost, I would like to extend my gratitude to my supervisors, Dr. Rupert Fray and Professor Graham Seymour: for your guidance, friendship and unwavering support throughout my PhD. I feel very lucky to have had your mentoring and invaluable guidance during my studies and thank you for having given me the opportunity to work in your research groups.

Thanks Mum and Dad for teaching me by example to always aim high academically and that I can achieve anything if I put my mind to it. Thank you for having sacrificed so much of your time for us and for your financial support throughout my studies. Thank you both for my laptop which really helped throughout the writing up. Mum, thank you for your love and positivity. Dad, you set the bar high and made those kind of achievements real for me without creating pressure to follow your footsteps. I would never have got through this without you both!

A big thank you to my lovely sisters Khemila, Mee-neilly and Leesa for all your support and for being there for me. A special thought goes to my grandparents, uncle Jean Marc and mame Indira. Many thanks to all my friends and colleagues from Mill Lodge for your best wishes. Last but by no means least, thank you James for everything you have brought into my life and for your emotional support, especially through the final stages of my studies.

Table of Contents

Table of Figures

List of Tables

Abbreviations

ABA	Abscisic acid
ABC	ATP-Binding Cassette
At	*Arabidopsis thaliana*
bp	Base pair
BR	Brassinosteroids
CaMV	Cauliflower Mosaic Virus
cDNA	Complementary Deoxyribonucleic Acid
CFP	Cyan Fluorescent Protein
CPR	Cytochrome P450:NADPH Reductase
CYP	Cytochrome P450 monooxygenase
CIP	Calf Intestinal Phosphatase
DBAT	10-deacetyl baccatin III-10-O-acetyltransferase
DBTNBT	3′-*N*-debenzoyl-2′-deoxytaxol *N*-benzoyltransferase
DMAPP	Dimethylallyl diphosphate
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
DTT	Dithiothreithol
DW	Dry Weight
EDTA	Ethylenediaminetetracetic acid
EthBr	Ethidium Bromide
ER	Endoplasmic Reticulum
EST	Expressed Sequence Tagged
EtOH	Ethanol
FDS	Farnesyl Diphosphate Synthase
FPP	Farnesyl Diphosphate
FW	Fresh Weight
GA	Gibberellic Acid
GC	Gas Chromatography
gDNA	Genomic DNA
GDP	Geranyl 3 phosphate dehydrogenase
GFP	Green Fluorescent Protein
GGPPS	GGPP Synthase
G3P	Glyceraldehyde 3-phosphate
IAA	Indole-3-Acetic Acid
IspD	4-diphosphocytidyl-2C-methyl-D-erythritol synthase
IspF	2*C*-Methyl-D-erythritol-2,4-cyclodiphosphate Synthase
IPP	Isopentenyl Diphosphate
Kg	Kilogram
kb	Kilobase pair
LB	Luria Bertani
MEP	2-*C*-methyl-D-erythritol 4-phosphate
mg	Miligram
ml	Mililitre
MVA	Mevalonate pathway
mRNA	Messenger RNA
MS	Mass Spectrometry

NMR	Nuclear Magnetic Resonance
OCT	5(12)-oxo-3(11)-cyclotaxane
OD	Optical Density
PCR	Polymerase Chain Reaction
PPP	phytyl diphosphate
Psi	Pounds per square inch
RNA	Ribonucleic acid
RT	Reverse Transcription
Spp	Species
SDW	Sterile Distilled Water
TCPR	*Taxus* Cytochrome P450:NADPH reductase
TE	Tris-EDTA
TXS	Taxadiene synthase
T5AT	Taxa-4(20), 11(12)-dien-5α-ol-O-acetyltransferase
T2OH	Taxadiene 2α hydroxylase
T5OH	Cytochrome P450 taxadiene 5αhydroxylase
T5OHNSC	Cytochrome P450 taxadiene 5αhydroxylase without Stop codon
T7OH	Taxoid 7β hydroxylase
T10BOH	Taxoid Cytochrome P450 10-β-hydroxylase
T13OH	Taxoid Cytochrome P450 13-α-hydroxylase
TBT	Taxane 2α-*O*-benzoyl transferase
UV	Ultraviolet
Wt	Wild type
YFP	Yellow Fluorescent Protein
µl	Microlitre
µg	Microgramme
°C	Degree celsius
µl	Microlitre
♂	Symbol for male organism
♀	Symbol for female organism

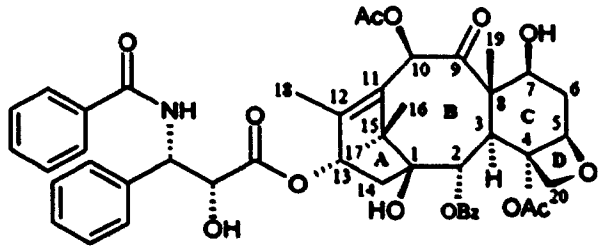
CHAPTER 1 : INTRODUCTION

1.1 *Taxus* and Taxoids

In late 1950s, an exploratory plant screening program was set up by the National Cancer Institution (NCI) in the United States to screen for anti-cancer substances from plants (Zubrod *et al.*, 1966). Many extracts were screened including a crude extract from the bark of *Taxus brevifolia,* which demonstrated cytotoxic activity against Melanoma B16 and other cancer cells. The active constituent was isolated and characterised as paclitaxel in 1971 (Wani *et al.,* 1971). Interests arose in paclitaxel when Susan Horwitz reported that this compound had a unique mode of action in promoting tubulin assembly and preventing depolymerisation, a process leading to cell death (Schiff *et al.*, 1979). In December 1992, paclitaxel was approved by the Food and Drug Administration (FDA) for the treatment of refractory ovarian and metastatic breast cancer (Suffness and Wall, 1995) and was subsequently marketed as Taxol™ (hereafter referred as Taxol) (Figure 1.1) by Bristol-Myers-Squibb. Taxol has become widely used for the treatment of a range of cancers including head and neck, breast, cervix, bladder and lung cancers as well as for AIDS-related Karposi's sarcoma.

The *Taxus* (yew) species are members of the *Taxaceae* family and are slow-growing shrubs widely distributed throughout Europe, the Far East, Northern and Central America. All *Taxus* species to date have been reported to produce Taxol (though levels may vary widely) (Itokawa, 2003), and the drug has been found to be most

abundant in the vascular cambial region (Strobel *et al.*, 1993) of the stem, needles and the root tissue of the yew.



Taxol

Figure 1.1 Structure and numbering of Taxol.

Taxol having a molecular formula $C_{47}H_{51}NO_{14}$ consists of an eight-membered B-ring fused to the A-ring and to the C-ring taxoid ring system to which an oxetane (D) ring is attached. The taxane core undergoes various esterifications to afford the baccatin core to which an *N*-benzoyl phenylisoserine side chain is attached to yield Taxol. OBz, benzoyl; AcO, acetyl.

The toxic nature of the yew trees has been known since the second century, when the king of Eburones poisoned himself with yew 'juice' (Reviewed in Wilson *et al.*, 2001). The accumulation of complex mixtures of taxane alkaloids in the needles and bark of the yews contribute to its poisonous nature. To date over 400 taxanes have been isolated from the different *Taxus* species (Baloglu and Kingston, 1999; Shi and Kiyota, 2005), among which Taxol and its precursor 10-deacteylbaccatin III exhibit strong antifeedant activity, explaining why the yew tree is not attacked by woodworm and its needles are resistant to insects (Daniewski *et al.*, 1998). Taxol, is believed to be produced as a secondary metabolite by the *Taxus* cells as part of plant defense responses to pathogenic attacks, hence possesses antibiotic (Elmer *et al.*,

1994) and antifungal activity (Young *et al.*, 1992) and is toxic to mammals (Odgen, 1988).

1.2 Biosynthesis of Taxol

The chemical structure of Taxol has led researchers to suggest that the pathway consists of 19 enzymatic steps. By making use of differential display of mRNA reverse-transcription PCR (DD-RT-PCR) method, homology based cloning approach and creating an EST library using methyl jasmonate induced *Taxus* cell culture, most of the genes and the steps of the Taxol biosynthetic pathway have been elucidated. Taxol biosynthesis (Figure 1.2) is believed to derive from the common isopentenyl diphosphate (IPP) and its isomer dimethylallyl diphosphate (DMAPP) produced via the mevalonate (MVA) or plastidial 2-*C*-methyl-D-erythritol 4-phosphate (MEP) pathway (Eisenreich *et al.* 1996; Cusidó *et al.*, 2002). Among the different enzymes which are responsible for the early steps of the Taxol biosynthetic pathway is the *Taxus* geranylgeranyl diphosphate synthase (GGPPS) (Hefner *et al.*, 1998). This enzyme catalyses the head-to-tail condensation of the five-carbon precursors, IPP and DMAPP to give the C10 geranyl diphosphate (GPP) followed by C15 farnesyl diphosphate and finally the C20 GGPP molecule.

The committed step of Taxol biosynthesis is represented by the cyclisation of the GGPP to taxa-4(5),11(12)-diene and its isomer taxa-4(20),11(12)-diene (Williams *et al.*, 2000a), a reaction catalysed by the plastidial taxadiene synthase (TXS) (Hezari *et al.*, 1995, 1997; Wildung and Croteau, 1996). The taxane core of these olefins undergo an allylic rearrangement of the 4(5) or 4(20) double bond to the 4(20)

position by a cytochrome P450 (CYP) mono-oxygenase taxadiene-5α-hydroxylase (T5OH) (Jennewein *et al.*, 2004) to yield taxa-4(20),11(12)-dien-5α-ol (Hefner *et al.*, 1996). The C-5-hydroxyl group of this alcohol is then acetylated by the taxadien-5α-acetyltransferase (T5AT) to form taxa-4(20),11(12)-dien-5α-yl-acetate (Walker *et al.*, 2000), representing the third step of the Taxol biosynthetic pathway. Two taxoids 10β-hydroxylases (Schoendorf *et al.*, 2001; Jennewein *et al.*, 2004b) have been isolated which both catalyse the hydroxylation of taxa-4(20),11(12)-dien-5α-yl-acetate at its C-10 position to afford taxa-4(20),11(12)-dien-5α-10β-ol-monoacetate. The oxygenated taxane core is further hydroxylated and decorated by a benzoyl group at the C2 position and an acetate group at the C10 position to afford the intermediate baccatin III, to which the C13 *N*-phenylisoserine side chain is appended in an additional five steps to yield Taxol (Figure 1.2).

Two CYP-dependent hydroxylases; taxoid 13α-hydroxylase (Jennewein *et al.*, 2001) and 14β-hydroxylase (Jennewein *et al.*, 2003) have been identified and are deemed to be involved in the bifurcation of the Taxol biosynthetic pathway, leading to the formation of dead-end metabolites in *Taxus* cell cultures (Ketchum *et al.*, 2007a). The enzyme that epoxidates the C4/C20 double bond to form the oxetane ring has not yet been identified, nor have those catalysing the C1 and C9 hydroxylation of the taxane core. Based on a survey of the relative abundances of the naturally occurring taxoids (Floss and Mocek, 1995) and the oxidation frequency of the taxoids found in cell cultures (Ketchum *et al.*, 2003, 2007a), the proposed sequence of oxidation is considered to progress from the C5 to C10, followed by C2 and C9 then C13, C7 and finally C1.

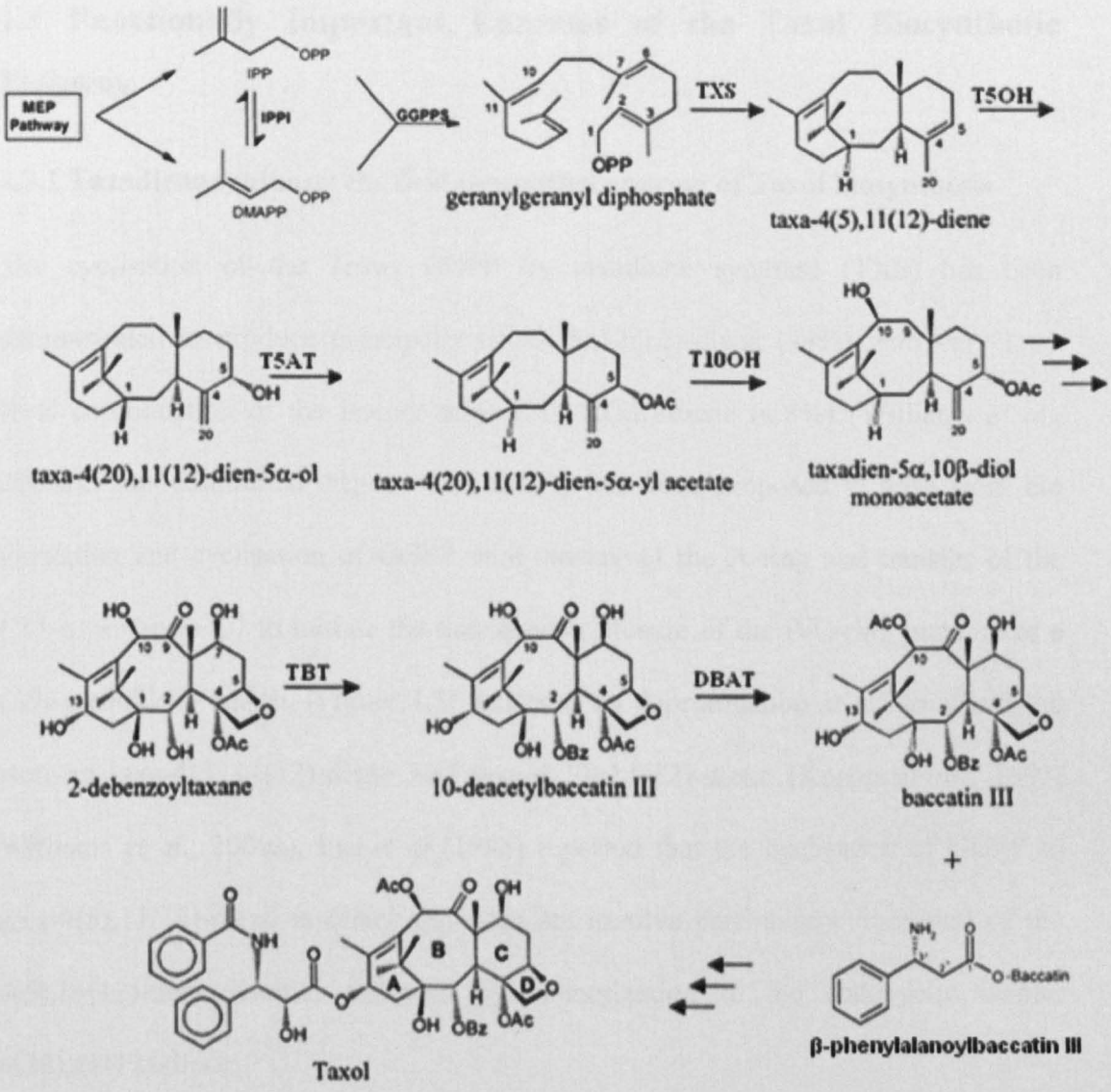


Figure 1.2 Outline of the proposed sequence of events in the Taxol biosynthetic pathway.

Taxol arises mainly via the plastidial 2-*C*-methyl-D-erythritol 4-phosphate (MEP) pathway through the supply of IPP and DMAPP. The cyclisation of GGPP to taxadiene by taxadiene synthase (TXS) is followed by the first oxygenation at the C5 by taxadien-5α-hydroxylase (T5OH). The taxane core undergoes further oxygenation (five alcohol functions, one carbonyl and one cyclic ether). Two subsequent acetylation (Ac), a benzoylation (OBz) by acyl and aroyl CoA-dependent transferases takes place followed by oxetane ring formation, and oxidation at the C9 leading to the late intermediate baccatin III, to which the C13 side chain is attached to afford Taxol. Multiple arrows represent undefined steps (Adapted from Guo *et al.*, 2006).

1.3 Functionally Important Enzymes of the Taxol Biosynthetic Pathway

1.3.1 Taxadiene synthase: the first committed enzyme of Taxol biosynthesis

The cyclisation of the *Taxus* GGPP by taxadiene synthase (TXS) has been demonstrated to produce principally taxa-4(5),11(12)-diene (94%) with very low-level coproduction of the isomer taxa-4(20),11(12)-diene (4.8%) (Williams *et al.*, 2000a). This committed step of the pathway has been proposed to arise from the ionization and cyclisation of GGPP with closure of the A-ring and transfer of the C11-α proton to C7 to initiate the transannular closure of the B/C-ring juncture to a (*1S*)-verticillene cation, (Figure 1.3) followed by deprotonation at C5 to afford the isomers taxa-4(5),11(12)-diene and taxa-4(20),11(12)-diene (Koepp *et al.*, 1995; Williams *et al.*, 2000a). Lin *et al* (1996) reported that the cyclisation of GGPP to taxa-4(5),11(12)-diene is direct and does not involve preliminary formation of the 4(5),11(12)-diene isomer followed by isomerization to the endocyclic isomer 4(20),11(12)-diene.

TXS was partially purified from the yew stem bark and adhering cambium and characterised as a monomeric protein of molecular weight 79 kDa. The enzyme was found to require Mg^{2+} as a cofactor, which is thought to bind the anionic diphosphate group of the substrate GGPP and to assist in the ionisation step of the reaction (Hezari *et al.* 1995). Eventually taxadiene synthase cDNA (Accession No.U48796) was cloned from a cDNA library of Pacific Yew (*T. brevifolia*) stem cells, through a homology-based PCR cloning strategy (Wildung and Croteau, 1996). TXS being the key enzyme in the Taxol biosynthetic pathway has been cloned from most of the

other *Taxus* species; *T. chinensis* (AY007207) (Wang *et al.*, 2002), *T. baccata* (AY424738) (Besumbes *et al.*, 2006) and *Taxus x media* (AY461450) (Kai *et al.*, 2005).

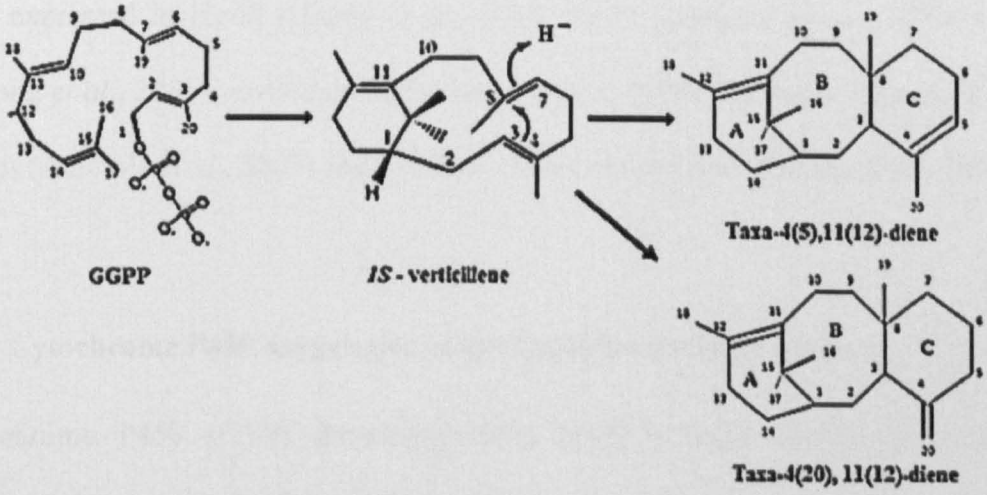


Figure 1.3 Outline of the cyclisation reaction of GGPP via an (1S)-verticillene to the formation of taxa-4(5),11(12)-diene and its isomer taxa-4(20),11(12)-diene (Adapted from Williams *et al.*, 2000a).

The very low levels of TXS activity in yew stem extracts and the trace amounts of this diterpene olefin intermediate in bark (5-10 µg/kg) (Koepp *et al.* 1995) suggested that this cyclisation step to establish the taxane skeleton could be a slow if not rate limiting step, in Taxol biosynthesis. However, a more detailed investigation into the enzyme activity of TXS over the course of Taxol production was performed using cell cultures of *T. canadensis* (Hezari *et al.*, 1997). The ability of the cell cultures to produce Taxol was not dependent on taxadiene synthase activity which led the authors to suggest that although the cyclisation step might be a slow step, it is not a

rate-limiting step in the pathway (Hezari *et al.*, 1997). Taxadiene is not sold commercially and is not readily available, requiring either 750 kg of dry bark from *T. brevifolia* for the isolation of only 1 mg of 85% pure taxadiene or a 25-step chemical synthesis (Koepp *et al.*, 1995). In order to achieve a high level of taxadiene as the first step towards the overexpression of Taxol in heterologous systems, TXS has been expressed in *E.coli* (Huang *et al.*, 1998, 2001; Ajikumar *et al.*, 2010), yeast (Dejong *et al.*, 2005), *Arabidopsis* (Besumbes *et al.*, 2004), the moss *Physcomitrella patens* (Anterola *et al.*, 2009) and a yellow fruited tomato line (Kovacs *et al.*, 2007).

1.3.2 Cytochrome P450 oxygenases of the Taxol biosynthetic pathway.

Cytochrome P450 (CYP) monooxygenases form a large family of enzymes distributed in many organisms from bacteria to higher eukaryotes, including mammals and higher plants (Nelson, 1999). CYP enzymes play an essential role in electron transfer of the respiratory chain and are also involved in the oxidation steps of plant secondary metabolism to form more complex structures of plant natural products, such as phenylpropanoids, flavonoids and terpenoids (Kim and Tsukaya, 2002; Morant *et al.*, 2003). CYPs catalyse the insertion of oxygen into an inactivated substrate (RH), while the second oxygen atom is reduced to water, as illustrated in the equation below:

$$RH + O_2 + NADPH, H^+ \rightarrow ROH + H_2O + NADP^+$$

The activities of most CYPs are usually associated to the outer face of the endoplasmic reticulum (ER) via a hydrophobic N-terminal transmembrane integration region required for anchoring (Werck-Reichhart and Feyereisen, 2000)

and receive their electrons from the flavoprotein NADPH Cytochrome P450 reductase (CPR) (Benveniste *et al.* 1977; Nelson, 1999). In eukaryotic organisms, their substrates are generally lipophilic, reflecting the integral nature of these enzymes. The CYPs are also characterised by the presence of a highly conserved proline, phenylalanine, glycine (PFG) heme-binding motif and a heme protein active centre which binds to the substrate protein through a highly conserved cysteine residue.

Approximately half of the proposed 19 distinct enzymatic steps of the Taxol biosynthetic pathway are considered to be catalysed by CYP mono-oxygenases (Croteau *et al.* 2006). By making use of differential display of mRNA-reverse transcription-polymerase chain reaction (DD-RT-PCR) method or homology based cloning and random sequencing, using transcripts isolated from methyl jasmonate-induced *T. cuspidata* cells, most of the CYPs involved in Taxol biosynthesis have been cloned.

1.3.2.1 Hydroxylation of the taxadiene nucleus by cytochrome P450 oxygenase taxadiene 5α-hydroxylase

The first and probably most important oxygenation step of the Taxol biosynthesis involves the hydroxylation at C5 of taxa-4(5),11(12)-diene or its isomer, with double bond rearrangement to yield taxa-4(20),11(12)-dien-5α-ol. Microsomal preparations from the woody stems of *T. brevifolia* and *T. cuspidata* were examined for their ability to transform partially soluble taxadiene to more polar products (Hefner *et al.*, 1996). The microsomal cuttings were found to catalyse the aerobic NADPH-

dependent conversion of taxadiene to an oxygenated substrate which structure was determined as taxa-4(20),11(12)-diene-5α-ol by NMR spectroscopy. The production of the taxadiene-5α-ol was also confirmed as the first oxygenated compound, by oxidising isomer taxa-4(20),11(12)-diene previously prepared by total synthesis (Rubenstein and Williams., 1995). The alcohol was demonstrated to be present in a 5-10 μg/kg range which was in the same range as taxadiene previously purified (Koepp *et al.*, 1995) suggesting that the oxygenation could be a less efficient step of the pathway (Wheeler *et al.*, 2001).

Due to its low level of induction, expression of this hydroxylase was previously undetected by the differential display-based cloning technique from methyl jasmonate-treated cell cultures (Wheeler *et al.*, 2001). Jennewein *et al.* (2004a) made use of a homology based PCR amplification with primers directed to regions of very high sequence conservation in CYP oxygenases of plant origin. Using this method, a clone of taxadiene-5α-hydroxylase (T5OH) (Accession no. AY289209) was identified containing an apparent 1509 bp open reading frame (ORF) encoding a protein of ~57 kDa and exhibited typical CYP characteristics including the oxygen binding domain and a N-terminal membrane anchor indicating that the enzyme is localised in the ER. To test the function of the encoded protein, the corresponding cDNA was expressed in *Saccharomyces cerevisiae* strain WAT11 (Pompon *et al.*, 1996) which harbours a galactose-inducible *Arabidopsis thaliana* NADPH-CYP reductase that is required for efficient redox coupling to the cytochrome. A cell-free protein expression system was used where the protein extracts from the transformed yeasts were fed exogenous taxane substrates including taxa-4(5),11(12)-diene and its

isomer taxa-4(20),11(12)-diene. Both isomers were found to have efficiently been converted to taxa-4(20),11(12)-dien-5α-ol.

To assess the substrate binding of T5OH, the cDNA was transferred to the baculovirus *Spodoptera fugiperda* (*Sf*9) expression system which coexpressed a *Taxus* CYP reductase (TCPR) (Jennewein *et al.*, 2005) for the efficient electron transfer to the cytochrome in microbial hosts. On feeding of the taxadiene isomers, the insect cell microsomes enriched in the recombinant T5OH bound to both isomers, with higher affinity to taxa-4(20),11(12)-diene. This correlated to what Lythgoe proposed in 1966 that taxa-4(20),11(12)-diene is the committed precursor of taxoid biosynthesis.

The oxygenation of taxadiene was shown to proceed by a CYP-mediated abstraction of hydrogen from the C20 methyl group of taxa-4(5),11(12)-diene or from C5 in the case of the 4(20) isomer, leading to the allylic radical intermediate **X** (Figure 1.4). This is followed by selective oxygen insertion from the 5α-face of **X** to accomplish the net oxidative rearrangement to yield to taxa-4(20),11(12)-dien-5α-ol (Jennewein *et al.*, 2004a).

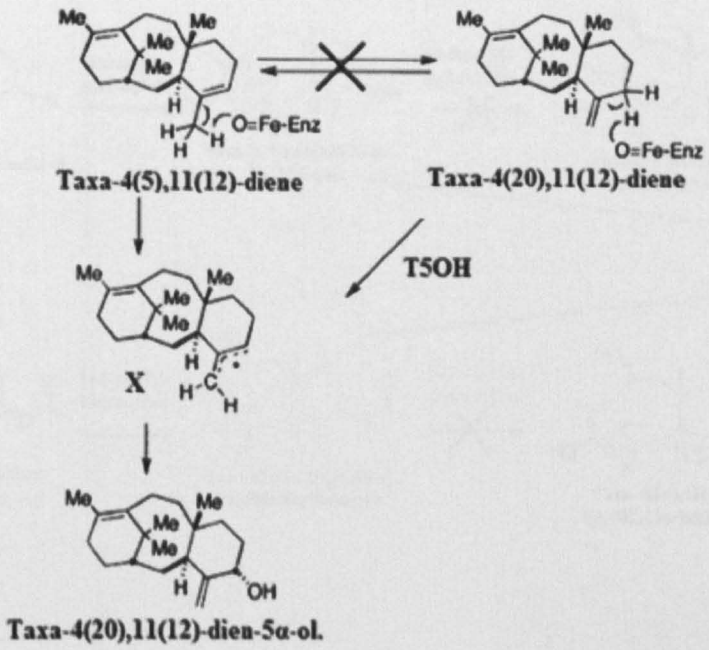


Figure 1.4 Proposed mechanism for CYP Taxadiene-5α-hydroxylase (Jennewein *et al.*, 2004a).

T5OH mediated oxygenation of taxa-4(5),11(12)-diene and taxa-4(20),11(12)-diene to taxa-4(20),11(12)-dien-5α-ol involved deprotonation at C20 in the 4(5) olefin or C5 in the 4(20) isomer leading to an intermediate **X** followed by oxygen insertion from the 5α-face to yield taxadien-5α-ol. Isomerisation of the 4(5) to 4(20) isomer was not observed.

1.3.2.2 Other CYPs involved in the Taxol biosynthetic pathway

Most of hydroxylases involved in the taxane decoration have been isolated, however, the sequence of oxygenations is still unknown. In an initial attempt to assess the product efficiencies and the order of hydroxylations of the taxadiene core, *Taxus* microsomes were fed with taxadien-5α-ol and taxadiene-yl acetate (third product of the Taxol pathway). Taxadien-5α-ol was observed to have been preferentially converted to a more polar compound identified as taxa-4(20),11(12)-dien-5α,13α-diol while taxadiene-yl acetate was converted to the compound taxa-4(20),11(12)-5α-10β-diol 5-acetate (Wheeler *et al.*, 2001) (Figure 1.5).

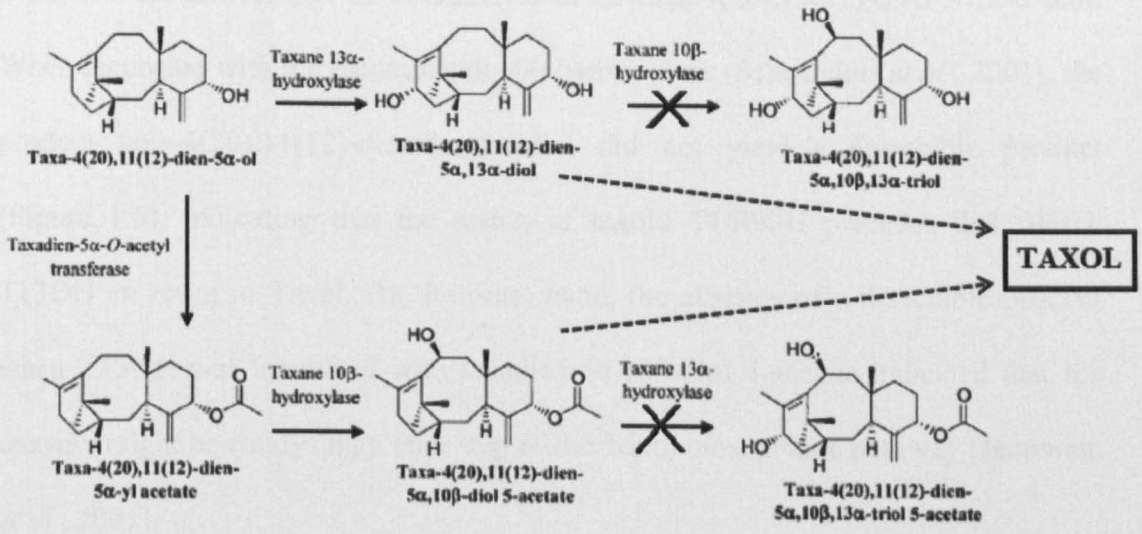


Figure 1.5 Outline of the synthesis of Taxol intermediates (Jennewein *et al* 2001).

The schematic diagram illustrating third step of the pathway, the acetylation of taxadien-5α ol to the ester by the enzyme taxadien-5α-O-acetyl transferase. The ester is further hydroxylated by 10β-hydroxylase while the taxadienol is preferentially hydroxylated by the 13α hydroxylase. Subsequent hydroxylations of the appropriate intermediates did not occur where the large X indicates that the illustrated reaction did not occur with the noted enzyme. The broken arrows signify undefined steps.

The clone responsible for the conversion of taxadien-5α-yl-acetate to taxa-4(20),11(12)-5α-acetoxy-10β-ol, was identified as a taxoid 10β-hydroxylase (T10BOH) (Accession no. AF318211) by Schoendorf and co-workers in 2001. Random sequencing using an EST cDNA library from the mRNA of *T. cuspidata* cells revealed the presence of another 10β-hydroxylase (Accession number AY563635) capable of mediating the same hydroxylation of taxadien-5α-yl acetate to taxa-4(20),11(12)-dien-5α-acetoxy-10β-ol (Jennewein *et al.*, 2004b).

In 2001, Jennewein and co-workers (2001) isolated the clone 13α-hydroxylase (T13OH) (GenBank accession no. AY056019) which produced a protein that

catalysed the conversion of taxadien-5α-ol to taxa-4(20),11(12)-dien-5α,13α-diol. When incubated with the recombinant 10β-hydroxylase (Schoendorf *et al.*, 2001), the product taxa-4(20),11(12)-dien-5α,13α-diol did not yield a detectable product (Figure 1.5), indicating that the action of taxoid T10BOH precedes that of the T13OH *en route* to Taxol. On the other hand, the absence of a detectable product when T13OH was incubated with taxadien-5α,10β-diol 5-acetate indicated that the enzyme might be catalysing a later step of the Taxol biosynthetic pathway (Jennwein *et al.*, 2001).

The uncertainties surrounding the central portion of the Taxol pathway and the lack of readily available test substrates have made the isolation of the genes involved difficult. Researchers have therefore made use of the dead-end metabolite of the yew heartwood Taxusin [taxa-4(20),11(12)-dien-5α,9α,10β,13α-tetraol tetra acetate] (Ho *et al.*, 1987), in an attempt to decipher the mid-pathway hydroxylases of the Taxol biosynthesis. The family of previously cloned CYP genes from the *T. cuspidata* cell culture (Schoendorf *et al.*, 2001) was expressed in WAT11 yeast cells which were incubated with the [^{3}H-acetyl]Taxusin. One clone was identified which catalysed the formation of 7β-hydroxyTaxusin and was characterised as a taxoid 7β-hydroxylase (Accession no AY307951) (Chau *et al.*, 2004). The hydroxylase was demonstrated to possess a higher binding affinity to polyoxygenated and acylated taxoids bearing a 5α-acetoxy group, hence placing this enzyme in the middle of the extended Taxol pathway.

A *T. cuspidata* EST library constructed by Jennewein and co-workers (2004b) allowed for the amplification of a range of cDNAs among which the taxoid 2α-

hydroxylase (GenBank accesssion no AY518383) (Chau and Croteau, 2004) was identified. This clone was found to preferentially catalyse the conversion of 7β-hydroxyTaxusin to a hexaol tetraacetate which led the authors to suggest that 2α-hydroxylase follows the 7β-hydroxylation *en route* to Taxol (Chau and Croteau, 2004).

1.3.2.3 Taxoid 14β-hydroxylase

Hydroxylations at position C14 is very rare in taxoids (Baloglu and Kingston, 1999), however 14β-hydroxy taxoids accumulate as major metabolites in *Taxus* cell cultures (Ketchum *et al.*, 2003). Jennewein *et al.* (2003) isolated the taxoid 14β-hydroxylase (GenBank accession no AY188177) which is responsible for the bifurcation of the Taxol pathway. This taxoid which hydroxylates the A-ring of the taxane core was found to possess several characteristics typical of CYP enzymes and resembles in sequence the taxoid 10β-hydroxylase which functionalises the B-ring of the core. Taxol is unsubstituted at C14, which led the authors to suggest that 14β-hydroxylase is not involved in the production of Taxol nor is the product taxadien-5α-acetoxy,10β,14β-triol (accumulated in *Taxus* cell culture (Ketchum et al., 2003)) an early metabolite of the drug.

1.3.2.4 C4, C20 epoxidation

The oxetane D-ring of the Taxol compound is considered as a very important feature for microtubule binding in cells, however little is known about the enzyme catalysing its formation or the substrate selectivity of the enzyme(s) which catalyse its synthesis. Mechanisms accounting for the formation of the oxetane ring has been proposed by several groups. The first step in formation of the ring is hypothesised as epoxide of the 4(20) double bond, however the configuration at C-5 and the oxetane ring preclude direct attack of the epoxide by the hydroxyl group.

Potier and co-workers in 1987 proposed that the oxetane formation involve an acetate rearrangement followed by the opening of the epoxide ring and a backside attack of a C5-acetate moiety onto C4 followed by a rearrangement to the oxetane ring with the migration of the 5α-acetoxy group to the C4α position of the taxane core (Guéritte-Voegelein *et al.*, 1987). However, it still remains unclear whether an inter or intramolecular reaction is responsible for the final formation of the oxetane ring (Rohr, 1997). Giner and Faraldos (2003) later proposed similar oxetane ring formation suggesting that intramolecular nucleophilic displacement gives rise to a dioxonium-ion intermediate (**1**) (Figure 1.6). Subsequent intramolecular displacement of the acetoxonium ion by the newly generated OH group would give rise to the oxetane ring.

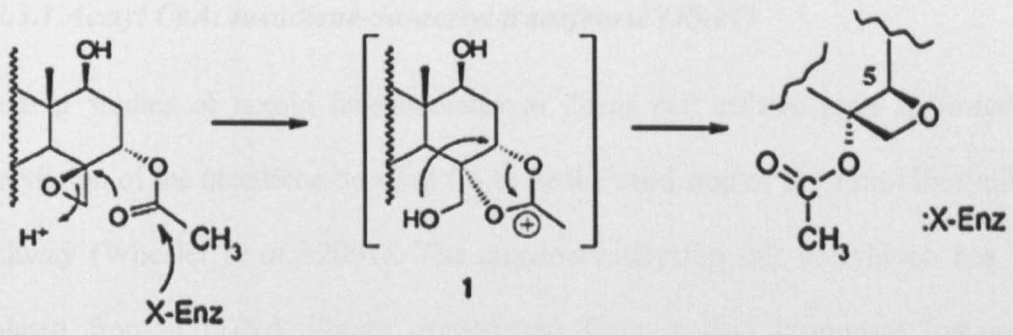


Figure 1.6 Proposed Biosynthesis Oxetane by Epoxy-Ester/Cyclic-Ether Rearrangement (Giner and Faraldos, 2003).

The reaction involves an acetate rearrangement with concomitant opening of the epoxide ring and a nucleophilic substitution reaction. The intramolecular exchange of the C5α-acetoxy group and the C4β-oxide function of an advanced taxoid is catalysed by a transferase mechanism to give a dioxonium-ion intermediate (**1**). **1** is proposed to undergo either an inter or intramolecuar reaction for the final formation of the oxetane ring. X-Enz is the undefined enzyme.

1.3.3 Acyl and aroyl transferases of the Taxane core

The CoA-dependent acyl and aroyl transferases are important members of the Taxol biosynthetic enzymes, constituting 5 of the 19 pathway steps. By making use of various cloning strategies, the genes encoding these transferases have been elucidated. The putative acyl group transfer motif (HXXXDG) generally found in a family of acyltransferases are; HAKSDG, HGVCDG, HSVSDG and HGICDG (Hu *et al.*,, 2002). This conserved motif has been suggested by St-Pierre *et al.*, (1998) to lie at or close to the active centre of the acyltransferases and to function in acyl group transfer from acetyl CoA to the substrate alcohol.

1.3.3.1 Acetyl CoA: taxadiene-5α-acetyl transferase (T5AT)

Feeding studies of taxoid intermediates to *Taxus* cell culture have indicated the acetylation of the taxadiene-5α-ol at C5 to be the third step of the Taxol biosynthetic pathway (Wheeler *et al.*, 2001). The enzyme catalysing this acetylation has been isolated from a cDNA library constructed from methyl jasmonate induced *T. cuspidata* cell culture (Walker *et al.*, 1999; 2000). This clone was functionally expressed in *E.coli* and the protein was found to convert the substrate taxa-4(20),11(12)-dien-5α-ol in the presence of acetyl-CoA, to its acylated derivative, taxa-4(20),11(12)-dien-5α-yl acetate, which correlated previous report (Wheeler *et al.*, 2001).

The isolated clone was characterised as an acetyl CoA: taxadiene-5α-acetyl transferase (T5AT) (Accession no. AF190130) which possesses characteristic features of other plant acyltransferases including the histidine motif (St. Pierre *et al.*, 1998) and a molecular weight of ~ 80kDa. Sequence comparison analysis have revealed high similarity between *Taxus* T5AT and other *Taxus* transferases as well as to other plant acetyltransferases (Hu et al., 2002). Walker *et al.* (2000) reported that T5AT sequence did not encode any N-terminal targeting information, however, the activity of this enzyme was found to reside in the microsomal extracts, possibly indicating that it is located to the endoplasmic reticulum (Walker *et al.*, 2000).

1.3.3.2. Benzoyl-CoA: taxane 2α-O benzoyltransferase (TBT) and 10-deacetyl baccatin III-10-O-acetyltransferase (DBAT)

A survey of the 350 naturally occurring taxoid intermediates characterised to date by Baloglu and Kingston (1999) suggests the second acylation of the taxane core to be the benzoylation of the C2α-hydroxyl group of an advanced taxane intermediate. Walker and Croteau (2000b) made use of a cell free assay of the expressed proteins previously isolated from *Taxus* cell cultures (Walker *et al.*,1999) to isolate the clone 2α-*O*-benzoyl transferase (TBT) (Accession no. AF297618), which mediates the benzoylation of the C2 position of the taxane core.

One of the late steps in the biosynthesis of Taxol is the acetylation at the C10 position of the late precursor 10-deacetylbaccatin III (10-DABIII) to afford the baccatin III. The acetyltransferase enzyme characterized as 10-deacetylbaccatin III-10-*O*-acetyl transferase (DBAT) (Accession no. AF193765), has been shown to catalyse the acetylation of 10-DAB III to baccatin III in the presence of acetyl CoA (Walker and Croteau, 2000a).

1.3.4 Enzymes of the C13 side chain biosynthesis

The taxane core and the C-13 phenylisoserine side chain of taxoids undergo both *O*- and *N*-acyl group transfer reactions, however, the precise order of some acylation and benzoylation reactions remain uncertain. Baccatin III was demonstrated as the immediate precursor of Taxol in labeling studies, to which the β-phenylisoserine side chain is attached to afford the end-product Taxol (Fleming *et al.*, 1994). Feeding studies of *Taxus* tissues with amino acid precursors and advanced taxoid

intermediates, metabolite profiling of *Taxus* cells in culture, and demonstration of relevant activities in cell free enzyme systems have demonstrated that the side chain synthesis requires a five-steps reaction, which is shown in Figure 1.7.

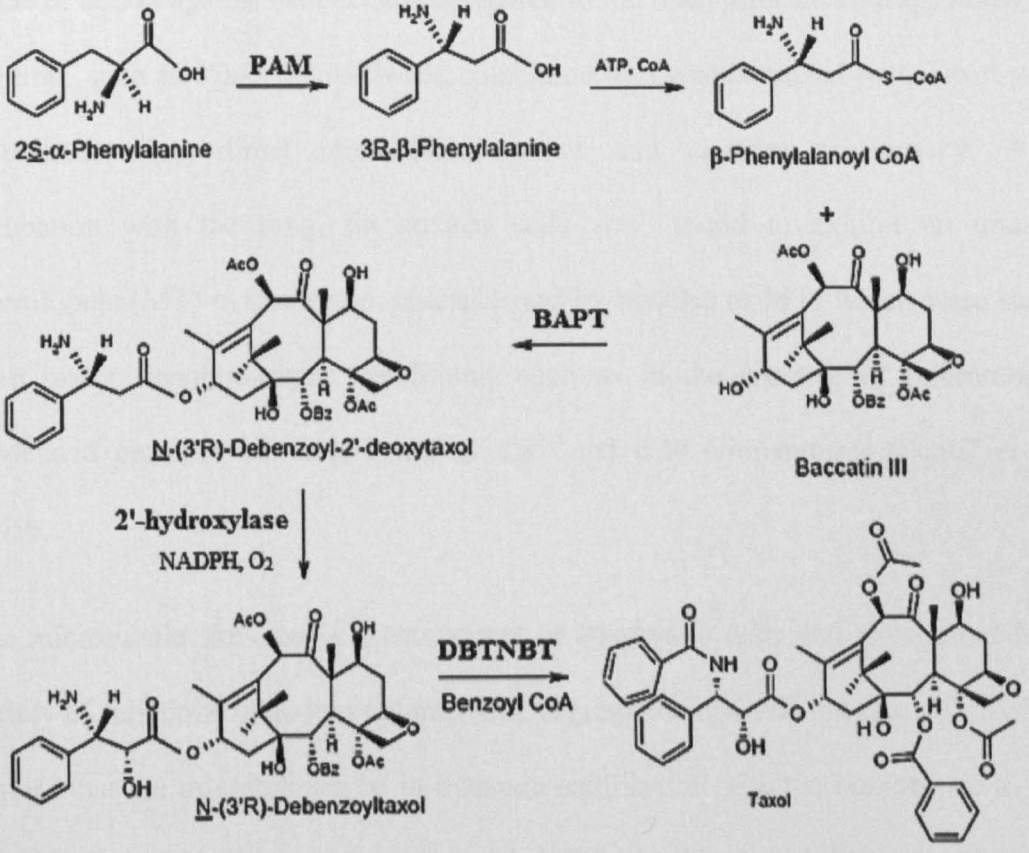


Figure 1.7 Outline of Taxol biosynthesis involving the five step assembly of the C13-side chain.

The enzymatic steps are as follows: the aminomutase PAM catalyses conversion of (*S*)-α-phenylalanine to (*R*)-β-phenylalanine. The amino acid is activated to the corresponding CoA ester and the aminoacyl group is then transferred to C13-hydroxyl of baccatin III. This reaction is catalysed by the enzyme BAPT to yield the product *N*-debenzoyl-2′-deoxytaxol. This intermediate then undergoes a CYP-mediated hydroxylation at the 2′-position in the side chain to give *N*-debenzoyltaxol, a reaction catalysed by a 2′-hydroxylase in the presence of NADPH and molecular oxygen. The last step is the *N*-benzamidation by the enzyme DBTNBT to afford Taxol in the presence of the benzoyl CoA. Ac is the acetyl group and Bz is the benzoyl group. PAM: phenylalanine aminomutase, BAPT: baccatin III 13-*O*-(3-amino-3-phenylpropanoyl) transferase, DBTNBT: 3′-*N*-debenzoyl-2′-deoxytaxol *N*-benzoyltransferase.

1.4 Mechanism of the Biological Activity of Taxol

Interest in Taxol as a drug candidate increased in 1979, when biologist Susan Horwitz and co-workers (Schiff *et al.*, 1979) reported the drug to employ a unique mode of action against cancer cells compared to the other anticancer drugs known at the time, such as *Vinka* alkaloids and colchinine. HeLa cells treated with Taxol were visualised using direct immunofluorescence and electron microscopy. After incubation with the drug, the treated cells were found to exhibit an unusual microtubule (MT) cytoskeleton, characterised by bundles of MTs which were stable even under depolymerising conditions, such as in the absence of microtubule-associated proteins, or in presence of Ca^{2+} and cold temperatures (Schiff *et al.*, 1979).

The microtubule skeleton is a component of eukaryotic cells and is required for a variety of functions including chromosome segregation and cell mitosis, all of which require that the microtubules be in dynamic equilibrium with the monomeric α- and β-tubulin. Taxol acts by promoting the MT assembly by shifting the dynamic equilibrium existing between tubulin dimers and microtubules towards microtubule assembly, hence promoting MT polymerisation. A schematic representation of the effect of Taxol on MT depolymerisation is shown in Figure 1.8.

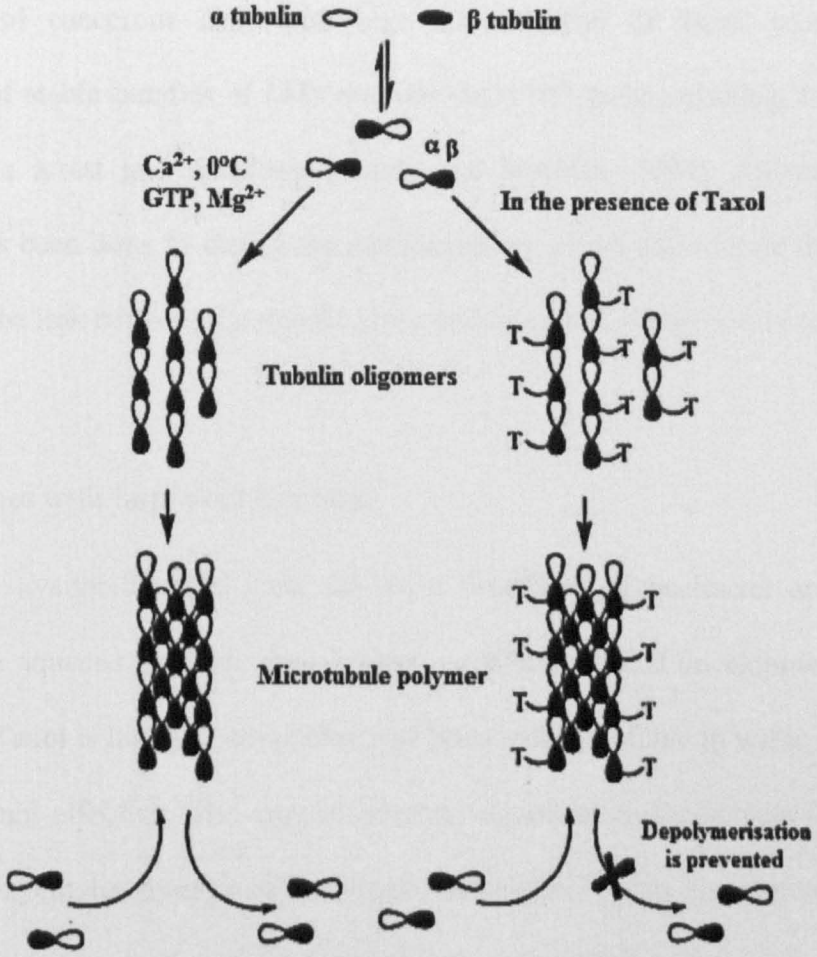


Figure 1.8 Microtubule polymerisation in the presence and absence of Taxol.

Under normal conditions, α- and β-tubulin bind reversibly to maintain a dynamic equilibrium between dimers and MTs. In the presence of Taxol, the drug molecules bind to the β-tubulin of the protofilaments, stabilising the MTs and inhibits the release of monomeric tubulin from the ends of the MT polymer causing a dynamic instability whereby the assembly and disassembly of the tubulin process is hampered.

Taxol was reported to induce different responses in a range of cancer cell lines, depending on the drug concentration. At low concentration, Taxol inhibits mitosis by altering MT dynamics initiating the formation of tubulin bundling, and micronuclei cells with abnormal DNA content, which lead to cell death due to gene dose problem and disproportionate distribution of chromosomes (Long and Fairchild, 1994).

Treatment of cancerous cells with high concentration of Taxol promotes the formation of stable bundles of MTs and increased MT polymerisation, resulting in G2/M phase arrest and apoptosis (Torres and Horwitz, 1998). Although, much research has been done to clarify the mechanism by which anti-mitotic drugs cause cell death, the link between the mitotic arrest and apoptosis is still poorly understood.

1.4.1 Taxanes with improved solubility

Apart from availability and cost, the main drawback of paclitaxel are its poor solubility in aqueous solution, drug delivery, cytotoxicity and development of drug resistance. Taxol is highly hydrophobic and being poorly soluble in water, its peroral delivery is not effective. The drug is currently dissolved in the solvent Cremophor EL (a polyoxyethylated castor oil derivative) and ethanol for its intravenous delivery. The main limitation to the clinical use of Cremophor EL is its toxicity, which has been associated with multiple side-effects (Expósito *et al.*, 2009). Recently, Abraxane®, a novel albumin-bound formulation of paclitaxel has been developed which has demonstrated better solubility and increased anti-tumour activity in breast cancer (Miele *et al.*, 2009).

The structure of Taxol has been modified in many ways to create novel drugs with improved bioavailability and solubility. One of the many Taxol derivatives widely used to date is Taxotere® (docetaxel) (Figure 1.9). Docetaxel was first semi-synthesised by Potier and co-workers in 1991 from the Taxol precursor 10-DAB III and carries a butoxyl group at the 3′-N-benzoyl position of the side chain attached to the C13 of the taxane core (Guéritte-Voegelein *et al.*, 1991; Guénard et al., 1993).

This important taxane has demonstrated improved solubility and anti-mitotic activity against a P388 leukemic cell line, compared to Taxol (Bissery *et al.*, 1991).

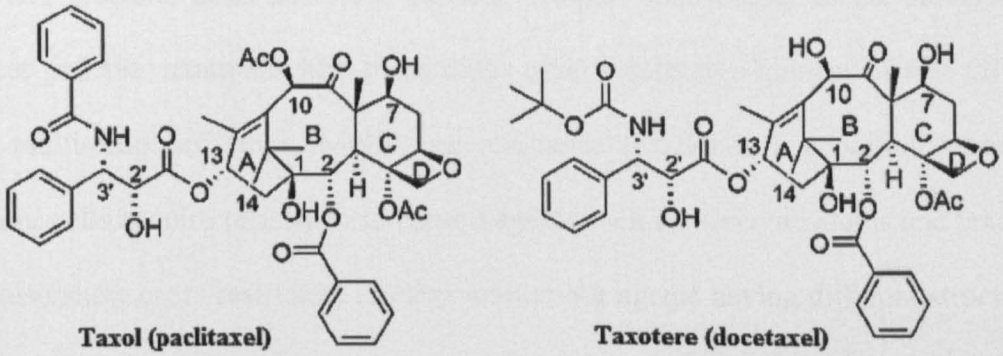


Figure 1.9 Structure of Taxol and docetaxel.

The structural differences between Taxol and its semi synthetic analogue Docetaxel are the presence of the *tert*-butoxycarbonyl group instead of a benzoyl group on the nitrogen atom at C-3′ of the C-13 side chain and the hydroxyl group instead of an acetate at the C-10 position of the B-ring.

A wide range of taxane-derived prodrugs has been synthesised by esterification of the 2′ of the *N*-phenylisoserine side chain and/or C7 of the taxane core which has improved the water solubility of Taxol. Replacing the hydroxyl group of C-2′ of the side chain of Taxol by a pyridine moiety (Nicolaou *et al.*, 1994 a) or addition of a highly ionisable phosphate moiety at the C7 position of paclitaxel-2′-ethylcarbonate (Ueda *et al.*, 1995) have respectively yielded in suitable cellular environments, products with high solubility and similar antitumour properties to Taxol.

1.4.2 Taxane resistance in cancerous cells

Taxol and Docetaxel are currently considered to be among the most important anticancer drugs used for the treatment of a range of cancers including breast, lung, ovarian, prostate, head and neck cancers. Despite contributing to the survival of cancer patients, treatment with these drugs often results in a number of side effects and multi-drug resistance. Multi-drug resistance (MDR) is a condition whereby tumour cells acquire resistance to natural agents such as *Vinca* alkaloids and taxanes but also show cross-resistance to other antitumour agents having different structures and modes of actions. Taxol resistance can also be the result of mutations occurring at various points in the β-tubulin, causing an alteration in the drug-binding pocket, therefore preventing the drug from effectively attaching itself to the MT (Reviewed in Galletti *et al.*, 2007).

The MDR phenotype in cancer cells is characterised by the overexpression of the ATP-binding cassette (ABC) superfamily of transport proteins which pump mainly hydrophobic cytotoxins out of the cells, keeping intracellular drug concentration below a cell-killing threshold. These transmembrane protein pumps include the "P-glycoprotein (P-gp) encoded by the *mdr1* gene, multidrug resistance protein 1 (MRP1), encoded by the *mrp1* gene and the breast cancer resistance protein (BCRP), encoded by the *mxr* gene" (Galletti *et al.*, 2007). Taxol and Taxotere are substrates for most of these ABC pumps, therefore several strategies for blocking their removal by modifying the structure of these drugs have been investigated.

1.4.3 Natural and synthetic taxanes overcoming transport-based resistance

Appropriate modifications at the C2, C7, C10, and C3′ positions of paclitaxel, docetaxel as well as other taxane derivatives have led to the synthesis of "second generation" taxane anticancer agents, prodrugs and noncytotoxic taxane-based reversal agents (tRAs) which act as modulators of the ABC pumps in cancer cells. The search of "second generation" taxoids possessing activity against drug-resistant cancer cells has led to the finding of the natural taxanes for instance Taxusin, Taxuspine C, taxinine NN-1, taxinine NN-11 and its 9,10-isomer (Figure 1.10) isolated from the Japanese yew, *T. cuspidata* (Kobayashi *et al.*, 1998; Bai *et al.*, 2005). Some of these natural taxanes possess a cinnamoyl group at different positions on the taxane core and have demonstrated the ability of modulating the P-gp pumps in human ovarian carcinoma MDR 2780AD cell line resistant to the *Vinca* alkaloid drug Vincristine. Among these taxoids isolated with antitumour activities, a few non-taxol type taxoids were also reported. These compounds when administered in combination with the antimitotic agent Vincristine were able to modulate the P-gp pump, inhibiting the active efflux of the antitumour agents from the tumour cells, hence overcoming MDR. (Kobayashi *et al.*, 1998; Kobayashi & Shigemori, 2002).

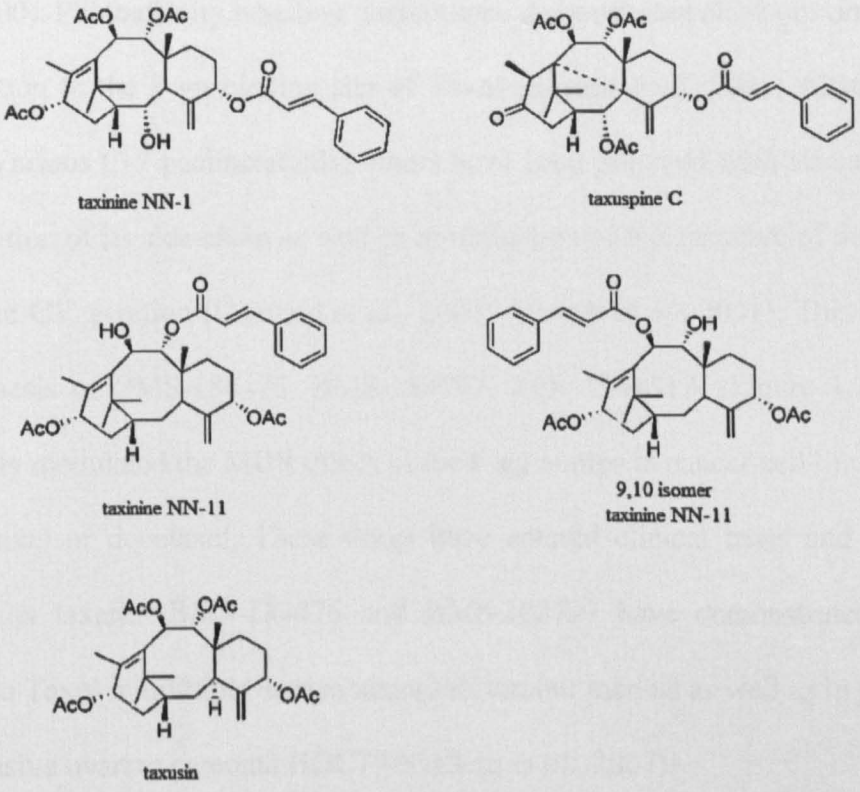


Figure 1.10 Natural taxanes with ability to modulate MDR (Galletti *et al.*, 2007).

A wide range of non-cytotoxic taxane-based reversal agents (tRAs) (Figure 1.11) has been synthesised from natural taxane 14β-hydroxy-10-deacetylbaccatin III (14-OH-DAB) isolated from the needles of *T. wallichiana* and from the Taxol precursors 10-DAB III and baccatin III (Brooks *et al.,* 2003; Ojima *et al.*, 1997; 1998). When orally administered along with paclitaxel, the most potent tRAs; tRA 96023 (Minderman *et al.*, 2004), tRA 98006 (Brooks *et al.*, 2003) and Ortataxel/ IDN 5109 (Nicoletti *et al.*, 2000) modulated a broad-spectrum of the ABC pumps. In particular, the P-gp of the intestinal tract were found to be modulated, allowing a greater accumulation of paclitaxel in the drug-resistant cancer cells. IDN5109, (Figure 1.11) currently marketed as Ortataxel has been found to be highly orally active with excellent bioavailability, thus providing the first orally active taxoid anticancer agent (Nicoletti

et al., 2000). Photoaffinity labelling studies have demonstrated close proximity of the C-7 position to the P-gp binding site of Taxol (Snyder *et al*, 2001; Altstadt *et al.*, 2001)). Various C-7 paclitaxel silyl ethers have been prepared with variation at the C-3′ position of its side chain as well as modification of the structure of docetaxel at its C4 and C3′ position (Guénard *et al.*, 2000; Altstadt *et al.*, 2001). This has led to the synthesis of BMS-184476, BMS-188797, RPR 109881A (Figure 1.11) which effectively modulated the MDR effect of the P-gp pumps in cancer cell lines resistant to paclitaxel or docetaxel. These drugs have entered clinical trials and the novel intravenous taxanes BMS-184476 and BMS-188797 have demonstrated superior activity to Taxol in multiple human xenograft tumour models as well as in paclitaxel-unresponsive ovarian carcoma HOC79 (Galletti *et al.*, 2007).

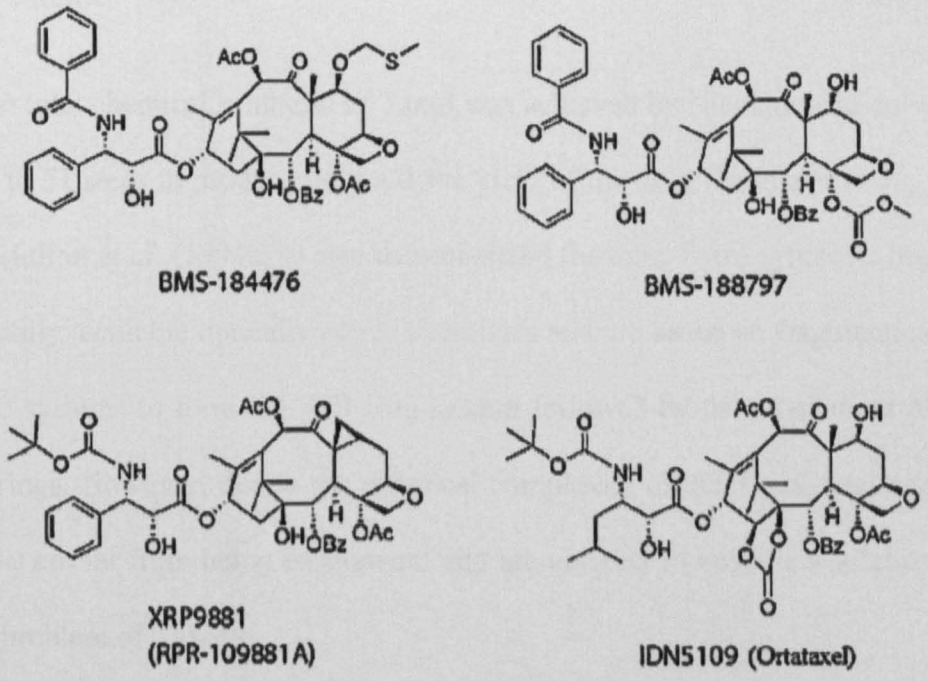


Figure 1.11 Taxanes with enhanced MDR modulation effects which have entered clinical trials (Galletti *et al.*, 2007).

1.5 Taxol Production and its Biotechnological Applications

The current market for Taxol relies on *Taxus* cell suspension cultures and on semi-synthesis using the Taxol precursors baccatin III and 10-deacetyl baccatin III (10-DAB III) which account for up to 0.2% dry weight of needles of the European yew *T. baccata.* The treatment of cancer per patient requires at least 3g of Taxol, and only 1 kilogram of Taxol is obtained from 1000 yew trees, which make the yew trees a limited supply of the drug (Rohr, 1997; Suffness and Wall, 1995). Furthermore, the very slow growth rate of the yew trees and the presence of analogues with very similar structure to that of Taxol make the purification of the drug difficult. Much research has been done in search of alternative methods to increase the production of Taxol, in order to meet its increasing demand for the treatment of cancer and for pre-clinical studies.

The first total chemical synthesis of Taxol was achieved by Nicolaou and co-workers in a 35 to 51 steps to produce only a 0.4% yield of the drug (Nicolaou *et al.,* 1994b, 1995). Holton *et al.* (1994a, b) also demonstrated the total Taxol synthesis beginning with readily available optically active chemicals and are based on fragmentation of a tricyclic systems to form the A/B ring system followed by the attachment of the C and D rings. However, due to the chemical complexity of the Taxol structure, these synthesis are far from being economical and are unlikely to become a solution to the supply problem of Taxol.

1.5.1 The semi-synthesis of Taxol from its immediate precursors

A semi-synthetic approach has been devised as an alternative source of Taxol or Docetaxel, which Bristol-Myers Squibb uses as a nine-step process by making use of the Taxol intermediate 10-deacetyl baccatin III (10-DAB III) (Reviewed in Malik *et al.*, 2011). 10-DAB III is readily obtained at level 1g/kg from the renewable needles of the European yew *T. baccata* (Chauvière *et al.*, 1981). Numerous reaction schemes have been developed, however production of the compounds in multi-kilograms is via the Holton β-lactam coupling method (Holton *et al.*, 1994a) which has provided a standard protocol and has been used and modified by other research groups for the development of short and practical semi-synthetic routes. The starting material, 10-DAB III from the needles of *Taxus baccata* is selectively acetylated and protected with triethylsilane (TES). The hydroxyl group at C13 of this C7-protected baccatin III derivative reacts with the selected β-lactam to give either Taxol or Docetaxel.

1.5.2 Biotechnological approaches and Tissue culture

The leading producer of Taxol, Bristol-Myers Squibb has a farm with 30 billion *Taxus* trees to supply the bark and needles necessary for the extraction of Taxol and taxane intermediates (Reviewed in Malik *et al.*, 2011). The extraction of taxanes from the bark and needles of the yew trees require expensive technology and complex system of purification techniques. One preferred alternative to these problems has been the use of *Taxus* cell suspension culture, which is an environmentally balanced source of the drug and its analogues. Cell suspension

culture also offers the possibility of growing the tissue material under controlled conditions and independently of its original and remote location (Jennewein and Croteau., 2001). Currently, Python Biotech is the largest producer of paclitaxel via plant tissue culture, employing a large-scale fermentor with a capacity of up to 75,000 L (Expósito *et al.*, 2009).

1.5.2.1 The use of elicitors on the production of taxanes

The production of many secondary metabolites including paclitaxel as part of plant defense in response to biotic and abiotic elicitors has prompted much research in their application to *Taxus* cell suspension culture for the increased production of the drug. Other strategies to optimise the production of taxanes in *Taxus* cell culture have been the use of different culture batch systems, selection of high-producing cell lines, effect of heat-shock (Zhang and Fevereino, 2007) and cell wall digesting enzymes (Roberts *et al.*, 2003).

Methyl jasmonate (MJ), a natural plant product produced as part of plant defense, is the most widely used elicitor for the enhancement of taxanes in *Taxus* cell suspension cultures. Its first use as an elicitor was demonstrated by Yukimune and co-workers in 1996 in *Taxus* cell suspension cultures, which resulted in a 5.1- and 20.2-fold increase in the yield of Taxol (in *T. media*) and baccatin III (in *T. baccata*) at levels ~110 mg/L and 53.02 mg/L by week two respectively. Since, this enhancement agent has been widely used for taxane optimisation in *Taxus* cell cultures (Ketchum *et al.*, 1999a), which has led to considerable progress in the

understanding of the pathway, enzymology, and molecular genetics of Taxol biosynthesis (Ketchum *et al.*, 2007a).

1.5.2.2 Synergistic effect of elicitors, additives or inducing factors

Different elicitors have enhanced the production of taxanes to various extents, however, when applied synergistically, they have demonstrated a more marked effect on the yield of these secondary metabolites. The combination effects of the elicitors such as methyl jasmonate, chitosan and silver ion on paclitaxel level in *T. chinensis* have been demonstrated to result into a 40-fold increase, yielding 25.4 mg/L of the drug in the elicited *T. chinensis* cell culture (Zhang *et al.*, 2000).

US. Patent. No. 5019504 (Christen *et al.*, 1991) described the first use of fungal elicitors for the enhanced production of Taxol by *T. brevifolia* cell suspension cultures. The release of ethylene from these fungal-elicited cultures was reported to interfere with the production of Taxol, hence addition of ethylene inhibitors such as silver ion was reported by Zhang and Wu (2003) to result in an increase in Taxol yield in fungal-elicited *Taxus* cell cultures.

Cusidó *et al.* (2002) demonstrated that the addition of *N*-benzoylglycine and mevalonate to a two-stage *Taxus media* cell culture elicited with methyl-jasmonate led to a 8.3 fold (21.12 mg/L) and 4 fold (56.03 mg/L) increase of paclitaxel and baccatin III respectively. The *N*-benzoylglycine was proposed to possibly undergo hydrolysis to release benzoyl moiety and a glycine residue which might be used for the synthesis of phenylalanine and benzoic acid required for Taxol synthesis. The

enhanced production of Taxol upon addition of mevalonate led the authors to speculate that Taxol can be synthesised via the mevalonate pathway under particular conditions (Cusidó *et al.*, 2002).

Recently, US. Patent. No. 0086397 (Bringi *et al.*, 2011) describes the ability of all known *Taxus* species to produce Taxol and its derivatives in very high yield in cell culture. Modifications of the culture conditions using a two-batch cell culture process with enhancement agents jasmonates, ethylene inhibitor silver nitrate, auxin and inhibitors of the phenylpropanoid pathway have enhanced the yield of a range of taxanes. The authors reported the highest yield obtained to date in cell culture of Taxol, baccatin III and other taxanes in the amount of 350 mg/L/day, 250 mg/L/day and 400 mg/L/day respectively.

Despite the effectiveness of *Taxus* cell culture for the production of Taxol and related taxanes, this method has proven challenging due to the low and often unstable pattern of production of the drug as well as high production cost and selectivity over unwanted by-products.

1.5.2.4 Agrobacterium transformation of Taxus tissue

The successful *Agrobacterium*-mediated transformation of *Taxus* cell lines (Han *et al.*, 1994) and the maintenance of the transgenic *Taxus* cell suspension culture on a long-term basis (Ketchum *et al.*, 2007b) have been documented. With most of the steps of the Taxol pathway elucidated and the enzymes involved cloned, this transformation methodology could allow the overexpression of the genes responsible for the rate-limiting steps, or block particular steps in the pathway by antisense, co-

suppression or RNA interference (RNAi) methods. For instance, the taxoid 14β-hydroxylase involved in the bifurcation of the Taxol biosynthetic pathway to a large family of 14β-hydroxy taxoid side-products (Menhard *et al.*, 1998; Ketchum et al., 2003) could be silenced. This method could also allow the redirection of these side routes for the production of novel taxanes with enhanced solubility and MDR modulating ability. Other important taxanes, for instance 13-acetyl-9-dihydrobaccatin III is present abundantly (2.5g/Kg of dry plant) in *T. Canadensis* (Canadian yew) (Ketchum *et al.*, 1999b) and has been used as starting material for the synthesis of 7-protected baccatin III (Zamir *et al.*, 2003). Genetically manipulating the genes towards the synthesis of this dihydrobaccatin III in *Taxus* cells might increase its production yield and it can subsequently be used for the synthesis of baccatin III and hence Taxol.

Methyl jasmonate elicitation affects the Taxol biosynthetic pathway at the mRNA expression level, which in turn has been reported to affect respective taxane accumulation (Nims *et al.*, 2006). By using *Agrobacterium*-mediated transformation of *Taxus* cells, the enzymes expressed at low levels in methyl jasmonate elicited cell cultures, can be overexpressed to increase Taxol production yields.

1.5.3 Production of taxanes by endophytic fungi

Taxol, an antifungal agent has been reported as a metabolite in the endophytic fungus *Taxomyces andreanae* associated with *T. brevifolia* (Stierle *et al.* 1993). This demonstrated that organisms other than *Taxus* species can produce Taxol and related taxanes, which has led to the hypothesis of horizontal gene transfer between *Taxus*

species and their corresponding endophytic organisms. Although the yield of Taxol from the fungus *T. andreanae* was as low as 24-50 ng/l, the short generation time and high growth rate of the fungi make it a potentially more reliable, cheap and renewable way of optimising the production of Taxol and related taxanes. Other reports of Taxol-producing endophytes include *Pestalotiopsis microspora* NE-32, isolated from the inner bark of a Himalayan yew tree *T. wallichiana* which yielded Taxol at titers 55±10 μg/L (Strobel *et al.*, 1996), and the *Tubercularia* fungi (strain TF5) isolated from *Taxus mairei* (Chinese southern yew) (Wang *et al.*, 2000).

The evergreens most resembling *Taxus* and growing in similar damp habitats but not producing Taxol, are bald cypress - *Taxodium distichum*. 9 of the 16 strains of the endophytic *Pestalotiopsis microspora* living in the bark, phloem and xylem of *T. distichum* were found to produce Taxol in a range of 24 – 1487 ng/L (Li *et al.*, 1996). Other reports of Taxol production in endophytic fungus not associated with the *Taxus* trees have been documented in *Capsicum annuum* plant pathogen *Colletotrichum capsici,* at level 687 μg/L, which was 13,740-fold more that reported in *T. andreanae* (Kumaran *et al.*, 2011). Pandi and co-workers reported the accumulation of 245 μg/L Taxol from the *Lasiodiplodia theobromae* (Pandi *et al.*, 2011), an endophytic fungus of the medicinal plant *Morinda citrifolia* which has in the past been reported to be used for treatment against cancers. These two recently isolated plant pathogenic fungi as well as the other Taxol producing fungi could serve as potential targets for genetic engineering to enhance the production of Taxol.

1.5.4 Biosynthesis of taxanes in transgenic *Saccharomyces cerevisiae*

Saccharomyces cerevisiae has been successfully employed as host for the engineered, multi-step production of terpenoids. In order to reconstitute the Taxol biosynthetic pathway in yeast, the cDNAs encoding a truncated version of GGPPS and TXS, the taxoid hydroxylases 5α, 10β and 13α and the acyltransferases T5AT, TBT and DBAT, were transferred into yeast under the control of GAL promoters (Dejong *et al.*, 2005). The protein levels for most hydroxylases was found to have increased steadily in the transformed yeast microsomal extracts, except for the T5OH. The engineered yeast cells successfully produced taxadiene via the MVA pathway, indicating that this pathway could be utilised for the synthesis of taxanes, as previously suggested by Cusidó *et al.* (2002). Taxadien-5α-ol was produced in very small amounts (~25 μg/L), but neither taxadien-5α-yl acetate nor taxadien-5α-acetoxy-10β-ol was detected during this study. These results postulated that the first two enzymes GGPPS and TXS cooperated well with each other yielding 0.5 mg taxadiene per g DW cells, however, a pathway restriction was encountered at the 5α-hydroxylation step. The authors hypothesised that this might have been due to the limited expression of this CYP or to the low endogenous levels of NADPH-CPR in yeast, indicating the need for a *Taxus* CPR to maximize redox efficiency with the CYP hydroxylases in microbes.

In 2008, Engels and co-workers developed a system to optimise the supply of GGPP in yeast as substrate for the production of taxadiene. The authors made use of a truncated version of 3-hydroxyl-3-methylglutaryl-CoA reductase (tHMG-CoA reductase), the rate-controlling enzyme of the MVA pathway, which converts 3-hydroxyl-3-methylglutaryl-CoA to mevalonate in yeast. This has prevented the

steroid-based negative feedback of the MVA pathway, resulting in an increased supply of IPP. A mutant allele *upc*2.1 of the key regulator of steroid uptake, was used to allow steroid uptake under aerobic conditions, hence diverting the IPP precursor flux away from steroid production and towards taxane biosynthesis. The authors also made use of the GGPPS from the thermophilic archeon *Sulfolobus acidocaldarius* which uses DMAPP as precursor for GGPP synthesis, reducing the competition for FPP by squalene synthase in yeast for steroid synthesis. A truncated form of TXS gene was expressed in combination with the *S. acidocaldarius* GGPPS, the tHMG-CoA reductase and the *upc2* transcription factor gene all under the control of the constitutive phosphate glycerol kinase (PGK) promoter. This system resulted in a 40-fold increase in taxadiene levels yielding the olefin at level 8.7±0.85 mg/l. (Engels *et al*, 2008).

1.5.5 Taxol precursor production in *Escherichia coli*

The faster growth of *Escherichia coli* (*E. coli*) compared to yeast makes it a better system to large-scale microbial fermentation while using inexpensive renewable feed stocks (carbon and nitrogen sources). The upstream MEP pathway of Taxol biosynthesis is native to *E. coli*, where IPP and DMAPP couple to yield GGPP which is naturally produced at low level due to the limited pools of its precursors.

Initial work in engineering the Taxol pathway in *E. coli* by Huang and co-workers (1998; 2001) has been limited to the production of taxadiene due to the absence of CYP-bound enzymes in the microbe. During the first attempt in 1998, the major limitation encountered was the formation of insoluble taxadiene synthase in inclusion

bodies. In 2001, the authors advanced this research by using a more soluble form of the TXS whereby the signal peptide was deleted in order to reduce membrane association. Engineering *E. coli* with a single vector carrying the enzymes 1-deoxyxylulose-5-phosphate synthase (DXS), IPP-DMAPP isomerase (IDI), GGPPS and the truncated form of TXS driven by a strong T7 promoter, led to an increase in production of soluble form of taxadiene with yields approaching 1.3 mg/L (Huang *et al.*, 2001).

Ajikumar *et al.* (2010) developed a combinational approach where the rate limiting enzymes of the MEP pathway (*dxs, idi, ispD* and *ispF*) were carefully optimised to find the most favourable order of genetic modification in transgenic *E. coli* to maximize the production of GGPP. A truncated form of *T. brevifolia* GGPPS and TXS were then genetically transplanted in this *E.coli* strain. A total of 32 constructs were prepared with varying promoter strengths (Trc, T5 and T7), plasmid copy number (5, 10 and 20) and genotype. Three *E.coli* strains were found to accumulate the highest yield of taxadiene, where the maximum accumulation of 1020 ± 80 mg/L taxadiene was achieved in strain 26. This strain harboured an additional copy of the *dxs-id-ispD-ispF* operon of the MEP pathway chromosomally integrated under the control of a Trc promoter. The TG (TXS-GGPPS) genes were driven by a strong T7 promoter while keeping a relatively low plasmid copy number. The authors also reported using fed-batch bioreactors with controlled glycerol feeding and a layer of organic solvent to prevent air stripping of the volatile taxadiene.

The limitation of bacterial platforms for the expression of T5OH has been due to the absence of the endoplasmic reticulum and CYP-reductases (CPRs) and has also been

hampered at the translational incompatibility level of the membrane signal modules of microsomal CYP enzymes. To address the issue of translational incompatibility, the membrane anchor of T5OH was altered by truncation at 8, 24 and 42 amino acids from the N-terminus of the transmembrane (TM) region of the gene to reduce membrane association. To provide for the efficient electron transfer to the cytochrome in host microbes, the three TM engineered T5OH were linked to the supporting 74 amino acids truncated *Taxus* CYP reductase (tTCPR) by a flexible peptide linker. The constructs At8T5αOH-tTCPR, At24T5αOH-tTCPR and At42T5αOH-tTCPR were respectively transformed into individual taxadiene producing strain 26. Among the chimeric constructs generated, At24T5αOH-tTCPR demonstrated the highest efficiency in converting 98% of the accumulated taxadiene to its oxygenated product taxadiene-5α-ol and its byproduct 5(12)-Oxa-3(11)-cyclotaxane (OCT), previously identified by Rontein and co-workers in 2008. The alcohol and the cyclotaxane both accumulated to ~24 mg/L, however no taxadiene accumulation was detected, which demonstrated its near total conversion to the oxygenated and cyclotaxane products. The yield of these oxygenated compounds were further optimised by small-scale fermentations in bioreactors which enhanced their production to 58±3 mg/L, demonstrating a 24,000 fold increase of taxadiene-5α-ol compared to that previously recorded in *S. cerevisiae* (Dejong *et al.*, 2005).

1.5.6 Taxane production in *Arabidopsis* and *Nicotiana sylvestris*

The use of heterologous plant systems as a biological production approach for the synthesis of Taxol precursors has led to the use of the model plant *Arabidopsis* for the production of taxadiene (Besumbes *et al.*, 2006). TXS expression was induced using a glucocorticoid-mediated system, which consistently resulted in an efficient recruitment of GGPP for the production of 600 ng taxadiene per g DW of mature leaves (Besumbes *et al.*, 2006). Transgenic *Arabidopsis* expressing TXS under the control of a constitutive promoter, was reported to produce taxadiene at level 20 ng/g DW, which was 30 fold lower than when TXS was under the control of an inducible promoter. Plants constitutively producing taxadiene could have accumulated this olefin at high levels which might have been toxic to the transgenic lines, therefore only those accumulating the lowest level of taxadiene might have survived. Furthermore, plants constitutively expressing the TXS showed reduced hypocotyl length and retardation of growth and flowering, and a pale, bleached phenotype due to the reduced level of chlorophyll and carotenoids. The production of TXS enzyme in the plant cells might have interfered with the endogenous balance of isoprenoid production for the synthesis of the GA hormones, photosynthetic pigments such as carotenoids and affected the production of the isoprenoid phytol side chain of chlorophyll (Besumbes *et al.*, 2004). A similar effect was observed in transgenic tomato plants constitutively overexpressing the phytoene synthase (*Psy1*) gene (Fray *et al.*, 1995). These authors suggested the reduced stature in the high Psy1 expressers to be due to the reduced GGPP pool available for GA and phytol biosynthetic pathway.

In an attempt to redirect the GGPP used for cembratrien-diol (CBR-diol) biosynthesis in the trichomes of *Nicotiana sylvestris* for the heterologous production of taxadiene, a TXS gene was engineered in the tobacco line where the CBR synthase (CBTS) was silenced using an intron hairpin RNAi construct (ihpCBTS) (Tissier *et al.*, 2006). This resulted into ~20 μg taxadiene/g of fresh leaves where the compound was reported to have successfully accumulated in the trichomes of this *N. sylvestris* line. Rontein and co-workers (2007) made use of this specific tobacco line to cross with transgenic *N. sylvestris* expressing T5OH under the control of a trichome specific promoter. When the TXS and T5OH were co-expressed, taxa-4(5),11(12)-diene was found to be absent from the GC-MS profile but taxadiene-5α-ol was not detected either. Instead a new isomer 5(12)-oxa-3(11)-cyclotaxane (OCT) was detected and quantified as 2μg/g FW leaf.

OCT was found to be a very hydrophobic taxoid with the same molecular weight as that of taxa-4(20),11(12)-dien-5α-ol, consistent with a single oxidation reaction on the taxadiene core. The three-dimensional model of OCT (Figure 1.12) displayed a globular shape featuring a C3/C-11 linkage and a C-5/C-12 ether bridge spanning the entire molecule. With this novel conformation of the cyclotaxane, the oxygen atom of the ether function is partially hidden inside the structure, explaining why OCT is very stable (at 200°C), insensitive to air oxidation and non-volatile in standard conditions.

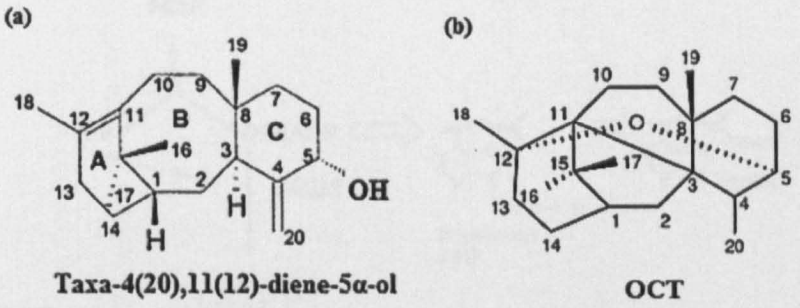


Figure 1.12 Structure of taxa-4(20),11(12)-diene-5α-ol and OCT.

The hydroxyl group at C5 in taxa-4(20),11(12)-diene-5α-ol is susceptible to modifications causing the compound to be unstable. OCT displays a globular shape with no free hydroxyl group, but an ether linkage between C5/C12 whereby the oxygen molecule is hidden inside the structure, hence protected from air oxidation and other modifications (Rontein *et al.*, 2008).

1.5.7 Redirection of GGPP for the production of taxadiene in tomato plant

With the large number of pathways requiring GGPP as precursor (Figure 1.13), this C20 compound is tightly regulated in plant tissues. In tomato fruit, however, most of the GGPP is devoted to the biosynthesis of the accessory pigments carotenoids, which accounts for up to 2% of the fruit dry weight (Kovacs *et al.*, 2007). The first step in the synthesis of carotenoids is the condensation of two molecules of the precursor GGPP to give the C_{40} phytoene, a reaction catalysed by the enzyme phytoene synthase (PSY-1) (Fraser *et al.*, 1999). The *yellow flesh* (*r*) tomato mutant lacks a functional fruit *Psy-1* gene and does not accumulate lycopene during ripening (Fray and Grierson, 1993). This mutation in *Psy-1* means that the carotenoid synthesis in ripening fruits does not occur and the conversion of GGPP into phytoene does not take place, resulting in a potentially large pool of unutilised GGPP.

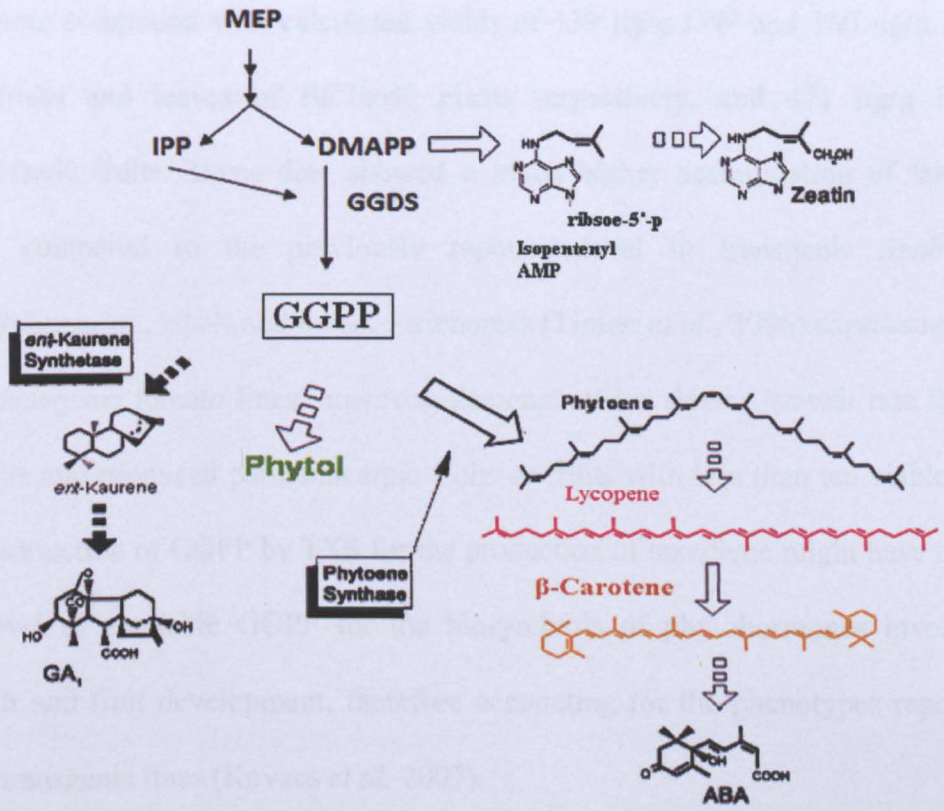


Figure 1.13 Schematic representation of the carotenoid pathway in tomato fruits (Fray *et al.*, 1995).

The pathways leading to production of carotenoids, phytol, gibberellins and other hormones all use GGPP as precursors except for the synthesis of zeatin, which uses the GGPP precursor DMAPP for its synthesis.

By engineering *r* tomato mutants with the *T. baccata* TXS gene, Kovacs *et al.* (2007) reported the redirection of the large available pool of GGPP for the production of taxadiene. The authors made use of two specific plant binary vectors pBCTaxK and pRD12TaxK, where TXS gene was placed under the control of a CaMV 35S promoter and polygalacturonase (PG) promoter respectively. Hexane extracts from the leaf and fruit materials of the pBCTaxK and fruit extracts from pRB12TaxK expressing plants gave GC peaks characteristic of taxadiene. The crude taxadiene extracts were purified by flash column chromatography to give a >95% pure

taxadiene compound with calculated yields of 339 μg/g DW and 160 μg/g DW in ripe fruits and leaves of BCTaxK plants respectively, and 471 μg/g DW in RD12TaxK fruits. These data showed a much higher accumulation of taxadiene when compared to the previously reported level in transgenic *Arabidopsis* (Besumbes *et al.*, 2004) and tobacco trichomes (Tissier *et al.*, 2006) expressing TXS. The transgenic tomato lines, however, demonstrated a slower growth rate than the controls and produced parthenocarpic fruits or fruits with less than ten viable seeds. The redirection of GGPP by TXS for the production of taxadiene might have reduced the level of available GGPP for the biosynthesis of phytohormones involved in growth and fruit development, therefore accounting for the phenotypes reported in these transgenic lines (Kovacs *et al,* 2007).

1.6 Manipulation of other Isoprenoid and Drug Pathways in Transgenic Plants

Humans have always used compounds derived from plants for treating diseases and to date many prescription drugs are still derived directly or indirectly from plants (Gómez-Galera *et al.*, 2007). Besides Taxol, other plant-derived compounds, for instance, the terpenoid alkaloids vinblastine and vincristine (from the Madagascar periwinkle *Catharanthus roseus*) and the antimalarial drug artemisinin (from *Artemisia annua*) have received FDA approval for clinical use. The low accumulation level of important drugs in plants however, makes the drugs very costly and unaffordable to developing countries, which are the most affected by diseases such as malaria. Biotechnological tools such as genetic engineering have

therefore been used to manipulate the key endogenous biosynthetic genes in plants catalysing the slow steps of the particular biosynthetic pathways in order to select, increase and improve the production of important metabolites.

Morphine alkaloids normally accumulate in the range 1.5-2.7% DW in the *Papaver* genus (poppies). The cultivation of *P. somniferum*, opium poppy remains critical for the production and supply of morphine, codeine, thebaine, oripavine and various semi-synthetic analgesics. Larkin *et al* (2007) reported an elevated level of these alkaloids in transgenic *P. somniferum* where the codeinone reductase (*PsCor1.1*) catalysing the last step conversion of morphine intermediates codeinone to codeine was overexpressed. The total alkaloids in the transgenic plants were increased following expression of the gene, with an elevated level of morphine (25%), codeine (58%) and thebaine (75%) in the seed capsules.

Amorpha-4,11-diene synthase (ADS) catalysing the committed step of artemisinin biosynthesis in *Artemisia annua* converting FPP into amorpha-4,11-diene has been the target for artemisinin metabolic engineering. Wu *et al.* (2006) engineered tobacco plants with a plastid-targeted ADS and avian farnesyl diphosphate synthase for the redirection of the intermediates IPP and DMAPP from the MEP pathway. Some of the transgenic tobacco lines accumulated more than 25 μg/g FW of amorpha-4,11-diene. This led the authors to suggest that the same methodology can be applied for the overexpression of these enzymes in the plastids of *A. annua*.

Plants produce a vast array of secondary metabolites that have nutritional and pharmaceutical properties in humans. One such group of metabolites are the carotenoids, whereby the intermediates lycopene and β-carotene have been reported

to reduce the onset of certain cancers and age-related macular degeneration (Reviewed in Farré *et al.*, 2010). Fruits and vegetables are good sources of certain carotenoids, however staple food of plant origin, for instance rice generally lack these compounds, including provitamin A. The deficiency of provitamin A causes various symptoms such as blindness, and has been estimated to affect a large number of children, especially in developing countries. Rice endosperm contains neither β-carotene nor C_{40} carotenoids. To improve the provitamin-A content of rice, Ye *et al* (2000) genetically engineered rice endosperm with carotenogenic genes phytoene synthase (*psy*) and lycopene β-cyclase (*lycb*) from *Narcissus pseudonarcissus* (daffodil) and phytoene desaturase (*crtI*) from *Erwinia uredovora,* which redirected the GGPP in the endosperm for the accumulation of 1.6 μg/g carotenoid in the endosperm of the 'Golden Rice'. In 2005, Paine *et al.* further increased the provitamin A level in another transgenic rice line (Golden Rice 2) by using a maize gene encoding *psy* in place of the daffodil. The authors reported the total carotenoid level in the homozygous transgenic rice to have increased by 23-fold to a maximum of 37μg/g compared to the original Golden Rice, with a preferential accumulation of 31μg/g β-carotene.

Tomato fruits and their processed products are one of the main sources of important carotenoids, hence much effort has been made towards genetically modifying food crops, especially tomato, to increase the level of carotenoids. Another successful example of plastidial isoprenoid metabolic engineering has been demonstrated by Fraser *et al.* (2002) who targeted a non-homologous bacterial phytoene synthase (*crtB*) from *Erwinia uredovora* to the chromoplasts of tomato plants. The use of the bacterial *crtB* with <35% homology to the endogenous tomato *Psy* under the control

of a fruit-specific polygalacturonase (PG) promoter circumvented gene-silencing of the *Psy1* gene and avoided the phenotypic effects previously observed while using the constitutive 35S CaMV promoter (Fray *et al.*, 1995). Transgenic tomato fruits expressing the *crtB* exhibited a 1.6 - 2.0-fold increase in fruit carotenoids, with increased accumulation of phytoene, lycopene, β-carotene and lutein. In 2004, D'Ambrosio *et al.* described the near conversion of all the lycopene to β-carotene in transgenic tomatoes, by the constitutive expression of the tomato lycopene β-cyclase cDNA which led to an increase in the total carotenoid accumulation in the fruits with a maximum of 202.6 μg/g FW β-carotene. Another successful method used for the manipulation of the carotenoid pathway has been demonstrated by silencing the tomato endogenous De-etiolated gene regulating photomorphogenesis (Davuluri *et al.*, 2005). As a result, the biosynthetic flux in the pathway was redirected towards carotenoid synthesis, which led to a 8-fold (130 μg/g DW) increase in β-carotene accumulation in transgenic tomato chromoplasts. Besides the carotenoid content, flavonoid levels were also found to have increased by 1.9- to 3.5-fold when compared to the wild type tomato fruits. Wurbs *et al.* (2007) reported the plastid expression of a bacterial *lycb* gene by genetic engineering of the chloroplast genome. The authors made use of a plastid *aptI* promoter and the lycopene β-cyclase gene from the carotenoid-producing eubacterium *E. herbicola* (*crtY*), which were transferred to tomato leaves by particle bombardment. The authors reported a four-fold enhanced provitamin A increase in the transplastomic lines with an accumulation of 286 μg/g DW β-carotene in tomato fruits.

Engineering plants with additional copies of carotenoid biosynthesis genes by nuclear transformation has frequently been associated with cosuppression (Fray and

Grierson, 1993; Fray *et al.*, 1995). Targeting the biochemical pathways of important isoprenoids and other pharmacologically important drugs to the plastid genome of several crops has demonstrated many advantages to nuclear transformation, one such includes the absence of epigenetic effects, therefore eliminating the hurdles of gene silencing. The high transformation efficiency and greatly enhanced level of provitamin A and amorphadiene accumulation in transplastomic tomatoes (Wurbs *et al.*, 2007) and in tobacco plastids (Wu *et al.*, 2006) respectively, highlight the advantage of generating plastid transformation systems. Plastid engineering also allows the generation of large quantities of therapeutic proteins that could be harmful if accumulated in the cytoplasm of the transgenic plant cells. Furthermore, plastid engineering permits transgene containment by maternal inheritance, thereby reducing the potential of toxicity exerted by transgenic pollen and the possibility of outcrossing transgenes to related plants (Reviewed in Daniell *et al.*, 2005).

1.7 Project Aims and Objectives

This study sought to investigate the heterologous expression of the genes involved in the early steps of the Taxol biosynthetic pathway engineered in *Nicotiana tabaccum*. This plant system was preferentially used for its ease of *Agrobacterium*-mediated transformation and for its rapid growth.

The aims of the first half of the study (chapter three) were to isolate and clone the genes encoding enzymes involved in the early steps of Taxol biosynthesis. Upon cloning and expression of such cDNAs in *Arabidopsis* and tobacco plants, the localisation of the enzymes were to be experimentally demonstrated using fluorescent tagged proteins (chapter four).

The aims of the second half of the study (chapter five) were to transfer the cloned genes into tobacco plants and to produce transgenic lines expressing each transgene separately. Each transgenic tobacco line was to be crossed with each other in order to "stack" the first three genes of the Taxol biosynthetic pathway into individual lines. Plants expressing the first two genes were to be analysed for the presence of the first oxygenated taxane of the Taxol biosynthetic pathway and the resulting lines were to be crossed with transgenic lines expressing the gene encoding the third enzyme of the pathway. Plant materials from these transgenic lines expressing the three enzymes would be subjected to GC-MS analysis for the detection of the downstream taxanes.

CHAPTER 2 : MATERIALS AND METHODS

2.1 Chemical Material

2.1.1 Laboratory reagents

All laboratory reagents were purchased from Sigma-Aldrich Co. Ltd. (Poole, Dorset, UK), VWR International Ltd. (Hunter Boulevard, Leicestershire, UK) or Fisher Scientific UK Ltd. (Loughborough, Leicestershire., UK), unless otherwise stated. Microbiological media were obtained from Oxoid Ltd. (Basingtoke, Hampshire, UK). Unless otherwise stated the reagents were of analytical grade or higher. ^{32}P radiolabelled nucleotides were obtained from Amersham International UK and Perkin-Elmer.

All media and solutions referred to in the text as sterile were autoclaved at 120 °C for 30 min at a pressure of 15 psi prior to use, or (when autoclaving was not suitable) were filter sterilised using either a 45 μm or a 20 μm syringe filter (Anachem Ltd., Bedforshire, UK)

2.1.2 Enzymes

Restriction and modifying enzymes were purchased from Fermentas (York, UK), Promega (Delta House, Southampton, UK) or New England Biolabs (NEB) UK Ltd. (Hitchin, Hertfordshire, UK).

2.1.3 Oligonucleotides

DNA oligonucleotides were purchased from Eurofins MWG Operon (Ebersberg, Germany) (see Appendix I).

2.1.4 pDestination vectors

The binary vectors with fluorescent protein tags (Cerulean, EGFP and Venus) used for localisation studies were kindly offered by Dr. Silin Zhong (Zhong *et al.*, 2008), in other cases the vectors were purchased.

2.2 Bacterial Strains

The bacterial host for all routine cloning work was *E. coli* DH5α (genotype: *lac*ZΔM15, Δ(*lac*ZYA-*arg*F), U169, *rec*A1, *end*A1, *hsd*R17 (r_k^-,m_k^+), *pho*A, *sup*E44λ-*thi*-1, *gyr*A96, *rel*A1). The *Agrobacterium tumefaciens* strain used in this study was GV3101 (pMP90).

2.3 Plant Materials and Growth Conditions

The seeds of *Arabidopsis thaliana* seeds of the control ecotype (Columbia, Col-0) were obtained from the Nottingham Arabidopsis Stock Centre (NASC). The tobacco wild type seeds (*Nicotiana tabacum var. Petit Havana)* of the control cultivar were kindly provided by Dr. Rupert Fray (Division of Plant and Crop Sciences, The University of Nottingham, UK). Unless specific conditions are stated, the *Arabidopsis* and *Nicotiana* plants were grown in Levington M3 compost consisting of 20 mg/L Intercept (Scotts, Ipswich, UK) insecticide. The glasshouse for maintain tobacco plants were kept at minimum 16 h day regime at about 22-26°C. The

greenhouse for growing Arabidopsis was maintained with 22 h photoperiod at about 21-23°C.

2.4 Screening Seeds for Control and Transgenic Lines

The collected seeds were first surface sterilised by soaking in 50% (v/v) bleach with a drop of Triton X-100 (BDH Chemical Ltd., UK) for 8 min and centrifuged in a bench-top microcentrifuge (Sigma 1-15) at 13000 g for 1 min. The bleach was discarded and the seeds were washed 5 times with SDW containing 0.5 % Triton by centrifugation each time at 13000 g for 1 min. This was followed by the addition of 70% ethanol to the seeds and this solution containing the seeds was poured immediately onto a sterile filter paper. The seeds were allowed to dry in a laminar flow hood. After drying, the seeds were either sprinkled (to screen for transgenics) or displaced horizontally (for seeds from transgenic lines) onto autoclaved MS agar media (1% (w/v) agar), containing appropriate antibiotics, in petri dishes or sterile 100mm x 100mm square plates for vertical growth respectively and sealed with parafilm. The petri dishes were kept at 4°C without light for 2 days before transferring to the tissue culture room (18 h photoperiod, 23-25°C) to allow germination. For transgenic seeds carrying a construct with a fluorescent tag, the seedlings were allowed to grow on sterile square petri dishes for 1-2 weeks before confocal microscopy. To obtain seedlings for further transgenic analysis, the *Arabidopsis* seedlings that developed dark green true leaves and an extending root system were transferred to compost in 9 cm pots and grown in the growth room under general conditions as previously described (section 2.3).

2.5 Plant Transformations

2.5.1 Arabidopsis transformation

The floral dip method (Clough and Bent, 1998) was used to transform *Arabidopsis* using *Agrobacterium tumefaciens* strains GV3101 (pMP90). 100 mL of the *Agrobacterium* overnight culture grown to OD_{600} 1.0 was spun down (3, 000 x g, RT) and resuspended in 200-300 ml 5% (w/v) sucrose solution. Silwet-L77 (Lehle Seeds, TX, USA) was added to the *Agrobacterium* solution to a final concentration of 0.05% (v/v) prior to dipping. The aerial parts of flowering *Arabidopsis* plants grown in 9 cm pots were dipped into the *Agrobacterium* solution for 15 s with gentle agitation. The dipped plants were covered with folded plastic sleeves for 24 h to maintain humidity. After transformation, plants were watered and grown normally for 3-4 weeks; the seeds of the transformed plants were then harvested and screened on MS media with appropriate antibiotics according to the selection marker present in the transgene construct. The collected seeds were screened under methods described in section 2.4 and grown on appropriate antibiotics to select for transformants.

2.5.2 Tobacco transformation

The surface sterilised tobacco seeds were first germinated and grown in sterilised plastic pots with MSR3 medium (Appendix II) for 4-5 weeks. The first healthy fully expanded leaves were chosen and cut into 1 cm^2 using surgical blades, forceps and placed in deep petri dish in sterile liquid MS, under sterile conditions.

Prior to day of tobacco transformation, a 10 ml overnight liquid culture of *Agrobacterium* (GV3101: pMP90) carrying the appropriate construct was spun down (3, 000 x g, 10 min, RT) and suspended in 100 ml of liquid MS (Appendix II) to about an OD_{600} of 0.2 in a sterile flow hood. The leaf pieces were transferred to the *Agrobacterium* culture and incubated with gentle swirling for 10 minutes. The explants were blotted dry on autoclaved filter paper (Whatman, UK) to remove the excess *Agrobacterim* then placed on M1 medium containing the appropriate antibiotics with the abaxial surface of the explants in contact with the callus inducing medium (MS with 0.5 mg/ml IAA and 0.75 mg/ml Zeatin). The explants were cultures at 25°C in the dark for 24-48 hours before being transferred to M13 medium (MS with 0.5 mg/ml IAA, 0.75 mg/ml Zeatin, 400 mg/L augmentin and appropriate antibiotic depending on the selection marker of the transgene construct) and sub-cultured to freshly prepared medium every two weeks. Once the shoots were regenerated from the calluses, they were to be separated from the explants with a sterile blade and grown on fresh M13 medium for two more weeks after which they were placed into sterile pots with root inducing medium M13 (without hormones) and allowed to root. When a branching root system had been established, the whole plant is either transferred to compost in a 9 cm pot and covered with a transparent plastic bag to maintain humidity. The small plants were then allowed to recover for one to two weeks before removing the plastic bag cover. The transgenic tobacco plants were re-potted and maintained in the greenhouse under general conditions as previously described (Section 2.3).

2.6 GENOMIC DNA EXTRACTION

A simple and rapid method for the preparation of plant genomic DNA for Polymerase Chain Reaction (PCR) analysis was based on the protocol developed by Edwards *et al.* in 1991. Leaf tissue samples were collected by taking up 1 cm^2 of material into the tube and the tissue was macerated using autoclaved sterile plastic pestels for each sample at room temperature (RT) in 400 µL of extraction buffer (see appendix II). The mixtures were left at room temperature until all the samples had been extracted. The extracts were heated at 65 °C for 10 min followed by centrifugation (13, 000 rpm, 5 min, RT) using a table microcentrifuge. 350 µL of each supernatant was transferred to a fresh microcentrifuge tube to which 350 µL isopropanol was added and mixed by pipetting. DNA was allowed to precipitate by storing the sampled at -20 °C for at least 1 h. Following centrifugation (14,000 rmp, 5 min, RT), the pellets were washed with 400 µL of 70% ethanol, vortexed and centrifuged as before. The ethanol was discarded and the pellets were air dried in a flow-hood or by vacuum, making sure to not over-dry the pellets. The pellets were dissolved in 50 µL SDW. The gDNA samples were RNase treated (section 2.7) and 1-2µL of the samples were sufficient for a standard 20 µL PCR which was carried out for genotyping transgenic plants.

2.7 RNase Treatment of DNA

20 µg of total DNA samples were digested with 1 µL of RNase enzyme (Ribonuclease A) (QIAGEN) and 1x RNase R reaction buffer and the mixture made up to a final volume of 200 µl to allow easy layer pipetting. The tubes were

incubated at 37°C for 15 minutes. 1µl of proteinase K (QIAGEN) was added to the tubes and incubated for a further 15 mins at 37°C.

100 µl of phenol:Chloroform (1:1 v/v) was added to each tube, vortexed and spun down for 10 mins at 14 000 rpm. The upper part of the aqueous phase was transferred to a sterile DNase-free tube to which 100µl of chloroform was added and the tubes vortexed and centrifuged for 10 mins at 14 000 rpm. The upper aqueous phase was transferred to a sterile eppendorf tube and $1/10^{th}$ of 3M sodium acetate was added to each tube and vortexed. The DNA was precipitated by the addition of 1 volume of isopropanol the tubes were vortexed and centrifuged for 15 mins at 14000 rpm. The supernatants were discarded and the pellets rinsed in 70% ethanol, centrifuged and again the ethanol removed. The pellets were either allowed to dry in a flow-hood or by vacuum, after which they were each resuspended in 50µl of DNase-free water.

2.8 RNA Extraction

2.8.1 MiniPrep

RNA extractions from *A. thaliana* seeds were carried out using the RNeasy Mini Kit (QIAGEN), according to the manufacturer's instructions with slight modifications. 150 mg of leaf tissue was ground to a fine powder under liquid nitrogen in a 2 ml eppendorf tube and 500 µl of extraction buffer RLT was added. Tubes were then vortexed vigorously and incubated in a 60°C heat block for 5 min. the mixtures were then loaded onto the mini-prep columns and the following procedures were performed according to the manufacturer's protocol. The RNA was cleaned up, prior

to use for reverse transcription, with DNaseI provided with the RNeasy kit according to manufacturer's instructions.

2.8.2 Phenol-chloroform method

RNA extractions from mature green tobacco leaves were carried out using the extraction procedure based on a method developed by Goldsbrough and Cullis, 1981. 1-7 g of fresh leaf tissue was ground to a fine powder in a pre-chilled mortar under liquid nitrogen. The frozen powder was transferred to a sterile acid-washed and autoclaved 50 ml tube which has been kept on ice while tissue preparation. 1ml/g of RNA extraction buffer (See appendix II) was added to the powdered tissue and the sampled were left on the bench to thaw with occasional vortexing. An equal amount of phenol:chloroform (See appendix II) was subsequently added and the tubes were mixed by vigorous vortexing for 1 minute. The samples were held on ice until all samples were processed and again they were vortexed for 1 minute. Immediately, the aqueous phase and the organic phase were separated by centrifugation in a precooled centrifuge machine (10 000 rpm, 15 min, 10°C). The nucleic acids containing upper aqueous layer was then transferred to a new sterile 50 ml tube and total nucleic acids were precipitated by the addition of $1/10^{th}$ volume of 3M sodium acetate (pH 6.0) and 2.5 volume of ethanol (or 1 volume of isopropanol) and placed at -20°C for at least 30 minutes. Precipitated nucleic acids were then pelleted by centrifugation (10 000 rpm, 10 mins, 4°C). The supernatant was decanted and the tubes were dried briefly on paper towels. The pellets were dissolved in 500 µl of RNase-free SDW and placed in an RNase-free microfuge tube, any insoluble debris was removed by snap spinning the samples. The RNA was subsequently precipitated from this

solution with an equal volume of 8M lithium chloride solution and kept at -20°C overnight to enrich and precipitate the RNA. RNA was recovered by centrifugation (13, 000 rpm, 20 mins, 4°C) and the supernatant discarded. The RNA pellets were washed with 500 µl of RNase-free 70% ethanol, microfuged for a further 5 minutes, and again the supernatant was discarded. The final RNA pellets were dried in a laminar flow-hood or under vacuum, taking care not to over-dry them. The pellets were re-suspended in SDW (10 µl-200µl) depending on the size of the pellet and stored at -70°C.

2.9 RNA Quantification and Analysis

Total RNA from *A. thaliana* and *N. tabacum* were quantified using NanoDrop ND-100 Spectrophotometer (NanoDrop Technologies, Wilmington, Delaware, USA). The integrity of the RNA was confirmed by running 3 µL of the total RNA on a 1% agarose TAE agarose gel at 100 V for 15 min, staining with ethidium bromide and checking for the presence of ribosomal bands when viewed under UV light. RNA samples were either stored frozen at -80°C or kept on ice for immediate use.

2.10 Gel Electrophoresis

2.10.1 Non-denatured agarose gel

0.7-1g Agarose (Bioline, UK) was melted in 1X TAE buffer (see appendix II) in a microwave oven and cooled down to ~60 °C before ethidium bromide (EtBr) was added to a final concentration of 0.05-0.1 µg/mL to the gel solution in the conical

flask, swirled and poured in the tray containing suitable number of combs. The gel was then left on the bench to set for 15-30 min and transferred to an electrophoresis tank with 1X TAE buffer. 1X loading dye (see appendix II) was added to the samples if the dye was not added prior to PCR. 5µl of ladder (HyperladderI used in most cases) as well as the samples (RNA/DNA) were loaded into separate wells and DNA was electrophoresed at 100 V for an appropriate time. The gel was subsequently exposed to UV light (Transilluminator, Syngene) and photographed using UVP Imagestore 5000.

2.10.2 DNA extraction from agarose gel

Gel extraction was carried out by using the QIAquick Gel Extraction Kit or MiniElute Kit (QIAGEN) according to the manufacturer's instruction. Alternatively, for rapid purification of DNA in an agarose gel, a spin column was assembled by placing a 0.5 mL tube inside a 1.5 mL tube and the bottom of the 0.5 mL tube was punctured with a needle. A small piece of sterile filter paper was placed inside the 0.5 mL tube to cover the hole. The agarose gel slice containing the desired DNA band was then placed into the 0.5 mL tube and was snap frozen in liquid nitrogen before centrifugation (18000 g, 2 min). The flow-through liquid was collected into a new tube. To this liquid, 5 µL of Dextran solution (10mg/mL) was added. The mixture was vortexed, followed by addition of 3 volumes of pure ethanol, vortexed again and kept at -20 °C overnight. The tubes were centrifuged (18000 g, 10 min), the pellet was washed with 75% ethanol, air-dried and re-suspended in 20 µL of SDW.

2.11 Polymerase Chain Reaction (PCR)

2.11.1 General PCR protocol for Biotaq™ DNA polymerase (Bioline)

The following PCR conditions (Table 2.1) were used to generate DNA fragments and probes for presence of the respective cDNA and northern analysis respectively, while using non-proof-reading DNA polymerase that also adds a single adenosine nucleotide (A-overhang) at the end of the PCR product. Variations in amplification conditions were required to accommodate primers of differing lengths during the amplification cycle.

Reaction Mixture Set-up (100 µL)	PCR Conditions
1 µL of forward primer (10 µM)	Lid: 100 °C
1 µL of reverse primer (10 µM)	1: 95 °C, 3 min (initial denaturation)
1 µL of dNTPs (10 mM; Promega)	2: 94 °C, 30 sec (denaturation)
1.5 µL of $MgCl_2$ (50 mM; Bioline)	3:56 °C, 30 sec (annealing)
5 µL of 10x PCR Buffer (Bioline)	4: 72 °C, 1 min (extension)
1 µL of DNA Template (~50 ng)	5: Go to 2, repeat for 35 cycles
32.5 µL of SDW	6: 72 °C, 5 min (final extension)
6 µL Cresol Red dye	7: Hold at 20 °C
1 µL of DNA polymerase (5 U/µL)	

Table 2.1 PCR conditions while using Biotaq™ Polymerase.

The PCR mixture, as in Table 2.1, was kept on ice before being placed into the PCR machine. The annealing temperature was optimised for each of the primer pairs. The extension time was directly proportional to the length of the PCR product (1 min per 1 kb). The PCR program used is shown above.

2.11.2 General PCR protocol for Phusion® High Fidelity DNA polymerase (FINNZYMES)

The following PCR conditions (Table 2.2) were used to generate DNA fragments for cloning while using a proof-reading polymerase that allows for amplification of DNA fragments with very low error-rate. Variations in amplification conditions were required to accommodate primers of differing lengths during the amplification cycle.

Reaction Mixture Set-up (50 µL)	PCR Conditions
2.5 µL of forward primer (10 µM)	Lid: 100 °C
2.5 µL of reverse primer (10 µM)	1: 98 °C, 30 sec (initial denaturation)
1.5 µL of dNTP (100 mM)	2: 98 °C, 10 sec (denaturation)
10 µL of 5x Phusion® HF Buffer	3: 58 °C, 20 sec (annealing)
1 µL of DNA Template (~50 ng)	4: 72 °C, 30 sec (extension)
33 µL of SDW	5: Go to 2, repeat for 30 cycles
0.5 µL of Phusion polymerase (2 U/µL)	6: 72 °C, 7 min (final extension)
	7: Hold at 10 °C

Table 2.2 PCR conditions while using Phusion® Polymerase.

Phusion is a modified *Pfu* DNA polymerase which has an artificial DNA binding domain fused to the DNA polymerase domain. It has a superior thermostability, the fastest extension rate (15-30 sec per kb DNA) and a very low error rate (4.4×10^{-7}) compared to other thermostable polymerases. The annealing temperature was dependent on the Tm of the oligos and the extension time was alternated depending

on the length of the fragment being amplified. PCR products produced by amplification with Phusion polymerase normally produce blunt ends.

2.11.3 A – tailing

A-tailing process is performed to add an overhang of adenosine residues to the blunt ends of a PCR product obtained by a proof-reading polymerase prior to cloning in TA cloning vectors such as pCR8/GW/TOPO (Invitrogen). The reaction conditions shown in Table 2.3 were used.

Reaction Mixture Set-up	Conditions
15 µL of Phusion PCR product 1 µL of dATP (100 mM) 1 µL of 10 x PCR buffer (Bioline) 1 µL of $MgCl_2$ 0.5 µL Biotaq™ Polymerase 1.5 µL of SDW	Lid: 100 °C 1: 70 °C, 60 min (extension)

Table 2.3 Conditions required for a-tailing.

2.12 Cloning DNA into Plasmid Vectors

2.12.1 Ligation

The appropriate plasmids were digested with the desired restriction endonuclease and dephosphorylated prior to gel extraction. Dephosphorylation was performed to inhibit recirculation in the self-ligation reaction by adding Antartic phosphatase

enzyme (NEB) (1U/μg vector) followed by a 37°C incubation for 30 minutes, 65°C for 5 minutes to deactivate the phosphatase. The tubes containing the reactions were immediately transferred on ice for 30 seconds to prevent any possibility of recirculation of the vectors. In a 10 μL ligation reaction, 50 ng of gel-extracted vector was mixed with 5-10 ng DNA fragment (typically 3:1 insert to vector molar ratio), SDW, 2 μL of 5x ligation buffer and 1μL of T4 DNA ligase (3 U/μL) were mixed and incubated overnight at 4 °C.

2.12.2 Gateway cloning

To generate PCR product for gateway cloning, an extra nucleotide sequence CACC AAA was added to the forward PCR primer in front of the first ATG (CAC CAA AATGNNNNNNN) in some cases. The purified PCR product containing CACC in the 5′ end was cloned to the pCR8 Entry vector (Invitrogen) through TOPO isomerase reaction according to the manufacturer's protocol. Once the DNA fragment was cloned into the entry vector (pCR8/GW/TOPO) which contains the *attL* sites, they were recombined into the desired destination vectors containing the *attR* sites through the LR reaction (Invitrogen).

For directional cloning, PCR products were cloned into pDonor221 vector through the BP reaction (Invitrogen) according to the manufacturer's protocol. The DNA fragment of interest was amplified in a two-step PCR approach. The first PCR was performed with gene specific primers (GSP) (13αOH SP and 13αOH ASP, See appendix I) with an additional *attB* tag in the 5′ end, the second round PCR was carried out with *attB* adaptor primers (Invitrogen) which recognise the tags of the

GSP primers. The GSP forward PCR primer has extra "aaaaagcaggctnn" in the 5′ end and the GSP reverse primer has extra "agaaagctgggtt" at the 3′ end. The sequences of the *attB* adaptor forward and reverse primers are described in Appendix I. The final PCR product was then cloned into pDonor221 vector through the BP reaction. In a 0.2 ml tube, 0.5 µl of pDonor221 (25 ng) was mixed with 1.5 µl of the purified PCR product (10-20 ng) and 0.5 µl of BP clonase II was added. The tube was incubated in a PCR machine at 25° for no less than 3 hours before 0.5 µl of proteinase K was added to inactivate the enzyme for 15 min at 37° C. The 3 µl BP reaction product was then used immediately for *E.coli* transformation and selected on solid LB medium (see appendix II) supplemented with 50 mg/l kanamycin. Once the DNA fragment was cloned into the pCR8 or pDonor vectors, which contain *attL* sites, they were recombined to the pDestination vectors containing the *attR* sites through the LR reaction (Invitrogen). In a 0.5 mL PCR tube (thin-walled; eppendorf), 1 µL (50 ng) of pEntr plasmid, 1 µL (50 ng) of the destination vector, 0.5 µL LR clonease II were mixed and incubated at 25 °C for two hours. 0.5 µL of proteinase K was added to the LR reaction mix to remove the enzymes prior to *E. coli* transformation.

2.12.3 *E. coli* transformation

The *E. coli* competent cells were thawed on ice for 15-20 min. 5 µL of the ligation reaction mixture was added to the cells. The tube was then incubated on ice for 10-15 min, heat shocked in a 42 °C water bath for 60-90 s, and immediately placed back on ice for 1 min. 600 µL of room temperature LB medium (see appendix II) was then added to the cells. Tubes were incubated at 37 °C for 1 h before being spread onto

LB plates containing the appropriate antibiotics. The plates were incubated overnight in an inverted position in a 37 °C oven to allow the transformed cells to form colonies.

2.12.4 Plasmid purification

The small scale plasmid preparations were performed using QIAprep Spin Miniprep Kit (QIAGEN), following the manufacturer's instructions. For high copy number plasmids, 10 ml of overnight culture was used. An *E.coli* glycerol stock was prepared for each relevant colony from 0.5 ml of the overnight culture prior to the plasmid prep by adding an equal volume of 20% (v/v) glycerol. The glycerol stock was then snap frozen in liquid nitrogen and stored at -70°C.

2.12.5 Large scale preparation of plasmid DNA

100 ml of sterile LB media was inoculated from a single bacterial colony grown on a selection plate containing the appropriate antibiotics and the culture was incubates overnight in a 37°C shaker. Cells were pelleted by centrifugation at 4000xg in a sterile, clean centrifuge tube for 10 min at 4°C. The bacterial pellet was resuspended in 5 ml of solution 1 containg the following: 25 mM Tris-HCl (pH 8.0), 50 mM EDTA (pH 8.0) and 1% (w/v) Glucose. This was added to 10 ml of solution 2 [200 mM sodium hydroxide and 1% (w/v) SDS] and the contents were mixed well by inversion. As the cells lysed the suspension becomes viscous after which 7.5 ml of solution 3 [5M Potassium acetate pH 4.8] was added, mixed gently and placed on ice for 20 min. Cell debris and chromosomal DNA was removed by pelleting at 9000xg

for 30 min at 4°C followed by filtration through sterile muslin. The supernatant was precipitated with two volumes of ethanol (~45 ml). After 10 mins of ice, nucleic acid was pelleted by centrifugation at 8000xg for 10min at 4°C. The pellet was drained well, resuspended in 2 ml of sterile distilled water and transferred to a 14 ml centrifuge tube. 400 mg of ammonium acetate and added to dissolve. After 15min on ice this was spun at 10000xg for 10min at 4°C and the supernatant transferred to a fresh tube containing 4 ml of ethanol. This was left on ice for 30 min then centrifuges at 10000xg for 10 min at 4°C and the pellet was drained and resuspended in 1 ml of 1mM EDTA. This was transferred to two Eppendorf tubes and 5µl of a 10µg/ml solution of ribonuclease A was added to each sample. Following incubation for 15 min at 37°C, an equal volume of phenol/chloroform was added, mixed by vortexing and the tubes were spun in table microcentrifuge for 5 min at maximum speed. The aqueous phases were transferred to new tubes containing an equal volume of chloroform and the phases were mixed by vortexing and centrifuged for 5 min at maximum speed. The aqueous phases were transferred to sterile tubes and 1/10 volume of Sodium acetate and two volumes of ethanol added to each sample. After 10 min on ice, DNA was precipitated by centrifugation for 10 min at high speed. The supernatant was discarded and the pellets were washed with 70% ethanol then resuspended in 200µl of 625 mM NaCl. These were pooled in one sterile tube and to this was added 100µl of 50% PEG-6000. DNA was precipitated for at least 1h at 4°C then pelleted at maximum speed in a microcentrifuge for 10 min. The pellet was washed with 70% ethanol, vacuum dried and resuspended in 200 µl sterile distilled water. The concentration was determined by reading the absorbency at 260 nm and the plasmid stock stored at -20°C.

2.12.6 Rapid boiling method for plasmid DNA extraction

The Rapid Boiling method developed by Holmes and Quigley (1981) was used for the extraction of plasmid DNA which were used for restriction digests. Prior to the boil prep, the appropriate *E.coli* colonies were restreaked on LB plates containing the appropriate antibiotic and incubated overnight at 37°C. The colonies were respectively scrapped with the help of a bent sterile pipette tip and resuspended in 200µl of STET solution (see appendix II) in sterile eppendorf tubes. The tubes were vortexed to break up the pellet and the tubes were immediately placed in an open-bottom rack placed in boiling water for 45-60 sec, ensuring that the tubes are at least half emerged in the boiling water. Here the cells are lysed allowing plasmids to escape, while the bacterial chromosomal DNA remains trapped in the cell debris. Boiling denatures the chromosomal DNA, after which reannealing allows the plasmids to reassociate. The tubes were spun down using a table microcentrifuge (14, 000 rpm, 1 min, RT) to remove the chromosomal DNA along with the cell debris. The supernatant containgin the plasmid was transferred to a fresh sterile eppendorf tube to which 200µl isopropanol was added and the tubes were vortexed and allow DNA precipitation at -20°C for at least 30 mins. The tubes were then centrifuged (14, 000 rpm, 10 min, RT) and the supernatants were decanted. The pellets were washed in 500µl of 70% DNase-free ethanol, vortexed and centrifuged for a further 5 mins (14, 000 rpm, RT) down for 5 mins. The supernatants were discarded and the pellets air dried and suspended in a minimum of 30µl of DNase-free SDW. The DNA was then treated with RNase as previously described in section 2.7 prior to restriction digest of the plasmid DNA.

2.12.7 Restriction enzyme digestion

Restriction digestion was performed to confirm the insertion into the vector. DNA digests were typically set up in a total volume of 20 µL or 50 µL in cases where the restricted fragment would be required for further cloning work. This reaction mix (20 µL) consisted of 2 µL of buffer (specific for each restriction enzyme), 6 µL of water, 2 µL of the desired restriction enzyme (in case of 2 enzymes being used then 1 µl each) and 0.1 – 0.5 µg of DNA (from the plasmid prep). Digests were incubated at 37 °C for 1.5 - 2 hrs. 5 µL of loading dye was added and the samples were then run on a 0.7% agarose gel. Where appropriate, 5µL of the plasmid prep of positive clones were sent for sequencing or the restricted fragments used for further cloning.

2.12.8 Agrobacterium transformation

1µl of ligation mix or 1µ of the binary vector obtained by LR recombination was added to One Shot® of *Agrobacterium tumefaciens* strain GV310 in a sterile eppendorf tube and allowed to stand one ice for 1 hour. The plasmid and were then transferred a cuvette and was transformed into *Agrobacterium* by electroporation (Zabarovsky and Winberg, 1990). 600 µl sterile LB was immediately added to the cuvette and the content transferred to a sterile 1mL eppendorf tube which was incubated for 2-3 hours in a 28-30 °C shaker. The bacterial sample were selected on LB agar plates containing 5 µg/ml tetracycline, 30 µg/ml gentamycin and suitable antibiotics depending on the resistance gene carried in the plasmid.

2.13 Northern Blot

2.13.1 Electrophoresis of RNA

The following procedures were carried out in the fume hood due to the toxicity of formaldehyde. To prepare the 100 mL formaldehyde northern gel, 90 mL of SDW, 2 mL of sodium phosphate buffer (1 M, pH 6.8) and 1 g of agarose were mixed in a 250 mL flask and melted in a microwave oven. 8 mL of de-ionised formaldehyde (adjusted to pH 7.0 with NaOH) was added when the gel solution was cooled to 50-60 °C. The gel solution was then poured into the gel caster tray and was left to set in the fume hood for 30 min before use. To prepare the RNA samples, in a 1.5 ml Eppendorf tube, 50% of the RNA (10 μg) is added to 50% of ethidium bromide sample buffer (see appendix II). The RNA samples are denatured at 65°C for 15 mins then snap cooled on ice for 1 min and then briefly centrifuged. 4 μL of Bromophenol blue loading dye (see appendix II) was added to the RNA samples immediately before gel electrophoresis. After loading the samples onto the formaldehyde gel, the gel was allowed to run at 80V for 30-40 min in gel running buffer (see appendix II) during which time the buffer was allowed to circulate using a circulator pump linked to the gel tank. The gel was then photographed under UV light to confirm equal loading and the gel run further for better rRNA separation if required.

2.13.2 Upward capillary blotting

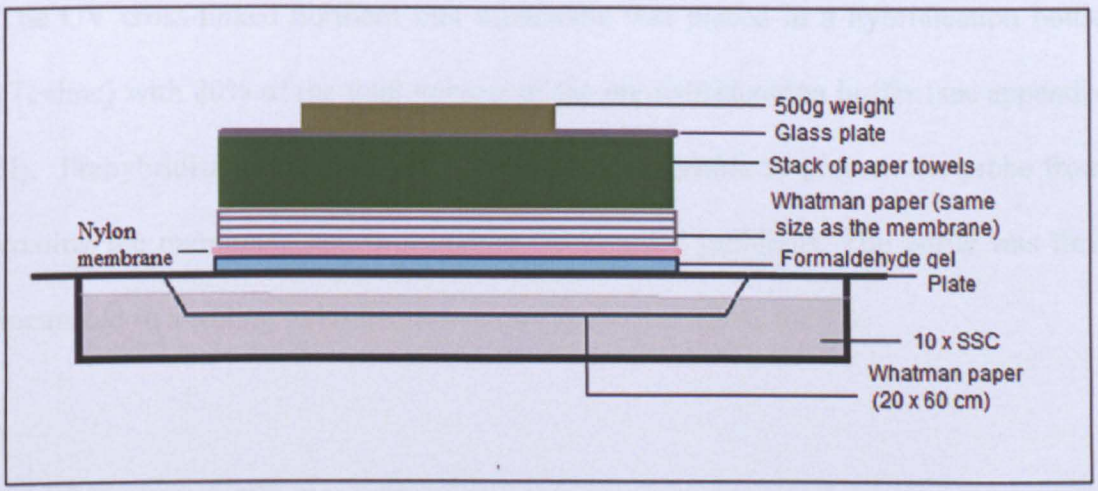


Figure 2.1 Schematic diagram of the blotting assembly.

From top: 500 g weight, paper towels, filter papers (3-4 layers), nylon membrane, formaldehyde gel, bridge-shaped filter paper (1 layer) and blotting buffer (10 x SSC)

The best low-tech method for agarose transfer is by a passive and slighltly alkaline buffer to increase blot sensitivity by efficiently moving RNA. Prior to blotting, the corner of the gel near the loading site of the control sample was cut as an indication of the position of the samples. The formaldehyde gel was inversed and placed on the bridge-shaped filter paper soaked in 10 x SSC (Figure 2.1). This was followed by transfer of the denatured RNA to a GeneScreen hybridisation transfer membrane (PerkinElmer Life Sciences) soaked in 2 x SSC prior to being placed on top of the gel. Care is taken so that no air bubble is trapped between the membrane and the gel. The upward capillary blotting assembled as shown in Figure 2.1 using 10 x SSC as transfer buffer as described by Sambrooks *et al.* 1989. The RNA was UV cross-linked to the membrane using Stratalinker™ 2400 (Stratagene) after the overnight blotting.

2.13.3 Hybridisation

The UV cross-linked northern blot membrane was placed in a hybridisation bottle (Techne) with 80% of the total volume of the pre-hybridization buffer (see appendix II). Prehybridisation is done prior to probe hybridisation to prevent the probe from coating the membrane and to minimise background problems. The bottle was then incubated in a rolling hybridization oven (Techne) at 42 °C for 4 h.

2.13.4 Radiolabelling of probes

The fragments for labelling generated in section 2.11.1 were diluted to a final concentration of 2.5-25 ng in a 45 µl of 10 mM Tris HCl, pH 8.0. The DNA was denatured by heating the sample to 95-100°C for 5 mins and snap cooled in ice for a further 5 mins. A ProbeQuant™ G-50 Micro Column (GE Healthcare) used per reaction was equilibrated with 3 ml TE (pH 8.0) containing 0.1 % (w/v) SDS. The snap cooled probe was added to an individual tube of Ready-To-Go DNA Labelling beads (Amersham Biosciences) to which 5 µl of 50 µCi [α-^{32}P] dCTP (3000 Ci/mmol) was added and gently mixed by pipetting the sample after which the sample was incubated at 37°C for 10 minutes. 400 µl of TE, 0.1% (w/v) SDS was added to the probe and this mixture was loaded onto the G-50 Micro column to remove any unincorporated nucleotides and allowed to flow through the column. The filtrate was collected in a sterile eppendorf tube and a further 400 µl of TE, 0.1% (w/v) SDS was added to the top of the column and collected in the same tube. The radio-labelled probe was denatured at 95° C for 5 mins and immediately transferred to the remaining hybridisation buffer (from section 2.13.3) and mixed gently. The

labelled mixture was carefully poured in the hybridisation bottle containing the membrane and the bottle incubated overnight at 42°C.

2.13.5 Washing of membrane

The hybridisation solution was decanted the next day and the membrane was washed in four steps to remove the unhybridised probe, as shown in table 2.5. Low stringency washes (2xSSC) remove the hybridisation solution and unhybridised probe. High stringency washes (0.1xSSC) remove partially hybridised molecules.

	SOLUTION	DURATION	TEMPERATURE
First wash	2 x SSC + 0.1% SDS	5 min	Room temperature
Second wash	2 x SSC + 0.1% SDS	15 min	65°C
Third wash	1 x SSC + 0.1% SDS	10 min	42°C
Fourth wash	0.1 x SSC + 0.1% SDS	10 min	65°C

Table 2.4 Details of wash steps after membrane hybridisation.

The duration of the fourth wash was optimized for each probe due to the differences in length and binding specificity of the probe. The remaining radioactivity of the membrane was monitored at the end of the washes by taking the membrane out of the hybridisation bottle, placing in a plastic wrap and scanned using the radioactivity monitor.

2.13.6 Detection and quantification of radioactivity

After all washes, the membrane was immediately sealed in a plastic wrap to keep it from drying out. The membrane was then placed under an X-ray (Kodak X-OMAT)

in an autoradiography cassette in the dark room under red light. The cassette was wrapped in a dark plastic bag to prevent any light exposure and stored -70°C for an appropriate length of time, depending on the counts on the membrane which can be determined using the radioactivity monitor. A membrane with <10 counts requires one week for exposure whilst those with >20 counts can produce enough signal on the film after 1-2 days.

2.13.7 Development of autoradiography film

The equipment and reagents used for developing films were purchased from Kodak. All procedures were carried out in a darkroom with red light according to the instruction of the manufacturer. The cassette containing the membrane was first warmed to room temperature or the film was immediately placed in a tank containing 10% (v/v) developer, for about 5 minutes with occasional shaking until the bands are visible. The film was then placed into a 5% (v/v) fixer solution in a separate tank: at this point the film was no longer light sensitive. The film was then washed abundantly in water and allowed to dry.

2.14 Reverse Transcription (Rt)

The total RNA (5 μg) was mixed with 1 μL of oligo$(dT)_{25}$ primer (0.5 μg/μL, Invitrogen) and 1 μL dNTP mix (10 mM each) in a 0.5 mL nuclease-free eppendorf tube. RNase-free water was then added to bring the final volume to 13 μL. The tube was incubated at 65 °C for 5 min and snap cooled on ice for at least 1 min. The contents of the tube was collected by brief centrifugation after which 4 μL 5X RT

buffer, 1 μL RNaseOUT™ Recombinant RNase inhibitor (Invitrogen), 1 μL 0.1M DTT and 1 μL Superscript™ III RT (200 U/μL, Invitrogen) were added to the tube and mixed by gentle pipetting. The tube was incubated at 50°C for 1 h. The reaction was terminated by incubating the tube at 70 °C for 15 min. The cDNA thus produced through reverse transcription was used for PCR.

After confirming the presence of the transgene by PCR of genomic DNA, an RT-PCR was performed to check whether the transgene mRNA was transcribed, using gene specific primers to PCR the cDNA. In some cases, RT-PCR was performed to check the right size of the mRNA transcribed, thus confirming that no post-transcriptional modifications have occurred, for instant RNA splicing.

2.15 GC-MS Analysis

2.15.1 Preparation of hexane extracts from plants

The extraction of crude oils and terpenes from 5 g of fresh leaf material of transgenic and control plants was done using a Polytron homogenizer (Kinematica PT10, Northern media supply LTP, North Humberside, UK) and 95% n-hexane as solvent. The extracts were centrifuged at 4000 rpm for 5 minutes after which the upper layer of supernatant was carefully transferred to a clean tube. Care was taken to not disturb the interphase between the aqueous layers. Alternatively, to obtain a more concentrated extract, the leaf tissue were ground in a mortar in hexane with the addition of a small quantity of refined sand. The extract was carefully transferred to a clean tube and the plant tissue re-extracted a further four times with the same volume of hexane to a total of 50 ml. Each of the subsequent extractions were pooled with

the first and the total extract was centrifuged for 5 minutes at 4000 g to remove any plant material, sand or any non-dissolved debris which might have been transferred. The solvent hexane containing the volatiles to be analysed were then individually dried in a rotary evaporator (BUCHI). The evaporation flask was washed twice with hexane then a final wash with ethanol, after drying each sample to prevent transfer of any contaminant to the following sample. The dried residue from each extract was taken up in a maximum of 2 ml of hexane and respectively transferred to 2ml eppendorf tubes and the lids closed properly to prevent the extracts from drying out.

2.15.2 Gas Chromatography- Mass Spectrometry (GC-MS) analysis of crude hexane extracts

In order to characterise the analytes within the crude hexane extracts, GC-MS analysis was performed on the samples by using a DSQ II mass spectrometer linked to a TRACE GC Ultra gas chromatograph (Thermo Scientific). 1µl of each sample was respectively injected with an AS3000 auto-sampler in a split-less mode (injector 250°C) onto a 30m x 0.25 mm ID fused silica capillary Zebron column ZB-1HT (Phenomenex, Macclesfield, UK) with a 0.25 µm film thickness. The ZB-1HT column was prefentially used due to its mid-polarity properties allowing the analysis to be raised to temperature up to 430°C. Unless otherwise stated, the initial oven temperature was 160°, this was ramped to 300°C at 10°C/min after a 1 min delay. The spectrum was produced using helium as carrier gas and the charged particles were formed by electron impact ionisation. The mass spectrum for the appropriate compounds was analysed by the Thermo Xcalibur™ mass spectrometry software.

2.16 Microscopy

2.16.1 Confocal Scanning microscopy

The Leica TCS SP2 AOBS confocal scanning microscope was used to study the localisation of proteins by imaging the transgenic plant materials carrying the transgene. The microscope is equipped with the Leica confocal software (LCS), a 100 mW multi-line Argon laser (458 nm, 476 nm, 488 nm, 496 nm and 514 nm) and a 1 mW He-Ne laser (543 nm) as excitation sources. The fluorescent protein signal was collected by the SP scanner (Table 2.5) and the auto-fluorescence of the chlorophyll was measured between 630 and 730 nm. The transmitted light image was collected in a separate detector.

Fluorescence Protein	Laser	Emission collected
CFP (cyan)	458 nm	465-600 nm; used with YFP: 465-505 nm
GFP (green)	488 nm	500-600 nm
YFP (yellow)	514 nm	525-600 nm; used with mRFP1: 525-560 nm
mRFP1 (red)	543 nm	590-670 nm; used with YFP: 600-725 nm

Table 2.5 Excitation and emission settings.

The plant tissues were mounted on to the microscopic glass slide with a drop of water and covered gently with the No. 1.5 cover slip (0.17 mm thick, Scientific Laboratory Supply). The slide was observed under UV light. Once the desired focus and object lens have been achieved, the UV shutter was closed and the sample was

scanned by Leica TCS SP2. The appropriate laser intensity was selected in the Leica Control Software (LXS) accordingly and the live image was acquired through continuous scan mode. The scanner and detector was set to xyz scanning mode, 514 x 514 image size, 8 bite image and 400 Hz scan rate unless otherwise mentioned. The pinhole was set to AE unit 1 as default. The voltage applied to the photomultiplier tube (PMT, AKA "detector gain value") was adjusted experimentally to obtain the best signal-to-noise ratio. The z-position and electric zoom were chosen accordingly and the series z-position scan and image maximal projection was carried out using LCS. Images were acquired using the 10x/0.4 HC PL APO CS object lens and the 40x/0.7 HCX PL FLUOTAR object lens. The 63x/1.3 HCX PL APO CS and 40x1.25-0.75 HCX PL APO CS object lenses were used to obtain images where fluorescent proteins were targeted to the endoplasmic reticulum.

2.16.2 Transient expression in onion epidermial cells

The gold powder (sphere 0.8-1.5 μm) was purchased from AlfaAesar and the Helium-driven PDS-1000/He biolistic Particle Delivery System was a product of Bio-Rad. The gold powder (30 mg) was first washed with 1 ml of ethanol and three times with SDW in a 1.5 ml Low-Bound centrifuge tube (Eppendorf) before suspension in 0.5 ml of SDW. The gold solution was then sonicated for 1 min in a water-bath sonicator (Fisher) and was dispensed to 30 μl aliquots in the Low-Bound centrifuge tubes.

To prepare the DNA-coated gold particles, 30 μl of the resuspended gold particles was used per transfection shot of 2-3 μg of plasmid DNA from isolated *E.coli*. The

gold solution was sonicated for 10 s and vortexed vigorously for 1 min. The plasmid DNA was coated onto the gold particles by adding 50 µl of 2.5 M $CaCl_2$ solution and 20 µl of 0.1 M Spermidine to the vortexing gold solution immediately. The tube was vortexed for 1 min before settling the gold on ice for 1-3 min. The gold was then pelleted by a 1 sec pulse centrifugation and the supernatant was removed gently without disturbing the gold pellet. Ethanol (150 µl) was gently laid on top of the gold pellet and the tube was again pulse centrifuged for 1 sec.

The gold pellet was then washed with 150 µl of ethanol three times without centrifugation and suspended in 30 µl of ethanol. The tube was vortexed for 2-3 sec and the gold powder was fully dispersed by pipetting. The gold solution was then loaded to three micro-carriers (Bio-Rad) and dried in a Petri dish with desiccators for about 20 min in a 30°C oven before shooting into the onion epidermal cells.

The inner epidermal cell layer of the onion (locally purchased) was peeled and cut in two rectangular (1.5cm x 3.5cm) slivers and laid flat next to each other onto Petri dishes with solid MS (1x MS, 3% Sucrose, pH 5.7 KOH) with the inner side down on the media. Water was sprayed onto the Petri dishes to keep the onion peel moisturised and to prevent it from rolling. The excessive water on the onion peels was removed by gently blotting with a piece of dry tissue. The gold powders coated with DNA were then bombarded into the onion cell under 26-28 cm Hg vacuum using 1,100 psi rupture disk. The Petri dish was then sealed with parafilm and covered with black cloth. The transformed onion peels were incubated overnight before scoring the transformed cells under the UV lamp.

2.17 Pollen Viability and Germination Assay

Flower buds or free anthers were collected before anthesis, when pollen was mature but the anthers are non-dehiscent. The anthers were placed on a microscopic slide and dissected to release the pollen, using a Stemi SV6 microscope (Zeiss, Germany). The leftover plant debris was carefully removed under a dissecting microscope and 2-3 drops of Alexander's stain added to the pollen on the slide. Any excess stain was thoroughly and carefully dried from the plant material with absorbent paper. To ensure that the stain has been completely absorbed into the pollen grains, 5-10 minutes were allowed before observing the pollen under the microscope. A cover-slip was placed over each sample and even pressure was applied on the cover-slip to ensure that all the plant components converge to one plane.

To assess the germination of the pollen, a growth medium was composed of MS, agar and 5% sucrose at pH 5.8. The medium was placed on microscopic slides and allowed to cool down. The anthers were placed on the MS-coated slides and dissected to release the pollen. The pollen were ensured to be distributed onto the germination medium. The slides were kept overnight in the dark at room temperature. The following day, 2-3 drops of Alexander's stain were applied to the coated pollen and allowed 5 minutes before observing under the microscope. The preparations were examined under a microscope (Nikon Optipot) and photographed by a Nikon DXM120 camera linked to the microscope. The counts were based on the principle of pollen tube growth among the germinated pollen.

CHAPTER 3 : CLONING OF TAXOL BIOSYNTHETIC GENES

3.1 Introduction

Using a range of techniques such as differential display of mRNA reverse-transcription PCR (DD-RT-PCR), homology-based cloning and use of an EST library from methyl jasmonate induced *Taxus* cell cultures, most of the steps and the genes involved in the Taxol biosynthetic pathway have been elucidated. The committed step of the pathway catalysed by the enzyme taxadiene synthase (TXS) establishes the taxane core, which subsequently undergoes extensive hydroxylations and further modifications to afford the advanced metabolite baccatin III to which the C13 side chain is attached in an additional 5 steps to yield Taxol.

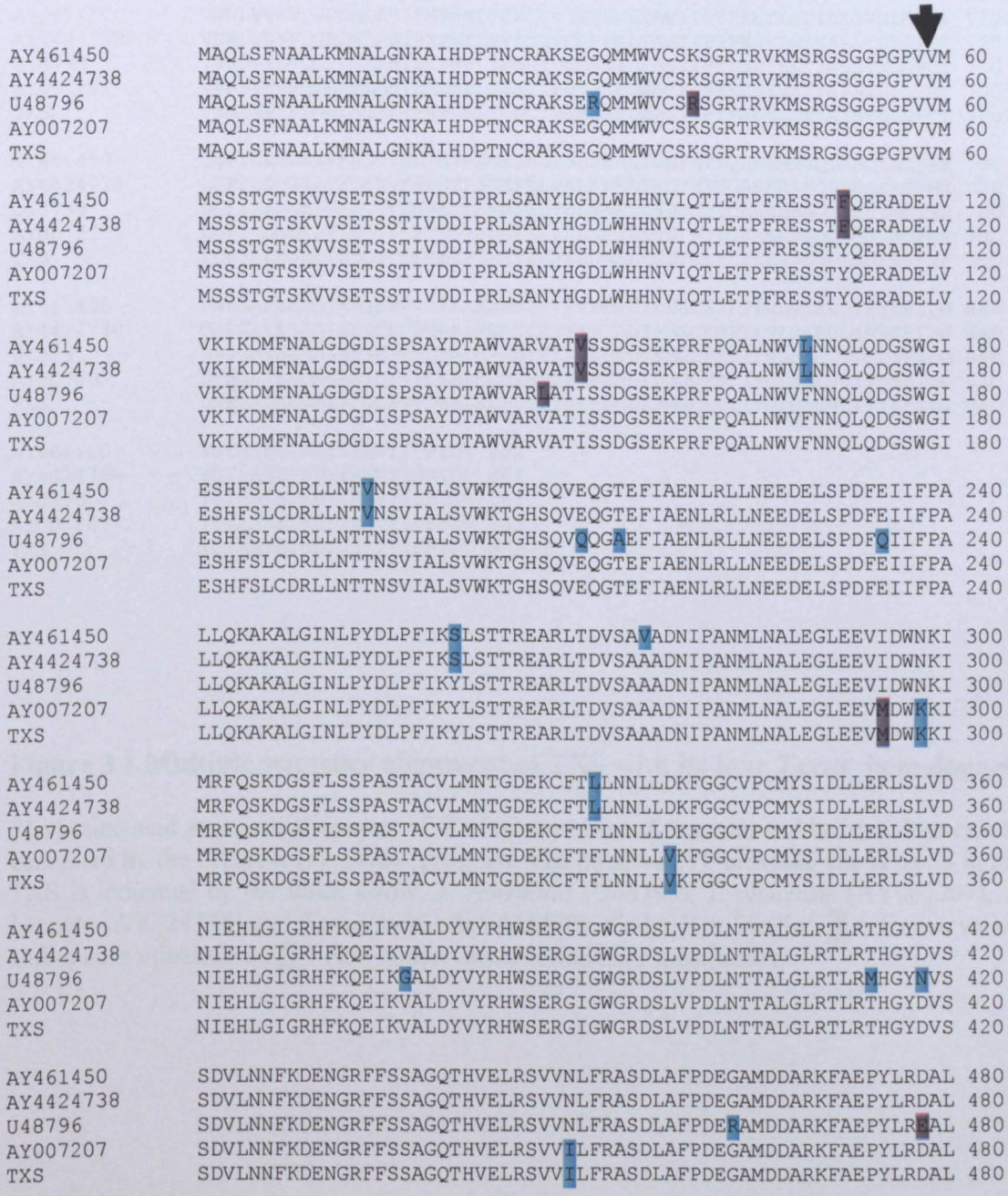
In yew tree extracts, most characterised taxanes have C5 modifications (hydroxylation or acetylation) with the next most common modifications being those with esterification at their C10/C9/C13/C2 followed by taxanes with C7 and C1 hydroxylation. However, 2α-hydroxylated taxoids have been reported to be in greater abundance than C-9 hydroxylated taxanes, which might suggest that C2 hydroxylation usually occurs prior to C9. The C7-hydroxylase has been reported to hydroxylate only highly polyoxygenated and acylated taxanes, hence placing this hydroxylase in the mid-section of the Taxol biosynthetic pathway. For taxanes bearing an oxetane ring, the more abundant taxanes are those having a hydroxyl/acyl/benzoyl group at C2, C7, C9 with less abundant taxanes bearing an

additional group at C10, which might suggest that the C10 hydroxylation occurs prior to the oxetane ring formation. The order of oxygenation of the taxane core is as yet undetermined, but based on the relative abundance of the naturally occurring taxoids (Baloglu and Kingston, 1999) and the relative frequency of oxygenations at their various positions, these data might suggest the favoured sequence of oxidation to progress from the C5 to C10, followed by C2 and C9 then C13, C7 and finally C1. Based on these published data, the genes involved in the early and mid-section steps of the Taxol biosynthetic pathway; C5-, C10-, C13-, C7-hydroxylases and the first acetyltransferase of the pathway were cloned as a starting material for this project with the aim of reconstituting the early steps of the pathway in heterologous plant system.

3.1.1 Sequence analysis of Taxadiene synthase (TXS)

Reverse transcription (RT)-PCR of the mRNA extracted from a *T. baccata* specimen growing in the University arboretum (Nottingham, UK) with primers, forward (5′ – GAA ATG GCT CAG CTC TCA TTT AAT GC-3′) and reverse (5′-TAG GAT CCT CAT ACT TGA ATT GGA TCA ATA TAA ACT TTT C-3′) designed from the sequences of previously cloned TXS sequence, amplified a cDNA with an approximately 2600 bp open reading frame encoding a predicted protein of 862 residues (Work previously carried out by Dr. Rupert Fray). The deduced full-length protein showed 100%, 98%, 98% and 97% homology to the previously cloned TXS enzymes from *T. chinensis, T. media, T. baccata* and *T. brevifolia* respectively (Figure 3.1). Sequence analysis of TXS revealed that the N-terminal domain of the

protein contains a cleavable putative chloroplast transit peptide (cTP) between amino acid residue 58 and 59, as predicted by the sub-cellular localisation prediction program PSORT (www.psort.org), indicating TXS to be located to the plastid. The TXS cDNA was previously cloned (work done by Dr. Rupert Fray) into a plant binary vector thus named pBCTaxK which has been used for plant transformation.



```
AY461450        ATKISTNTKLYKEIEYVVEYPWHMSIPRLEARSYIDSYDDDYVWQRKTLYRMPSLSNSKC 540
AY4424738       ATKISTNTKLYKEIEYVVEYPWHMSIPRLEARSYIDSYDDDYVWQRKTLYRMPSLSNSKC 540
U48796          ATKISTNTKLFKEIEYVVEYPWHMSIPRLEARSYIDSYDDNYVWQRKTLYRMPSLSNSKC 540
AY007207        ATKISTNTKLFKEIEYVVEYPWHMSIPRSEARSYIDSYDDDYVWERKTLYRMPSLSNSKC 540
TXS             ATKISTNTKLFKEIEYVVEYPWHMSIPRSEARSYIDSYDDDYVWERKTLYRMPSLSNSKC 540

AY461450        LELAKLDFNIVQSLHQEELKLLTRWWKESGMADINFTRHRVAEVYFSSATFEPEYSATRI 600
AY4424738       LELAKLDFNIVQSLHQEELKLLTRWWKESGMADINFTRHRVAEVYFSSATFEPEYSATRI 600
U48796          LELAKLDFNIVQSLHQEELKLLTRWWKESGMADINFTRHRVAEVYFSSATFEPEYSATRI 600
AY007207        LELAKLDFNIVQSLHQEELKLLTRWWKESGMADINFTRHRVAEVYFSSATFEPEYSATRI 600
TXS             LELAKLDFNIVQSLHQEELKLLTRWWKESGMADINFTRHRVAEVYFSSATFEPEYSATRI 600

AY461450        AFTKIGCLQVLFDDMADIFATLDELKSFTEGVKRWDTSLLHEIPECMQTCFKVWFKLMEE 660
AY4424738       AFTKIGCLQVLFDDMADIFATLDELKSFTEGVKRWDTSLLHEIPECMQTCFKVWFKLMEE 660
U48796          AFTKIGCLQVLFDDMADIFATLDELKSFTEGVKRWDTSLLHEIPECMQTCFKVWFKLMEE 660
AY007207        AFTKIGCLQVLFDDMADIFATLDELKSFTEGVKRWDTSLLHEIPECMQTCFKVWFKLIEE 660
TXS             AFTKIGCLQVLFDDMADIFATLDELKSFTEGVKRWDTSLLHEIPECMQTCFKVWFKLIEE 660

AY461450        VNNDVVKVQGRDMLAHIRKPWELYFNCYVQEREWLEAGYIPTFEEYLKTYAISVGLGPCT 720
AY4424738       VNNDVVKVQGRDMLAHIRKPWELYFNCYVQEREWLEAGYIPTFEEYLKTYAISVGLGPCT 720
U48796          VNNDVVKVQGRDMLAHIRKPWELYFNCYVQEREWLEAGYIPTFEEYLKTYAISVGLGPCT 720
AY007207        VNNDVVKVQGRDMLAHIRKPWELYFNCYVQEREWLDAGYIPTFEEYLKTYAISVGLGPCT 720
TXS             VNNDVVKVQGRDMLAHIRKPWELYFNCYVQEREWLDAGYIPTFEEYLKTYAISVGLGPCT 720

AY461450        LQPILLMGELVKDDVVEKVHYPSNMFELVSLSWRLTNDTKTYQAEKARGQQASGIACYMK 780
AY4424738       LQPILLMGELVKDDVVEKVHYPSNMFELVSLSWRLTNDTKTYQAEKARGQQASGIACYMK 780
U48796          LQPILLMGELVKDDVVEKVHYPSNMFELVSLSWRLTNDTKTYQAEKARGQQASGIACYMK 780
AY007207        LQPILLMGELVKDDVVEKVHYPSNMFELVSLSWRLTNDTKTYQAEKARGQQASGIACYMK 780
TXS             LQPILLMGELVKDDVVEKVHYPSNMFELVSLSWRLTNDTKTYQAEKARGQQASGIACYMK 780

AY461450        DNPGATEEDAIKHICRVVDRALKEASFEYFKPSNDIPMGCKSFIFNLRLCVQIFYKFIDG 840
AY4424738       DNPGATEEDAIKHICRVVDRALKEASFEYFKPSNDIPMGCKSFIFNLRLCVQIFYKFIDG 840
U4879           DNPGATEEDAIKHICRVVDRALKEASFEYFKPSNDIPMGCKSFIFNLRLCVQIFYKFIDG 840
AY007207        DNLGATEEDAIKHICRVVDRALKEASFEYFKPSNDIPMGCKSFIFNLRLCVQIFYKFIDG 840
TXS             DNLGATEEDAIKHICRVVDRALKEASFEYFKPSNDIPMGCKSFIFNLRLCVQIFYKFIDG 840

AY461450        YGIANEEIKDYIRKVYIDPIQV 862
AY4424738       YGIANEEIKDYIRKVYIDPIQV 862
U48796          YGIANEEIKDYIRKVYIDPIQV 862
AY007207        YGIANEEIKDYIRKVYIDPIQV 862
TXS             YGIANEEIKDYIRKVYIDPIQV 862
```

Figure 3.1 Multiple sequence alignment of TXS with its four *Taxus* homologues.

The amino acid sequence alignment of TXS cloned from *T. baccata* and its homologues was generated by the "Protein Boxshade" program. The putative cTP cleavage site at aa 58 in the TXS is indicated by the black arrow. *T. brevifolia* (U48796), *T. chinensis* (AY007207), *T. baccata* (AY424738) and *Taxus media (*AY461450). Code: X = Similar, X = Conservative amino acid substitution, X = Non-conservative amino acid substitution.

3.1.2 CYP monooxygenases and acetyl transferase of the early steps of the Taxol pathway

All the oxygenases and acetyltransferase cloned during this research study were isolated by RT-PCR amplification from the mRNA extracted from a *T. baccata* specimen growing in Warwick, UK (*T. baccata* mRNA provided by Professor Graham Seymour, Plant Sciences Division, University of Nottingham, UK). The first hydroxylation of the olefin taxadiene is catalysed by the enzyme taxadiene-5α-hydroxylase (T5OH), which adds a hydroxyl group at the C-5 position of the taxane core. This enzyme, which is deemed to catalyse the slow step of the Taxol biosynthetic pathway, has been previously isolated by homology-based cloning from *T. cuspidata* (AY289209) and *T. chinensis* (AY741375).

Based on the previously cloned T5OH sequences, the oligonucleotides 5Tax5alph (5′- CAC CAA AAT GGA CGC CCT GTA TAA GAG C-3′) and 3Tax5alph 3′(5′-TTC TCC TTC AAT TGA CTA TGG TCT CGG-3′) were used to amplify the ~1600 bp *T. baccata* T5OH, with a proof-reading polymerase as described in section 2.11.2 (Cloning done by former BSc student Sophia Sumal). The 1600 bp amplified product was incubated for an additional 1hr with a non-proof reading polymerase (Taq) as described in section 2.11.3, to provide A overhangs, and cloned directly into pCR8/GW/TOPO vector (Invitrogen, USA), as described in section 2.12.2. The pCR8 vector carries a spectinomycin resistance gene (SpnR) (Figure 3.3) and is a T-tailed vector, with the T-overhangs to complementarily anneal to the A-overhang produced by A-tailing. A colony PCR was performed on a few of the colonies obtained on LB plates containing 100 μg/ml spectinomycin and two of the clones indicated the presence of the T5OH cDNA. These colonies were selected for plasmid

preparation (Section 2.12.4) and the plasmid DNA (Clone X and Clone Y) were sequenced (MWG, Germany) using the GW1 and GW2 primers (Appendix I) located on the pCR8 vector.

Sequencing of the clones indicated the insertion of the T5OH cDNA in the wrong orientation in clone Y and nucleotide substitution in Clone X from T to C at position 615 position (Figure 3.2). This change in nucleotide sequence resulted in a non-conservative amino acid mutation from TTG (Leucine) to TCG (Serine). This mutation might have affected the structure of T5OH, hence the shape and/or function of the enzyme.

```
Clone X     CACCAGGAGGGGTGAAGACCATATAGTTATGCGCTCTGCTCTTGCAGGTTTTTTCGGCCC 480
AY289209    CACCAGGAGGGGTGAAGACCATATAGTTATGCGCTCTGCTCTTGCAGGTTTTTTCGGCCC 480

Clone X     TGGTGCGCTGCAGAGTTACATTGGTAAAATGAATACAGAGATCCAGAGTCATATCAACGA 540
AY289209    TGGTGCGCTGCAGAGTTACATTGGTAAAATGAATACAGAGATCCAGAGTCATATCAACGA 540

Clone X     AAAATGGAAGGGAAAAGATGAGGTGAATGTACTTCCTTTGGTAAGAGAGCTCGTCTTCAA 600
AY289209    AAAATGGAAGGGAAAAGATGAGGTGAATGTACTTCCTTTGGTAAGAGAGCTCGTCTTCAA 600
                                S
Clone X     CATTTCGGCCATCTCGTTTTTCAACATATATGATAAGCAGGAACAGGATCGTCTGCATAA 660
AY289209    CATTTCGGCCATCTTGTTTTTCAACATATATGATAAGCAGGAACAGGATCGTCTGCATAA 660
                                L
Clone X     GCTTTTGGAAACTATTCTGGTCGGAAGTTTTGCTCTTCCGATTGACTTGCCCGGATTTGG 720
AY289209    GCTTTTGGAAACTATTCTGGTCGGAAGTTTTGCTCTTCCGATTGACTTGCCCGGATTTGG 720

Clone X     TTTCCATAGAGCACTCCAGGGACGGGCCAAGCTCAACAAAATTATGCTGTCTTTAATTAA 780
AY289209    TTTCCATAGAGCACTCCAGGGACGGGCCAAGCTCAACAAAATTATGCTGTCTTTAATTAA 780

Clone X     AAAGAGAAAAGAAGATCTGCAGTCTGGATCGGCAACAGCCACGCAGGATCTGCTCTCTGT 840
AY289209    AAAGAGAAAAGAAGATCTGCAGTCTGGATCGGCAACAGCCACGCAGGATCTGCTCTCTGT 840
```

Figure 3.2 Sequence alignment of Clone X and T5OH from *T. cuspidata*.

Nucleotide sequence alignment showing the 421 to 840 bp of the T5OH of Clone X aligned against that of T5OH *T. cuspidata* using Clustal W1.2 Multiple sequence alignment program. Base change at position 615 from T to C results in amino acid change from Leucine to Serine. Colour code: X = PstI sites, X = non-similar, X = similar.

3.2 Results

3.2.1 Correcting the mutation in T5OH cDNA

Clone Y carried the T5OH cDNA in the wrong orientation but no nucleotide substitution at position 615 was observed, unlike in the T5OH cDNA of Clone X. The pCR8 vector carries an EcoR1 site on each of the TOPO cloning sites, but the presence of an EcoR1 site within the T5OH cDNA meant that it could not be digested with the enzyme for reinsertion into the correct orientation.

Sequence analysis indicated that the nucleotide change from TTG to TCG occurred between two *PstI* sites (Figure 3.2 & 3.3A). Hence, clones X and Y were both digested with *PstI* (section 2.12.6) to release a 309 bp fragment which was followed by dephosphorylation of Clone X as described in section 2.12.1. The fragments of size 309 bp and 4027 bp from clone Y and X respectively (Figure 3.3B) were gel extracted (section 2.10.2) and the 309 bp fragment from clone Y with the correct sequence was ligated (section 2.12.1) to the dephosphorylated linearised clone X vector. Following ligation, the mixture was transformed into one shot *E.coli* Max Efficiency cells (Invitrogen, UK) (section 2.12.3) and selected on LB plates containing 100 μg/ml spectinomycin. The presence of the 309 bp insert was checked by PCR on the colonies resistant to spectinomycin, using 5Tax5alp and 3Tax5alph as primers (see appendix I). Figure 3.3C shows the presence of the correct size band in colony 2 indicating the insertion of the 309 bp fragment in the vector. The pure plasmid from colony 2, thus named, pCR8::T5OH was sequenced with the GW1 and GW2 primers prior to transfer in the plant binary vector pGWB8.

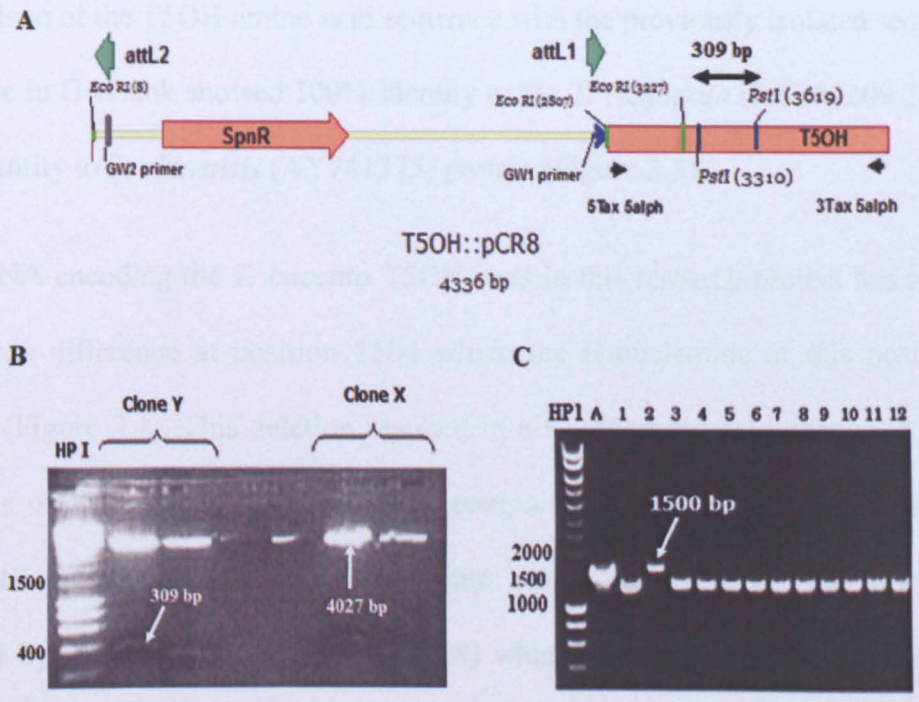


Figure 3.3 Clone pCR8CloneX

(A) The linear schematic diagram of Clone X construct generated by TOPO cloning, showing the positions of the primers used for PCR amplification of the full length T5OH cDNA. *Pst*I restriction sites are also indicated where the clone X and Y were digested (B) Gel picture of Clone pCR8Clone X and pCR8CloneY digested with *Pst*I enzyme to release a 309 bp and a 4027 bp fragment. (C) Colony 2 showing the presence of the correct size of T5OH (~1500 bp), where A is clone X used as a control. The other bands having a lower weight than 1500 bp consist of the clones for which the digested vector from clone Y was not linearised to the 309 bp fragment (See B).

3.2.2 Sequence analysis of T5OH and its homologues

The deduced amino acid sequence of the T5OH protein indicated that the mutation has been successfully corrected. The T5OH clone contains an ORF of 1563 bp which translates into a protein of 503 amino acids. It has been shown by BLAST analysis (http://www.ncbi.nlm.nih.gov/BLAST/) that T5OH is a member of the cytochrome P450 family as previously reported (Jennewein *et al.*, 2004a). Multiple sequence

comparison of the T5OH amino acid sequence with the previously isolated sequences available in Genbank showed 100% identity to the *T. cuspidata* (AY289209.2) and a 98% identity to *T. chinensis* (AY741375) protein (Figure 3.5).

The cDNA encoding the *T. baccata* T5OH used in this research project has a single nucleotide difference at position 1504 where the G-nucleotide at this position is deleted (Figure 3.4). This deletion resulted in a slight amino acid change in the C-terminus of the protein sequence when compared to the previously published *T. cuspidata* protein AY289209.1. The same amino acid change was previously reported by Rontein and co-workers (2008) which led to the GenBank entry being modified accordingly to AY289209.2. Instead of the IKLFPETIVN C-terminus previously described for AY289209.1, AY289209.2 and *T. baccata* T5OH used in this case, have an amino acid sequence IKLFPRP (Figure 3.5). This change in nucleotide sequence has resulted in an amino acid change from glutamic acid (E) (GAG) to arginine (R) (AGA) and a frame-shift generating an alternative stop codon differing by five amino acids from the original one.

Sequence analysis of T5OH (Figure 3.5) by the "Predictprotein" software (http://www.predictprotein.org/) revealed the presence of a putative transmembrane domain TESFSIALSAIAGILLLLLLF at amino acids 22 to 43. This transmembrane region of T5OH might serve as a membrane anchor and halt translocation into the endoplasmic reticulum through the membrane translocation channel, suggesting that the activity of T5OH is associated to the outer face of the endoplasmic reticulum via this hydrophobic N-terminal membrane integration region.

```
AY289209.1      AAAATATCAGGGGATCCACTCCCTCCTCTTCCTTCCAAGGGATTTTCCATTAAACTGTTT 1500
AY289209.2      AAAATATCAGGGGATCCACTCCCTCCTCTTCCTTCCAAGGGATTTTCCATTAAACTGTTT 1500
T5OH            AAAATATCAGGGGATCCACTCCCTCCTCTTCCTTCCAAGGGATTTTCCATTAAACTGTTT 1500

                   E
AY289209.1      CCCGAGACCATAGTCAATTGAAGGAGAAAACCACAGTGCAGAACTGCTATTCTTGAATCC 1560
AY289209.2      CCG-AGACCATAGTCAATTGAAGGAGAAAACCACAGTGCAGAACTGCTATTCTTGAATCC 1559
T5OH            CCG-AGACCATAGTCAATTGAAGGAGAAAAGGGCG------------------------- 1534
                   R
```

Figure 3.4 Nucleotide sequence alignment of T5OH and its homologues.

The sequence alignment of the T5OH nucleotide and its *Taxus* homologues AY289209.1 and AY289209.2 was formatted by ClustalW Multisequence alignment program. The nucleotide difference at position 1504 in the republished AY289209.2 and T5OH demonstrates a deletion of the G-nucleotide at this position leading to a frame-shift and generation of an alternative carboxyl terminus different by five amino acids. Colour code: X = similar, X = nucleotide change, X = nucleotide coding for stop codon, X = alternative stop codon

Sequence analysis of T5OH also revealed other characteristic sequence elements of CYP enzymes, including a highly conserved proline, phenylalanine, glycine (PFG) motif, a heme-iron ligand FGGGQRSCVG at amino acid 438-448 and the essential cysteine at position 445 of the protein which is a well-conserved region of the heme-binding domain of CYP proteins essential for enzyme binding and activity (Nelson, 1999). Following confirmatory sequencing, the T5OH cDNA was cloned via an LR recombination reaction (section 2.12.2) into the plant binary vector pGWB8 (Invitrogen) (Figure 3.6 A) and selected on LB plates containing Kanamycin (50 µg/ml). Two colonies were checked by PCR for the presence of the T5OH cDNA transfer by using the set of primers 5Tax5alp and 3Tax5alp (Figure 3.6 B). The LR recombinated vector thus named pGWB8::T5OH, was transformed into *Agrobacterium* strain GV3101 as described in section 2.12.7 which was used to transform a wild type *Nicotiana tabacum var. Petit Havana* line.

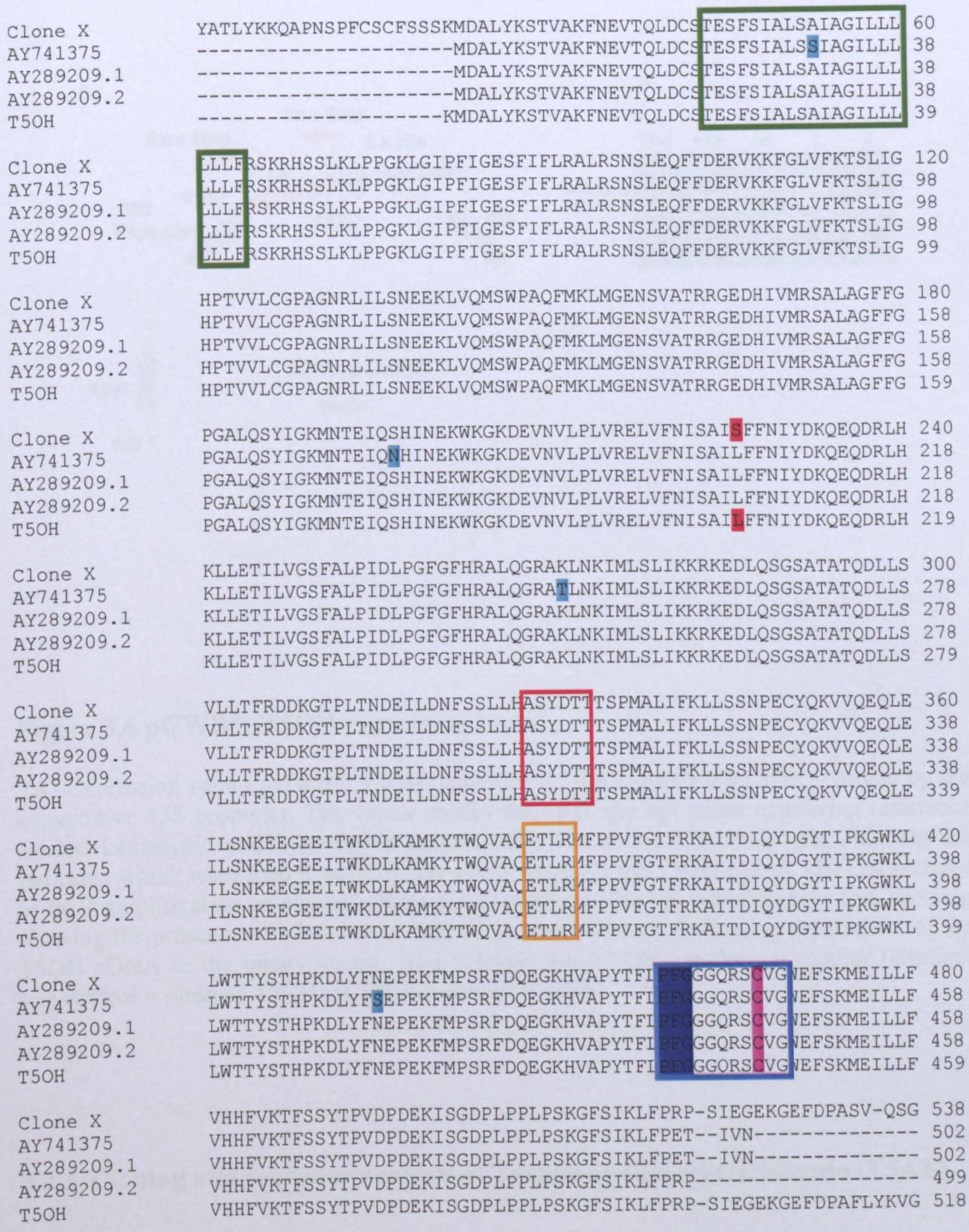


Figure 3.5 Amino acid sequence alignment of the corrected T5OH and its *Taxus* homologues.

The sequence alignment of the corrected mutation in T5OH and its *Taxus* homologues was formatted by T-coffee and produced by "Protein Boxshade" programs. The transmembrane helix is indicated by the green rectangle, the Glu-X-X-Arg indicated by the yellow box, the CYP cysteine heme-iron ligand signature is indicated by the blue rectangle and another conserved domain is indicated by the red rectangle. AY289209.1 AND AY289209.2 from *T. cupidata*; AY741375 from *T. chinensis*. Colour code: X = Similar, X = Non-conservative amino acid substitution, X = Conserved Cysteine residue within the heme-binding region, X = PFG domain, X = amino acid change in Clone X corrected in T5OH

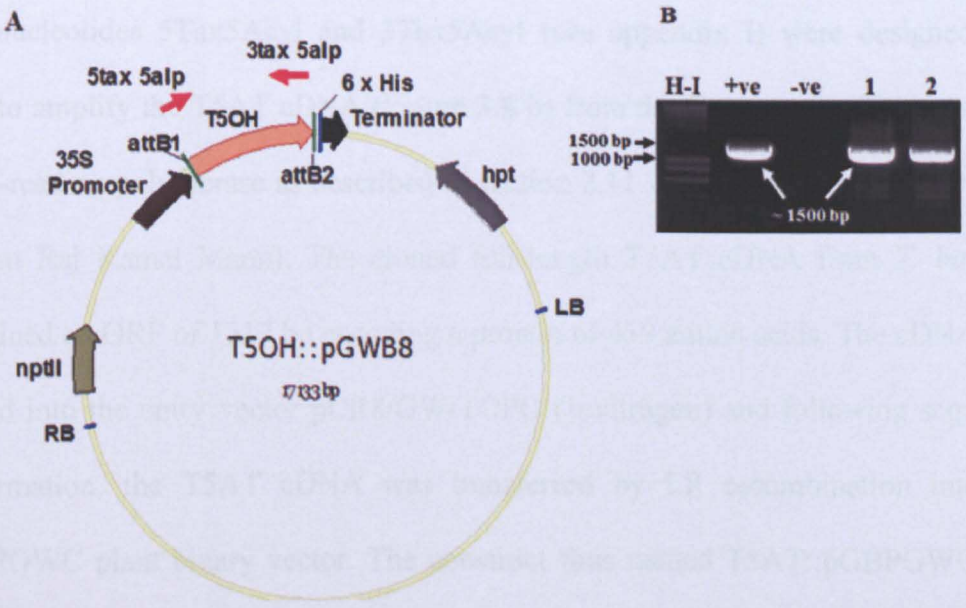


Figure 3.6 pGWB8::T5OH expression vector

(A) Expression vector pGWB8::T5OH consisting of full length T5OH cDNA driven by the constitutive 35S promoter. The vector carries the nptII and hpt genes conferring resistance against kanamycin and hygromycin respectively. The primers 5tax5alp and 3tax5alp are indicated which were used to amplify the whole length of the T5OH cDNA. (B) Gel pictures of PCR amplification of plasmids from *E.coli* colonies 1 and 2, with 5tax5alp and 3tax5alp showing the presence of a band size 1500 bp in both plasmids, indicating the transfer of the T5OH cDNA to the binary vector. H-I = HyperladderI DNA molecular marker (Bioline) +ve control = plasmid T5OH::pCR8, -ve control = water

3.2.3 Cloning and sequence analysis of Taxadiene-5acetyl-transferase (T5AT)

Biochemical studies have suggested that the third specific step of the Taxol biosynthesis pathway is the acetylation of the C5-hydroxyl group of taxadiene-5α-ol to afford taxa-4(20),11(12)-dien-5α-yl acetate (Wheeler *et al.*, 2001). The taxadiene-5acetyl-transferase (T5AT) catalysing this acetylation reaction has been isolated from the *Taxus* species; *T. cuspidata* (AF190130), *T. chinensis* (AY078285) and

Taxus x media (AY453402). Based on these published sequences, the oligonucleotides 5Tax5Acyl and 3Tax5Acyl (see appendix I) were designed and used to amplify the T5AT cDNA (Figure 3.8 b) from the *T. baccata* mRNA, using a proof-reading polymerase as described in section 2.11.2 (Work done by former MSc student Raj Kamal Mann). The cloned full-length T5AT cDNA from *T. baccata* contained an ORF of 1317 bp encoding a protein of 439 amino acids. The cDNA was cloned into the entry vector pCR8/GW/TOPO (Invitrogen) and following sequence confirmation, the T5AT cDNA was transferred by LR recombination into the pGBPGWC plant binary vector. The construct thus named T5AT::pGBPGWC was then transferred to *Agrobacterium* which was used for plant transformation.

The cloned T5AT from *T. baccata* had a 99% identity to both *Taxus x media* AY453402 and *T. chinensis* AY078285 and a 98% identity to *T. cuspidata* AF190130 (Figure 3.7). The isolated clone possesses a typical acyltransferase motif HXXXDG, characteristic of other *Taxus* acyltransferases, where the histidine may function in acyl group transfer from acetyl-CoA to the substrate alcohol taxadiene-5α-ol. This conserved motif has been suggested by St-Pierre *et al.*, (1998) to lie at or close to the active centre of acyltransferases.

Walker *et al.* in 2000 suggested that the sequence of the T5AT cloned from *T. cuspidata* did not appear to encode an organellar targeting information; however, "Predict protein" software indicated the presence of a transmembrane helix at amino acids 147-164 of the T5AT cloned from *T. baccata* (Figure 5.7), which might be involved in the translocation of enzyme into the ER.

```
AY453402        MEKTDLHVNLIEKMMVGPSLPPPKTTLQLSSIDNLPGVRGSIFNALLIYNASPSPTMISA 60
T5AT            MEKTDLHVNLIEKVMVGPSLPPPKTTLQLSSIDNLPGVRGSIFNALLIYNASPSPTMISA 60
AY078285        MEKTDLHVNLNEKVMVGPSLPLPKTTLQLSSIDNLPGVRGSIFNALLIYNASPSPTMVSA 60
AF190130        MEKTDLHVNLIEKVMVGPSPPLPKTTLQLSSIDNLPGVRGSIFNALLIYNASPSPTMISA 60

AY453402        DPAKLIREALAKILVYYPPFAGRLRETENGDLEVECTGEGAMFLEAMADNELSVLGDFDD 120
T5AT            DPAKLIREALAKILVYYPPFAGRLRETENGDLEVECTGEGAMFLEAMADNELSVLGDFDD 120
AY078285        DPAKLIREALAKILVYYPPFAGRLRETENGDLEVECTGEGAMFLEAMADNELSVLGDFDD 120
AF190130        DPAKPIREALAKILVYYPPFAGRLRETENGDLEVECTGEGAMFLEAMADNELSVLGDFDD 120

AY453402        SNPSFQQLLFSLSLDTNFKDLPFLVVQVTRFTCGGFVVGVSFHHGVCDGRGAAQFLKGLA 180
T5AT            SNPSFQQLLFSLSLDTNFKDLPLLVVQVTRFTCGGFVVGVSFHHGVCDGRGAAQFLKGLA 180
AY078285        SNPSFQQLLFSLPLDTNFKDLPLLVVQVTRFTCGGFVVGVSFHHGVCDGRGAAQFLKGLA 180
AF190130        SNPSFQQLLFSLPLDTNFKDLSLLVVQVTRFTCGGFVVGVSFHHGVCDGRGAAQFLKGLA 180

AY453402        EMARGEVKLSLEPIWNRELVKLDDPKYLQFFHFEFLRAPSIVEKIVQTYFIIDFETINYI 240
T5AT            EMARGEVKLSLEPIWNRELVKLDDPKYLQFFHFEFLRAPSIVEKIVQTYFIIDFETINYI 240
AY078285        EMARGEVKLSLEPIWNRELVKLDDPKYLQFFHFEFLRAPSIVEKIVQTYFIIDFETINYI 240
AF190130        EMARGEVKLSLEPIWNRELVKLDDPKYLQFFHFEFLRAPSIVEKIVQTYFIIDFETINYI 240

AY453402        KQSVMEECKEFCSSFEVASAMTWIARTRAFQIPESEYVKILFGMDMRNSFNPPLPSGYYG 300
T5AT            KQSVMEECKEFCSSFEVASAMTWIARTRAFQIPESEYVKILFGMDMRNSFNPPLPSGYYG 300
AY078285        KQSVMEECKEFCSSFEVASAMTWIARTRAFQIPESEYVKILFGMDMRNSFNPPLPSGYYG 300
AF190130        KQSVMEECKEFCSSFEVASAMTWIARTRAFQIPESEYVKILFGMDMRNSFNPPLPSGYYG 300

AY453402        NSIGTACAVDNVQDLLSGSLLRAIMIIKKSKVSLNDNFKSRAVVKPSELDVNMNHENVVA 360
T5AT            NSIGTACAVDNVQDLLSGSLLRAIMIIKKSKVSLNDNFKSRAVVKPSELDVNMNHENVVA 360
AY078285        NSIGTACAVDNVQDLLSGSLLRAIMIIKKSKVSLNDNFKSRAVVKPSELDVNMNHENVVA 360
AF190130        NSIGTACAVDNVQDLLSGSLLRAIMIIKKSKVSLNDNFKSRAVVKPSELDVNMNHENVVA 360

AY453402        FADWSRLGFDEVDFGWGNAVSVSPVQQQCELAMQNYFLFLKPSKNKPDGIKILMFLPLSK 420
T5AT            FADWSRLGFDEVDFGWGNAVSVSPVQQQCELAMQNYFLFLKPSKNKPDGIKILMFLPLSK 420
AY078285        FADWSRLGFDEVDFGWGNAVSVSPVQQQCELAMQNYFLFLKPSKNKPDGIKILMFLPLSK 420
AF190130        FADWSRLGFDEVDFGWGNAVSVSPVQQQSALAMQNYFLFLKPSKNKPDGIKILMFLPLSK 420

AY453402        MKSFKIEMEAMMKNYVAKV 439
T5AT            MKSFKIEMEAMMKKYVAKV 439
AY078285        MKSFKIEMEAMMKKYVAKV 439
```

Figure 3.7 Sequence comparison of T5AT and its *Taxus* homologues.

The sequence alignment of the T5AT protein and its homologues was generated by the "Protein Boxshade" program (http://www.fr33.net/boxshadeprotein.php.).The transmembrane anchor helix found at amino acid 147-164 is indicated by the red rectangle and the putative acyl HXXXDG group transfer motif is represented by the yellow rectangle. *Taxus x media* - AY453402, *T. chinensis* - AY078285, *T. cuspidata* - AF190130. Colour code: X = Similar, X = Conservative amino acid substitution, X = Non-conservative amino acid substitution.

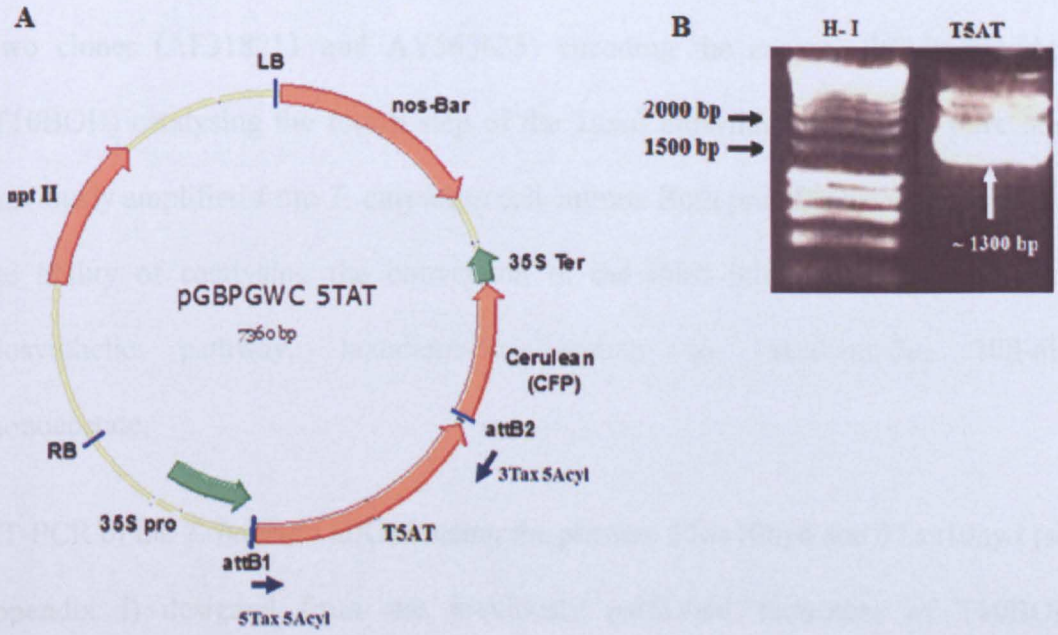


Figure 3.8 pGBPGWC::T5AT expression vector

(A) Expression vector pGBPGWC::T5AT consisting of full length T5AT cDNA driven by the constitutive 35S promoter. The vector carries the nos-bar gene conferring resistance against basta. The primers 5tax5acyl and 3tax5acyl are indicated which were used to amplify the whole length of the T5AT cDNA. (B) Gel picture of PCR amplification of the T5AT cDNA with 5tax5acyl and 3tax5acyl showing the presence of a band size of around 1300 bp, indicating the presence of the T5AT in the destination vector pGBPGWC. H-I - DNA molecular marker HyperladderI

3.2.4 Cloning and sequence analysis of taxoid 10β-hydroxylase (T10BOH)

Two clones (AF318211 and AY563635) encoding the enzyme 10β-hydroxylase (T10BOH) catalysing the fourth step of the Taxol biosynthetic pathway have been previously amplified from *T. cuspidata* cell culture. Both proteins have demonstrated the ability of catalysing the conversion of the third intermediate of the Taxol biosynthetic pathway, taxadiene-5α-yl-acetate to taxadiene-5α, 10β-diol monoacetate.

RT-PCR of the *T. baccata* mRNA using the primers 5Tax10hyd and 3Tax10hyd (see appendix I) designed from the previously published sequences of T10BOH, amplified a cDNA with a 1458 bp open reading frame encoding a predicted protein of 485 amino acid residues (work done by former BSc student Zac Bourn).

The deduced full-length T10BOH cDNA was sequenced and the amino acid sequence was aligned to other T10BOH *Taxus* homologues using ClustalW sequence alignment program. T10BOH cloned from *T. baccata* showed 95% and 63% homology to the *T. cuspidata* AY563635 and AF31811 respectively (Figure 3.9). Sequence analysis indicated the presence of two transmembrane helices at amino acid positions 11-28 (QLQSILFLTLLALLLLCLF) and 289-306 (LHASYETSV SPTVL) as predicted by the sub-cellular localisation prediction program Predictprotein (www.predictprotein.org). This suggested that T10BOH is of a membrane protein localised to the ER. Sequence analysis also indicated other features of the CYP family including the PFG motif and the cysteine heme-iron ligand signature FGAGARICPG at amino acids 437-447.

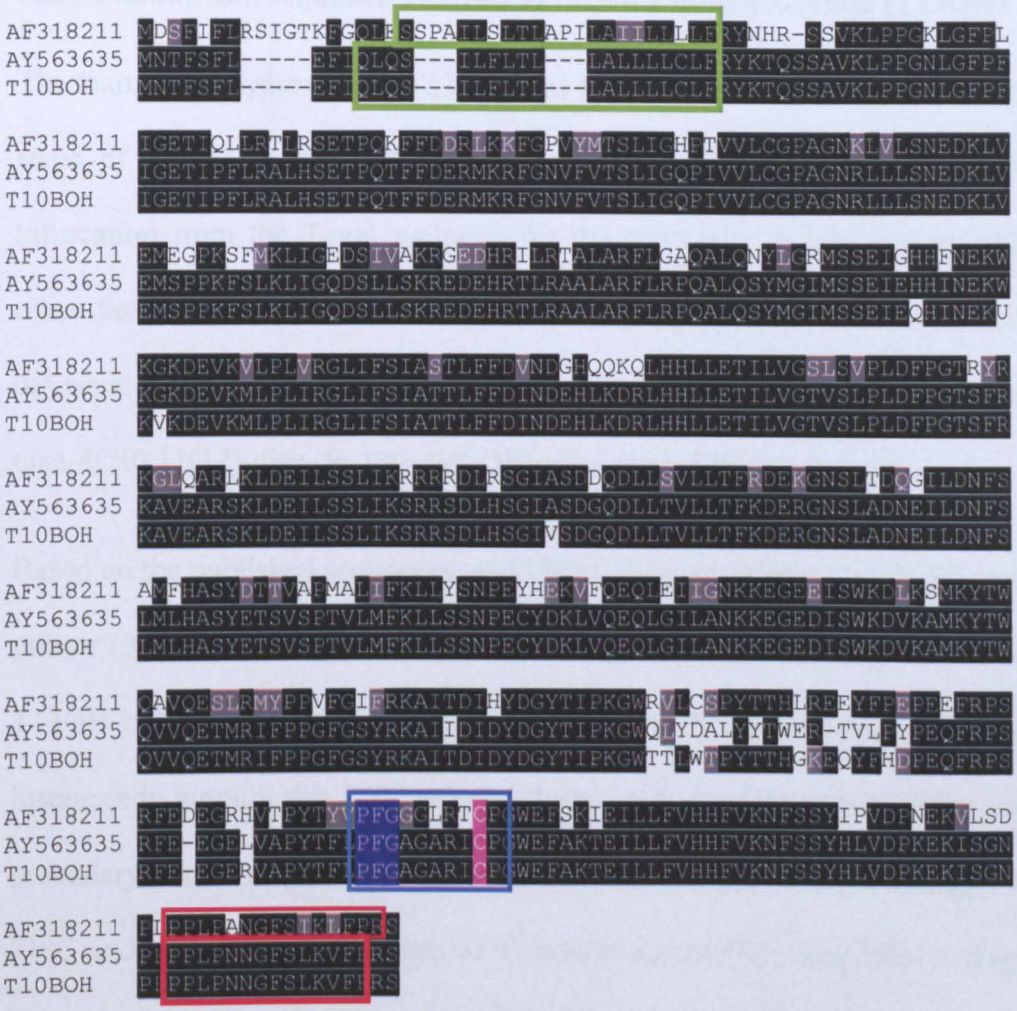


Figure 3.9 Amino acid sequence alignment of T10BOH to its *T. cuspidata* homologues

The sequence alignment of the T10BOH protein and its homologues was aligned by the ClustalW program and produced by "Protein Boxshade" program. The transmembrane anchor helices found at amino acid 11-28 and 289-306 are indicated by the green and red rectangles respectively. The CYP cysteine heme-iron ligand domain is represented by the blue rectangle. Colour code: X = Similar, X = Conservative amino acid substitution, X = Non-conservative amino acid substitution, X = Conserved cysteine residue within the heme-binding region, X = PFG domain. *T. cuspidata* AY563635 and *T. cuspidata* AF31811

3.2.5 Cloning and sequence analysis of taxoid 13α-hydroxylase (T13OH)

The taxoid 13α-hydroxylase (T13OH) has been proposed to be involved in the early steps of the Taxol biosynthetic pathway and is also responsible for the early bifurcation from the Taxol pathway for the synthesis of dead-end metabolites or other taxanes. This enzyme demonstrated a preference for the conversion of the oxygenated compound taxadiene-5α-ol to form a more polar compound identified as taxa-4(20),11(12)-dien-5α,13α-diol (Wheeler *et al.*, 2000).

Based on the published sequences of T13OH, forward primer 13aOH-SP and reverse primer 13aOH-ASP (Figure 3.10A) were designed and used to amplify the full length T13OH cDNA with an *attB* tag at each end. PCR amplification failed on many instances to amplify this CYP enzyme from the *T. baccata* cDNA giving unspecific secondary bands. A gradient PCR was done using the temperature range of 51°C - 71°C and annealing temperature 61°C which successfully amplified a fragment of size ~1400 bp (Figure 3.10B). The attB1 and attB2 adapter sites were then attached to the PCR product and the final cDNA was cloned into a pDonor221 entry vector (Figure 3.10A) by a BP cloning reaction (section 2.12.2). The entry vector thus named T13OH::pDonor was transformed into *E. coli* strain DH5α and the presence of the insert in the colonies resistant to kanamycin was verified by PCR using internal primers 13aOH-MidF and 13aOH-MidR (Appendix I). An amplified band of 179 bp was obtained in only one colony (Figure 3.10 C). The plasmid from this colony was sequenced using the vector specific primers M13F and M13R. A BLAST search of T13OH from *T. baccata* indicated a 99% similarity to T13OH from *T. cuspidata* (AY056019) and *T. media* (AAX20147) and a 97% similarity to *T. wallichiana var. chinensis* (AAX59903) (Figure 3.11).

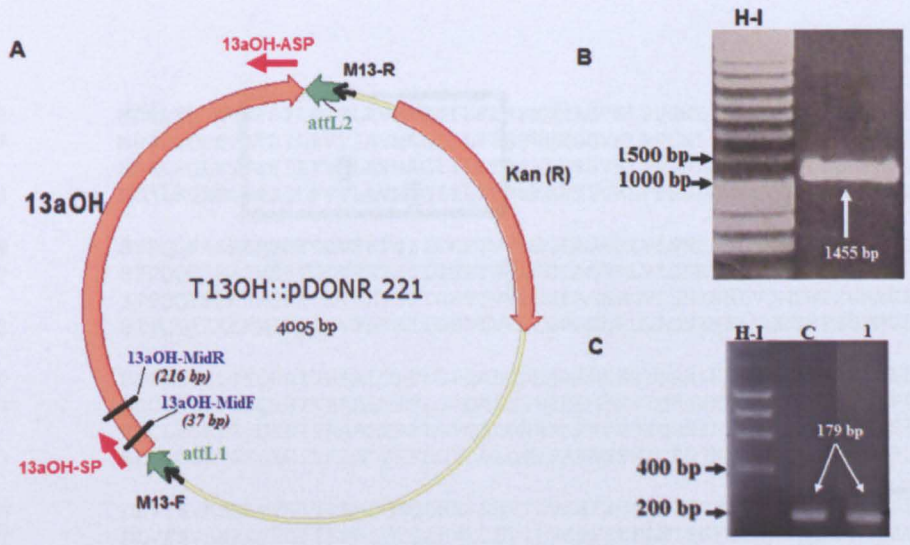


Figure 3.10 Entry clone pDonor221::T13OH

(A) Entry vector pDonor221::T13OH consisting of full length T13OH cDNA cloned between the *attL*1 and *attL*2 sites. The positions of the full-length primers 13aOH-SP and 13aOH-ASP and the internal primers 13aOH-MidF and 13aOH-MidR are indicated. (B) Gel picture of PCR amplification of the T13OH cDNA from *T. baccata* cDNA pool using primers 13aOH-SP and 13aOH-ASP amplified a fragment of size ~1400 bp. (C) Gel picture of PCR amplification using the internal primers showing the presence of the amplified fragment of ~180 bp in colony 1.

Amino acid changes from Asparagine (N) to Threonine (T) at position 17, and from Valine (V) to Alanine (A) at position 173, are present in the *T. baccata* T13OH sequence, and might alter the protein folding and activity of this enzyme. Sequence analysis by "Predictprotein" software indicated that T13OH contains two transmembrane domains at amino acids 10-28 and 290-303, consistent with the activity of the enzyme being associated with the ER. Sequence analysis also revealed other characteristic sequence elements of CYP enzymes, including a highly conserved PFG motif, a heme-iron ligand FGGGMRVCPG at amino acids 423-433 and the essential cysteine at position 430 for enzyme binding and activity.

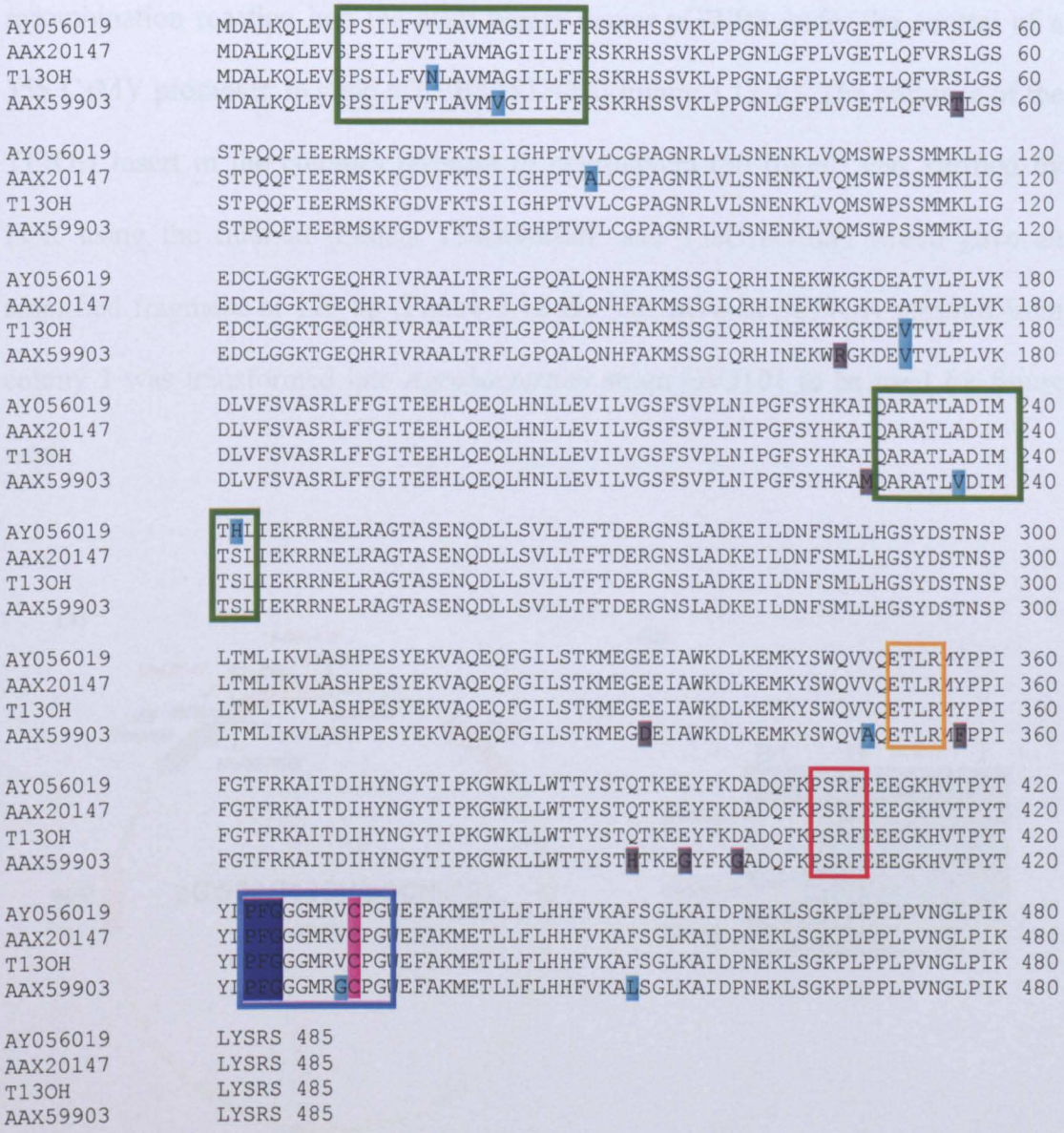


Figure 3.11 Amino acid sequence comparison of T13OH and its *Taxus* homologues.

The sequence alignment of the T13OH protein and its homologues was aligned by the ClustalW program and produced by "Protein Boxshade" program. The transmembrane anchor helices found at amino acid 10-28 and 290-303 are indicated by the green rectangles. The CYP cysteine heme-iron ligand located between amino acids 422-433 is indicated by a blue rectangle. The CYP conserved domain PSRF and the Glu-X-X-Arg motif are indicated by the red and yellow boxes respectively. *T. cuspidata* (AY056019), *Taxus x media* (AAX20147), *Taxus wallichiana var. chinensis* (AAX59903). Colour code: X = Similar, X = Conservative amino acid substitution, X = Non-conservative amino acid substitution, X = putative PFG domain, X = conserved cysteine residue within the heme-binding region.

Following confirmatory sequencing, the T13OH was cloned via an LR recombination reaction into the plant binary vector pGWB8 under the control of a 35S CaMV promoter, to yield pGWB8::T13OH (Figure 3.12 A). The presence of the T13OH insert in the colonies resistant to hygromycin (50 µg/ml) was verified by PCR using the internal primers 13aOH-MidF and 13aOH-MidR which gave an amplified fragment of 179 bp (Figure 3.12 B). The pGWB8::T13OH plasmid from colony 1 was transformed into *Agrobacterium* strain GV3101 to be used for future work.

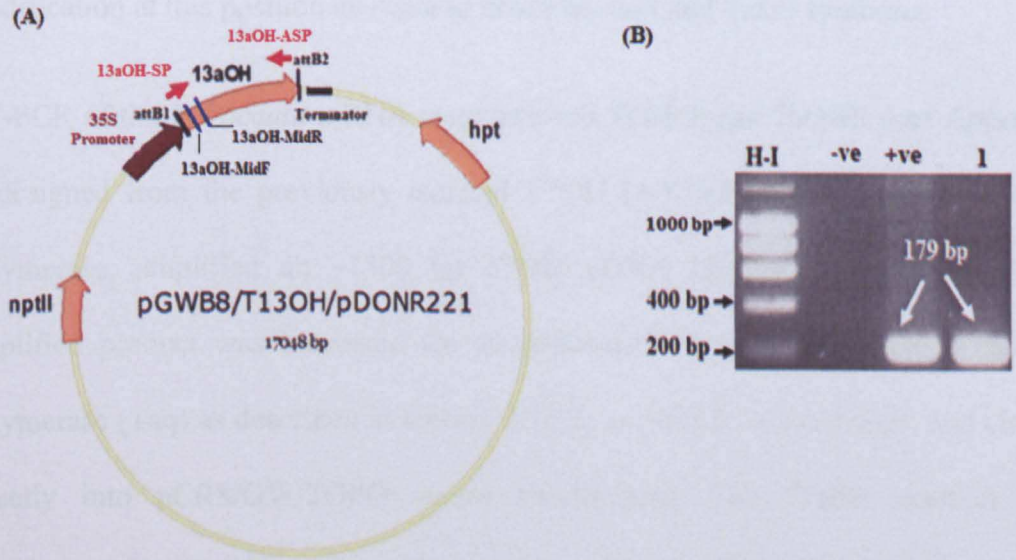


Figure 3.12 T13OH::pGWB8 expression vector

(A) Expression vector pGWB8::T13OH consisting of full length T13OH cDNA driven by the constitutive 35S promoter. The vector carries the nptII and hpt genes conferring resistance against kanamycin and hygromycin respectively. (B) Gel pictures of PCR amplification of plasmids from *E.coli* colonies with internal primers 13aOH-MidF and 13aOH-MidR, showing the presence of a band size 179 bp in both the positive control and colony 1, indicating the transfer of the T13OH cDNA to the binary vector in this colony.

3.2.6 Cloning and sequence analysis of taxoid 7β-hydroxylase (T7OH)

The taxoid 7β-hydroxylase (T7OH) was demonstrated to possess a higher binding affinity to polyoxygenated and acylated taxoids bearing a 5α-acetoxy group, hence placing this 7β taxoid in the midsection of the Taxol pathway. Although the 7β-hydroxyl group has been demonstrated not to be essential for antimitotic activity, this functional group is the most reactive and accessible hydroxyl of the taxane core, and has been modified for the synthesis of taxanes with improved activity. Hence, T7OH catalysing the hydroxylation of the C7 position of the taxane could prove useful for the synthesis of 7β-hydroxy taxoids as starting materials for the subsequent modification at this position *en route* to novel taxanes and Taxol synthesis.

RT-PCR of the *T. baccata* mRNA using primers 7bOHF and 7bOHR (see Appendix I) designed from the previously isolated T7OH (AY307951) and a proof-reading polymerase, amplified an ~1500 bp T7OH cDNA (Figure 3.13 B). The PCR amplified product was incubated for an additional 1hr with a non-proof reading polymerase (Taq) as described in section 2.11.3, to provide A-overhangs, and cloned directly into pCR8/GW/TOPO vector (Invitrogen). The TOPO reaction was transformed into *E.coli* and the presence of the cDNA insert in the colonies resistant to spectinomycin was checked by PCR using the same full-length primers as above (Figure 3.13). The plasmids of colonies 1, 2 and 4 thus named T7OH::pCR8 carrying the T7OH were sequenced using the pCR8 specific primers GW1 and GW2.

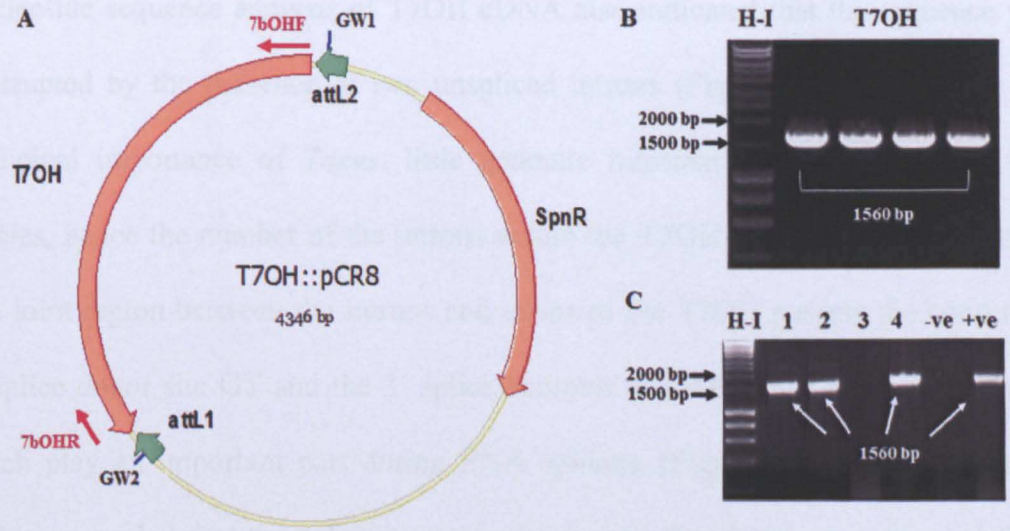


Figure 3.13 Entry vector T7OH::pCR8

(A) Schematic diagram of the entry vector T7OH::pCR8 carrying the SpnR gene and the T7OH cDNA between the attL1 and attL2 insertion sites. The primers 7bOHF and 7bOHR are indicated which were used to amplify the whole length of the T7OH cDNA. The GW1 and GW2 primers indicated were used for sequencing. The diagram also indicates the attL1 and attL2 sites where the T7OH cDNA was inserted (B) Gel picture of PCR amplification of T7OH fragment using 7bOHF and 7bOHR primers shows the presence of a band size 1560 bp (C) Gel picture of PCR amplification of the T7OH::PCR8 transgene from *E.coli* colonies 1 to 4 using the full-length primers 7bOHF and 7bOHR. –ve = water as control, +ve = T7OH fragment as positive control.

Sequence comparison of the T7OH cloned from *T. baccata* and its *T. cuspidata* AY307951 homologue indicated that there has been a nucleotide change from G to A at position 488 (Figure 3.14 A). This resulted in an amino acid change from Tryptophan (TGG) to a premature STOP codon (TAG) and hence would result into a truncated protein. Site directed mutagenesis by using an oligonucleotide containing the correct sequence TGG instead of TAG could be used as part of the future work to amplify the gene, hence omitting the stop codon.

Nucleotide sequence analysis of T7OH cDNA also indicated that the sequence was interrupted by the presence of two unspliced introns (Figure 3.14 B). Despite the medicinal importance of *Taxus*, little genomic information is available for yew species, hence the number of the introns within the T7OH could not be ascertained. The joint region between the introns and exons of the T7OH possess the conserved 5′ splice donor site GT and the 3′ splice acceptor site AG at the end of the intron, which play an important part during RNA splicing (Figure 3.14 B, C). Sequence analysis revealed that the splice acceptor site for the first intron was changed from AGG to AGA (indicated by the arrow in Figure 3.14 B) which could explain the failure of first intron splicing.

Following sequence analysis, the T7OH nucleotide sequence was translated using the ExPASy translate tool (http://expasy.org/tools/dna.html) which omitted the introns within the T7OH sequence. The amino acid sequence was aligned to the previously isolated T7OH *Taxus* homologues using ClustalW sequence alignment program and showed 99% homology to the *T. cuspidata* AY307951 and 97% identity to *T. wallichiana var chinensis* AAR21106 (Figure 3.15).

Analysis of the deduced T7OH sequence by "PredictProtein" program revealed several features characteristic of CYP enzymes, such as an N-terminal membrane anchor at position 23-43, the highly conserved heme binding motif with PFG element (aa 438-440) and the essential cysteine at position 446 (Figure 3.15).

(A)

```
T7OH          ATACTCTGCGGCCCTGCGGGAAACCGCTTAGTTCTTTCCAACGAGGAAAAACTGTTGCAC 480
AY307951      ATACTCTGCGGCCCTGCGGGAAACCGCTTAGTTCTTTCCAACGAGGAAAAACTGTTGCAC 395

                    Ter     Lys(K)
T7OH          GTGTCGTAGTCCGCCAAAATTGCCAGAATCCTGGGTCTCAATTCTGTTGCAGTGAAAAGG 540
AY307951      GTGTCGTGGTCCGCCCAATTGCCAGAATCCTGGGTCTCAATTCTGTTGCAGTGAAAAGG 455
                    Trp     Gln(Q)
```

(B)

```
T7OH          TCTTCCAATCCAGAATGCTTTGAAAAAGTAGTTCAAGGTATGCAACGCTTCTGCTCTGTT 1140
AY307951      TCTTCCAATCCAGAATGCTTTGAAAAAGTAGTTCAAG---------------------- 1032

T7OH          TTCTTGATTTCTCGATGTTTGTGCTAGGCGCAATGTAATTTGTGTGGATACATTGATATA 1200
AY307951      ------------------------------------------------------------

T7OH          ATATTACAGAGCAATTGGAGATAGCTTCAAATAAAAAGGAGGGAGAAGAAATCACAATGA 1260
AY307951      ---------AGCAATTGGAGATAGCGTCAAATAAAAAGGAGGGAGAAGAAATCACAATGA 1083

T7OH          AGGATATCAAAGCCATGAAATACACATGGCAAGTGCTCCAGGAAAGTCTACGGATGCTTT 1320
AY307951      AGGATATCAAAGCCATGAAATACACATGGCAAGTGCTCCAGGAAAGTCTACGGATGCTTT 1143

T7OH          CTCCAGTATTTGGAACACTTCGTAAGACCATGAATGACATTAATCATGATGGTTACACAA 1380
AY307951      CTCCAGTATTTGGAACACTTCGTAAGACCATGAATGACATTAATCACGATGGTTACACAA 1203

T7OH          TTCCAAAAGGATGGCAGGTAAATCACTCCATACTTTTCATATGTCAATTTATGTTTTTCC 1440
AY307951      TTCCAAAAGGATGGCAG---------------------------------------- 1220

T7OH          TTCTGAGATTCAATATATGAAATCCTGTTTTCCCCTTTGCACAATTTGTTACAGGTTGTA 1500
AY307951      ------------------------------------------------GTTGTA 1226
```

(C)

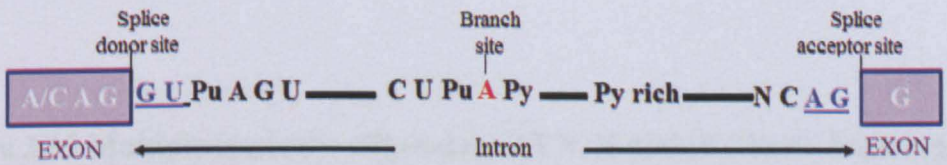


Figure 3.14 Nucleotide sequence comparison of T7OH from *T. baccata* and its T. cuspidata homologue.

(A) Sequence alignment indicating the nucleotide change in the coding region of the T7OH gene at position 488 from TGG to TAG to give the stop codon (*Ter*). (B) The nucleotide sequence alignments indicating the presence of the introns, the splice donor (GT) and acceptor (AG) sites required for splicing. The black arrow indicates the expected position of a Guanine nucleotide required for splicing. (C) The consensus sequence for splicing. The 5′ splice site includes the GU at the 5′ end of the intron and the 3′ splice site terminated the intron with an AG sequence, upstream of which is found the pyrimidine rich site including the conserved adenine nucleotide important for splicing. Pu= A or G; Py = C or U. Colour code: X = similar, X = non-similar nucleotide, X = consensus features of plant introns.

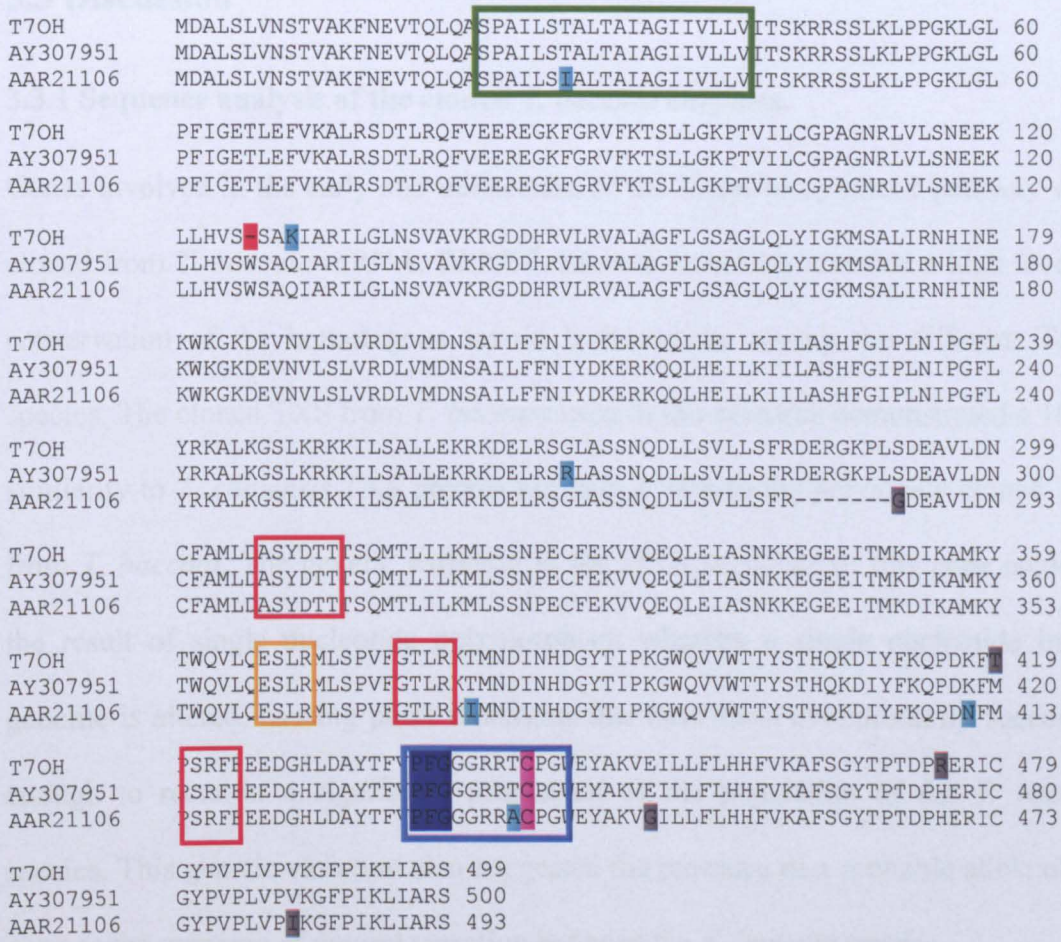


Figure 3.15 Multiple sequence alignment of T7OH and its *Taxus* homologues.

The sequence alignment of the T7OH protein and its homologues was aligned by the ClustalW program and produced by "Protein Boxshade" program. The transmembrane anchor helix found at amino acid 23-43 is indicated by the green rectangle. The CYP cysteine heme-iron ligand domain is represented by the blue rectangle and the Glu-X-X-Arg is represented by the yellow box. The other CYP conserved domains are indicated by the red rectangles. *T. cuspidata* AY307951, *T. wallichiana var chinensis* AAR21106. Colour code: X = Similar, X = Conservative amino acid substitution, X = Non-conservative amino acid substitution, X = Conserved cysteine residue, X = PFG domain, X = STOP codon

3.3 Discussion

3.3.1 Sequence analysis of the cloned *T. baccata* enzymes.

Genes involved in the early and midsection of the Taxol biosynthetic pathway were cloned from *T. baccata* mRNA. BLAST database searching revealed a high level of conservation of the homologous taxoid hydroxylases among the different *Taxus* species. The cloned TXS from *T. baccata* used in this research demonstrated a 100% similarity to *T. chinensis* TXS protein but only a 98% to the previously cloned TXS from *T. baccata.* The genetic variation in the DNA sequence of this gene could be the result of single nucleotide polymorphism whereby a single nucleotide in the genome is altered, causing point mutations that have been evolutionarily successful enough to recur in a significant proportion of the population of the *T. baccata* species. This genetic variation also suggested the presence of a probable allele of the gene or the outcome of natural variation between the *T. baccata* species.

The analysis of the deduced amino acid sequence of the hydroxylases cloned during this project revealed several typical characteristics of cytochrome P450 monooxygenases, including an N-terminal hydrophobic helix required for anchoring in the ER membrane. The highest structural conservation among CYPs is found around the heme domain and comprise the heme-binding loop, containing the most characteristic CYP heme-binding structure (Phe-X-X-Gly-X-Arg-A-Cys-X-Gly or FXXGXXXCXG), with the conserved cysteine residue that serves as a ligand to the heme-iron (Hasemann *et al.,* 1995). The region around the cysteine of the heme-binding domain is almost identical in most of the aligned hydroxylases sequences. The Glu and Arg of the E-X-X-R motif which has been proposed to be required for

the stabilisation of the core structure of the enzyme and the Arg (R) of the PXRF motif may altogether form an E-R-R triad salt bridge that locks the cysteine pocket in position and assures heme association with the protein (Hasemann *et al.,* 1995). The conserved domain ASYDTT which contains a highly conserved Threonine that is thought to be involved in binding and activation of the enzyme was also observed in the oxygenases T5OH and T7OH cloned from *T. baccata.*

Based on relative abundances of hydroxylated taxanes in *Taxus* species, almost all known taxoids are oxygenated at the C-5 positions (Baloglu and Kingston, 1999; Shi and Kiyota, 2005) suggesting the C-5 oxygenation to be the first hydroxylation of the taxadiene core. The enzyme T5OH catalysing this oxygenation reaction was cloned from *T. baccata* and nucleotide sequence analysis revealed this clone to have a nucleotide change from TTG to TCG at position 616 which resulted into an amino acid substitution from Leucine to Serine. This mutation was corrected and the final sequence was confirmed by sequencing to share a 100% and 98% identity to the *T. cuspidata* and *T. chinensis* protein sequences respectively. The T5OH used in this work has a slight nucleotide change in the C-terminus of the protein sequence which generated an alternative stop codon differing by five amino acids from the originally cloned *T. cuspidata* T5OH (AY289209.1) . Instead of the IKLFPETIVN C-terminus previously described for T5OH, the CYP-T5OH used in this case has an amino acid sequence IKLFPRP, more closely related to that of other CYP oxygenases from the yew tree belonging to the CYP725 family. The minor differences observed are likely attributable to the species differences, to that the putative protein encoded by T5OH

should have the same function as that of the previously cloned taxadiene-5α-hydroxylases.

Nucleotide sequence analysis of T7OH cloned from *T. baccata* indicated the presence of two introns which failed to splice out. The sequence analysis of the splice acceptor site at nucleotide position 1209 indicated the absence of the G-nucleotide which is required at the 5′-start point of the exon to aid the splicing process. The T7OH arisen by gene duplication could have subsequently acquired mutations at the splice site that might cause it to be non-functional and results into the formation of a non-processed pseudogene. The second intron possesses the correct splice acceptor site GAGG, hence this intron would have been spliced out during post-translational modification in plants. The presence of a functional protein could thus be verified by a western blot analysis to check the correct size of the protein formed. The premature stop codon at amino acid position 488 might have been the result of a PCR bias or the amplification of a pseudogene that shares similar sequences as the wild type T7OH. This sequenced cDNA possessed the same CYP features as the wild type one from *T. cuspidata* and the loss of the pseudogene's functionality probably has no effect on the *T. baccata*'s capability of producing 7β-taxoids or Taxol, since an intact functional copy of the gene would still exist in the yew. To circumvent this premature stop codon issue, the cloned T7OH cDNA can be site directed mutagenated by using an oligonucleotide containing the correct sequence TGG instead of TAG, hence omitting the stop codon. In order to remove the unspliced introns, primers can be designed having appropriate restriction sites,

which would amplify the sequence omitting the introns and the premature stop codon.

The enzyme T5AT catalysing the first acetyltransferase reaction of the Taxol biosynthetic pathway was successfully cloned from *T. baccata* mRNA. Amino acid sequence analysis of this cloned T5AT indicated a few non-conservative amino acid substitutions when compared to the published sequences of the *Taxus* T5AT homologues. These changes in amino acids substitution might be the result of natural variation among the species and might not cause a conformational change in the protein structure and function of the acetyl transferase. BLAST database search indicated the *T. baccata* T5AT to be more closely related to that of *T x media* and *T. chinensis* than to the T5AT from *T. cuspidata*, contradictory to the fact that *T x media* is known to be the hybrid arisen from *T. cuspidata* and *T. baccata*. However, the four *Taxus* T5AT homologues showed high percentage of identity among themselves, suggesting that the transferase shares a common evolutionary origin based on their similar roles and conserved structural and sequence characteristics, such as the conserved HXXXDG domains. This conserved motif has been suggested by St-Pierre *et al.*, (1998) to lie at or close to the active centre of the acyltransferases where the histidine residue of this element is essential for catalytic activity of the enzyme for acyl group transfer from acyl CoA to the substrate alcohol.

3.3.2 Evolutionary relationships of the enzymes of the Taxol biosynthetic pathway

A phylogram (Figure 3.16) of the taxoid oxygenases cloned from *T. baccata* and the *Taxus* 14-β hydroxylase and 2α-hydroxylase previously isolated, was constructed by using "ClustalW alignment program" based on the amino acid sequence alignment. The pattern shows the early emergence of C14 hydroxylation suggesting a side route after the C5 hydroxylation. This data correlates with previous reports that taxoid 14β-hydroxylase is involved in the bifurcation of the Taxol biosynthetic pathway after the C5 hydroxylation and acetylation, leading to the formation of dead-end metabolites in *Taxus* (Ketchum *et al.*, 2003; 2007a). Interestingly, the T5OH from *T. baccata* which oxygenates the C ring of the taxane core displays the highest similarity to this 14*β*-hydroxylase, which oxygenates the taxane A-ring.

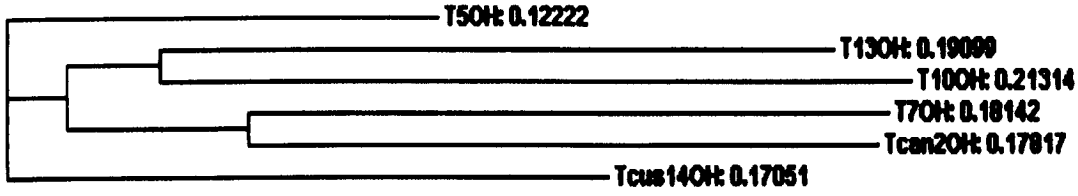


Figure 3.16 Phylogram for the cloned taxoid hydroxylases from *T. baccata* and other *Taxus* hydroxylases.

The phylogram indicates the early emergence of the Tcus14OH after C5 hydroxylation. T7OH and Tcan2OH are clustered next to each other, indicating the high sequence similarity between these two hydroxylases. T13OH appears to be more closely related to T5OH than it is to T10BOH. *T. canadensis* 2α-hydroxylase (Tcan2OH) (AY518383), *T. cuspidata* 14β-hydroxylase (Tcus14OH) (AY188177).

T10BOH and T13OH were found to be closely related in the phylogram constructed, which might support the previous report of their overlapping substrate selectivity (Wheeler *et al.*, 2001). "ClustalW alignment program" indicated that T5OH displayed a higher similarity to the T13OH (62%) than to T10BOH (61%). This could explain the preferential hydroxylation of the taxadiene-5α-ol by T13OH previously reported by Wheeler and co-workers (2001), than to the expected T10BOH which has been shown to catalyse the fourth step of the Taxol biosynthetic pathway. The 2*α*- and 7*β*-hydroxylases which hydroxylate highly functionalised taxoids were found to be more similar to each other (64% identity) than to the early pathway hydroxylases. These observations indicate the near impossibility of predicting hydroxylation regiochemistry based solely on homology.

3.3.3 Evolutionary relationships of the *T. baccata* hydroxylases and other plant CYPs.

Sequence and structure analysis have indicated that all the CYP genes diverge from only one evolutionary origin and can be depicted in a single phylogenetic tree (Nelson, 1999). The clones isolated during this research were subjected to extensive analysis and their sequences compared pairwise to the known plant CYP oxygenases sequences available in the NCBI database (http://www.ncbi.nlm.nih.gov/) to provide a cladogram of these relationships. BLAST database searches indicated that the CYP oxygenases of the Taxol pathway show high sequence similarity with each other (~60%) but distant relationships (<50%) to other plant CYPs (Figure 3.17). Thus, the genes encoding the respective sequential oxygenation steps of the Taxol pathway

almost certainly arose by gene duplication and differentiation of an ancestral gene, which is mostly likely to be the T5OH gene.

The sequences from all the CYPs shown in Figure 3.17 were aligned using the "Clustal W1.2 Multiple sequence alignment program". The percentage similarities between the *Taxus* hydroxylases cloned and the other plant CYPs was generated using the Clustaw program. It was shown from the sequence alignment that the *Taxus* clones were related to the CYP725 family and resembled most closely the CYP707 (28-29%), CYP90 (23-30%), CYP85 (24-26%), and CYP88 (23-28%) families which are known to be involved in the synthesis of other large terpenoids (e.g. gibberellins and steroids). The *Taxus* clones were also found to resemble CYP716 (33-44%) which have unknown function, however the other closest gene to the *T. baccata* clones outside the CYP725 family was identified in *Arabidopsis* as ArthCYP718, which is as yet uncharacterised.

The CYP707 family sharing a 28-29% similarity to the cloned *T. baccata* hydroxylases consists of the hydroxylases involved in abscisic acid (ABA) biosynthesis. The CYP85 and CYP90 families sharing a 23-30% similarity to the taxoid hydroxylases, have been shown to catalyse the early steps of the growth-promoting plant steroids brassinosteroid (BR) biosynthetic pathway from the triterpenoid campesterol to brassinoide (Kim *et al.*, 2005; Clouse and Sasse 1998). The early three-step oxidation of *ent*-kaurene to *ent*-kaurenoic acid involved in gibberellins (GA) biosynthesis and the late analogous oxidation to GA have previously been demonstrated to be catalysed by the sub-families CYP701A and CYP88A respectively (Helliwell *et al.*, 1999; 2001a). These two CYP families were

observed to split within two clades, raising the possibility that these genes existed independently in a common progenitor of angiosperms and gymnosperms.

The Loblolly pine (*Pinus taeda*) PitdCYP720B1 gene encodes the multifunctional abietadienol/abietadienal oxidase which catalyses at least two oxidation steps in the conversion of abietadiene to abietic acid as part of diterpene resin acid (defence chemicals in conifers) biosynthesis (Ro *et al.*, 2005). The CYP720B1 found in *Picea sitchensis* (Sitka spruce) (PisiCYP720B1) was separated from the PitdCYP720B1 into two clades, with PisiCYP720B1 demonstrating a closer homology to the CYP725 family. However, considering the close distance of the pine and the spruce to the *Taxus* CYP clones, it can be speculated that these genes form a functionally related family, acting on similar substrates or could possibly share a common ancestry. Additional members of the conifer CYP gene family with high similarity to loblolly CYP720B1 and the *Taxus* hydroxylases have emerged in the spruce database (Hamberger and Bohlmann, 2006). These enzymes, involving PisiP450-1 which is yet uncharacterised have been found to share close homology to the *Taxus* hydroxylases, possibly encoding an enzyme involved in the BR or GA biosynthetic pathway in conifers. Furthermore, the high sequence homology between the *Taxus* hydroxylases and the CYP716B *Picea sitchensis* family could be possibly due to both groups belonging to the taxa Coniferae or both families might each act on overlapping substrates, thereby enhancing diterpenoid chemical diversity in conifer defence.

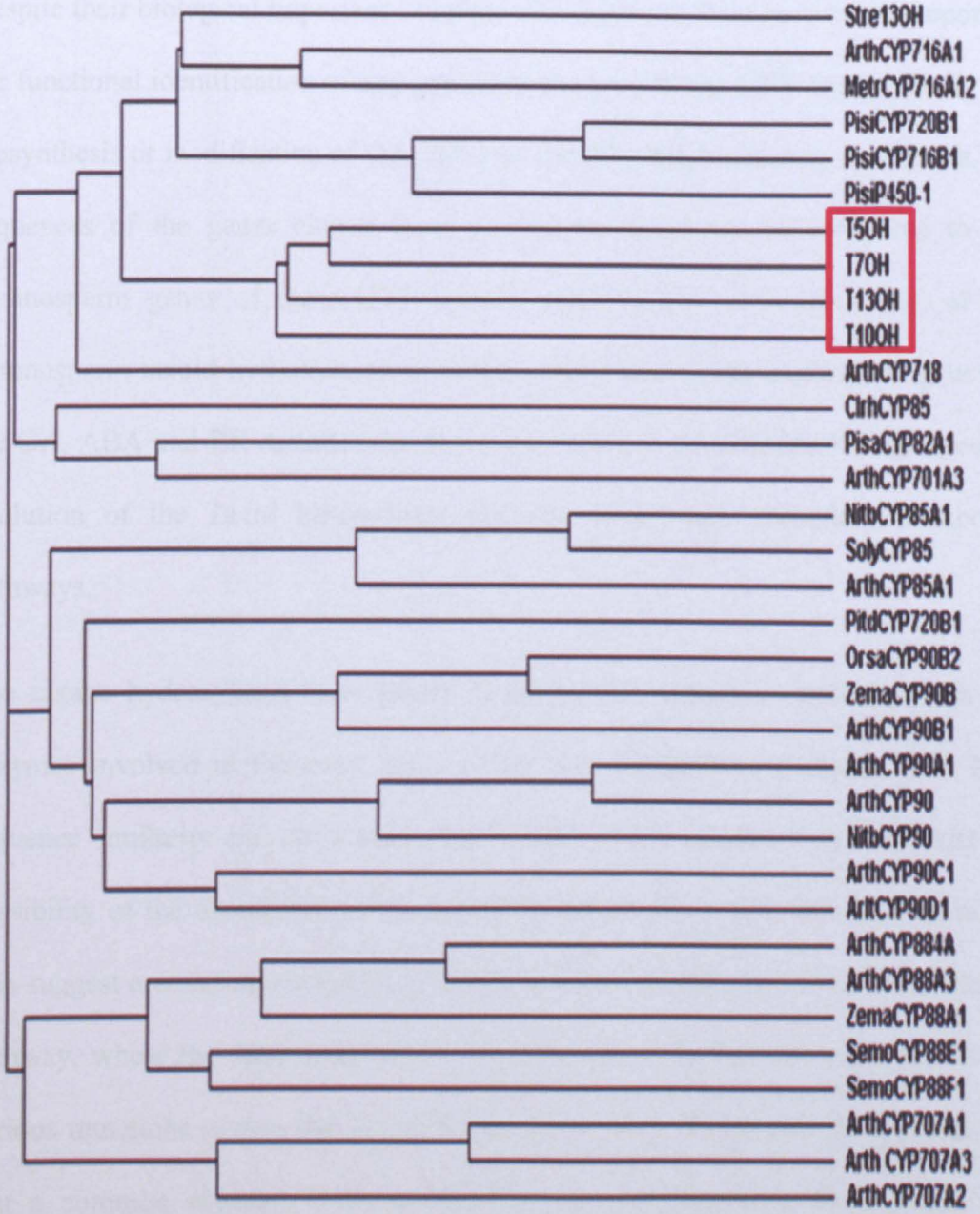


Figure 3.17 Cladogram showing peptide sequence relationships between some published, related plant cytochrome P450s and those cloned from *T. baccata*.

The Taxol biosynthetic genes cloned from *T. baccata* are represented by the red box. For the published sequences, the first four letters of each name are genus and species abbreviations, CYP is the abbreviation or cytochrome P450, the following two numbers indicate the P450 family, and any additional letter or numbers refer to the subfamily. The genus and species abbreviations are as follows: Arth – *Arabidopsis thaliana*, Arlt - *Arabidopsis lyrata*, Clrh - *Chlamydomonas reinhardtii,* Metr - *Medicago truncatula*, Nitb – *Nicotiana tabacum*, Orsa – *Oryza sativa*, Pisa - *Pisum sativum* (pea), Pisi - *Picea sitchensis* (Sitka spruce), Pitd - *Pinus taeda* (loblolly pine), Soly – *Solanum lycopersicum*, Semo - *Selaginella moellendorffii*, Stre13OH - *Stevia rebaudiana* ent-kaurenoic acid 13-hydroxylase, Zema – *Zea mays*

Despite their biological importance in plant development, there has been no report of the functional identification of any gymnosperm gene of the CYP related family for biosynthesis or modification of GA, ABA or the BR phytohormones. As a result, the sequences of the genes cloned from *T. baccata* could not be compared to the gymnosperm genes of these CYP related families. The close clustering of the gymnosperm taxoid hydroxylases (CYP725) with those of the angiosperm genes of the GA, ABA and BR metabolism, however, indicated the possible lineage-specific evolution of the Taxol biosynthetic pathway from other secondary metabolite pathways.

The taxane hydroxylases were found to share high sequence similarity with the enzymes involved in the early steps of the GA biosynthetic pathway. This high sequence similarity and close clustering of these CYP families might indicate the possibility of the divergence of the taxoid family from the GA family. This might also suggest a common evolutionary origin between the GA and Taxol biosynthesis pathway, where the duplicated GA biosynthetic genes in the yew trees underwent various mutations to give the Taxol biosynthetic genes. It can also be hypothesised that a common ancestry CYP involved in GA phytohormone biosynthesis and secondary metabolism prior to the divergence of angiosperms and gymnosperms, evolved to acquire the novel biochemical functions of the taxane hydroxylases in the yew trees.

CHAPTER 4 : LOCALISATION OF THE FIRST THREE ENZYMES OF THE TAXOL BIOSYNTHETIC PATHWAY

4.1 Introduction

The amino acid sequence of TXS catalysing the first committed step of Taxol biosynthetic pathway has indicated that the N-terminal domain of this protein contained a putative chloroplast transit peptide. This roughly 60-80 residue amino-terminal transit sequence has been previously suggested to be cleaved on maturation following plastid import (Williams *et al.*, 2000b). This suggests that this protein is plastid-located, although there has been no report where the localisation of this protein has been experimentally demonstrated.

Sequence analysis of T5OH indicated the presence of a transmembrane helix (Section 3.1.1), suggesting that this enzyme is membrane associated. Although the localisation of the T5OH has not yet been identified, previous reports have documented the enzyme to be restricted to the microsomal fraction prepared from *Taxus* suspension culture cells, further implying this enzyme to be localised to the ER (Jennewein *et al.*, 2004a).

With the cyclisation step of the Taxol biosynthetic pathway being localised to the plastid and the first oxygenation step speculated to take place in the ER, it was difficult to reconcile these assumptions with the putative ER membrane-bound

T5OH. The oxygenation of the taxadiene core could not take place if the taxadiene olefin remained in a different sub-cellular location from the T5OH. Hence, it was reasoned that the non-polar taxadiene most likely partition between the plastid membrane from where it is translocated by the T5OH which may be either spatially distributed between the plastid and the ER or associated with both organelles. As a result, the plastid-associated T5OH would appear to facilitate the movement of taxadiene from the plastid to the ER, as well as catalysing the formation of taxadiene-5α-ol.

As mentioned in the previous chapter, T5OH appears to be closely related to the *A. thaliana* CYP701A3 encoding the *ent*-kaurene oxidase (AtKO1) (Section 3.3.3). This GA biosynthetic enzyme has been shown to catalyse the three-step oxidation of the plastid synthesised *ent*-kaurene to *ent*-kaurenoic acid *en route* to GA synthesis (Helliwell *et al.*, 1999). AtKO1 has been reported to be associated to both the ER and the plastid envelope, thus appears to link the plastid with the subsequent ER-located steps (Helliwell *et al.*, 2001b). Based on the sequence similarity and close relatedness of T5OH and the AtKO1, it was speculated that T5OH might have similar localisation as that of AtKO1. This would suggest that taxadiene might bind to the T5OH associated on the plastid envelope and is transported to the ER where further modifications of the taxane core take place.

The third step *en route* to Taxol has been reported to be catalysed by taxadiene-5α-acetyl transferase (T5AT) (Walker *et al.*, 2000). The authors documented that the sequence of this protein did not encode any N-terminal targeting information, however sequence analysis by the "predictprotein" software has indicated a

transmembrane anchor helix at amino acid 147-164, as described previously (Section 3.2.4). The authors also reported the activity of T5AT to reside in the microsomal extracts, thus agreeing with the prediction of the protein to be located to the ER. Considering the third step of the Taxol biosynthetic pathway to be located in the ER, it could be speculated that the T5OH links the plastidial and the endoplasmic reticulum steps of the Taxol biosynthesis.

The aim of this part of the study was to determine the subcellular localisation of the enzymes involved in the early steps of the Taxol biosynthetic pathway, which might also indicate whether the production of the taxadiene-5α-ol compound is associated with the movement of the taxadiene across the plastid membrane. The key question was how the plastid-located taxadiene interacted with the T5OH if they have different sub-cellular localisations. Confocal microscopy was thus used to localise protein fusions with green fluorescent proteins (GFP), yellow fluorescent protein (YFP) and cyan fluorescent protein (CFP) using stable transgenic plants.

4.1.1 Fluorescent Proteins

The green fluorescent protein (GFP) was first discovered in the jellyfish *Aequorea victoria* as a companion to the bioluminescent protein aequorin (Shimomura *et al.*, 1962). Since then, GFP has been widely used in various types of cells as a reporter gene due to its unique *in vivo* fluorescence ability and low toxicity to the host cell. The use of GFP in higher plants was not feasible until a cryptic intron inside the coding sequence of the GFP was removed (Haseloff *et al.*, 1997). In addition, extensive mutagenesis screens have been carried out and numerous GFP variants

with distinct fluorescence characteristics have been generated. For instance, the S65T (Serine65 changed to Threonine) GFP stabilised the fluorochrome in a permanently ionized form with a single absorbance peak at 489 nm, which became the backbone of the commercially available enhanced GFP (Clontech). The identification of the spectra-shifted GFP variants, such as the blue-shifted cyan fluorescent protein (CFP) and the red-shifted yellow fluorescent protein (YFP) enables multiple proteins to be visualised simultaneously in the same cell (Heim *et al.*, 1994).

4.2 Results

4.2.1 Taxadiene synthase is targeted to the plastids

4.2.1.1 Preparation of Constructs

The protein sequence of TXS, the terpene cyclase of Taxol biosynthesic pathway, was analysed using the ChloroP program which predicted the protein to contain a plastid transit peptide and a cleavable site at amino acid 58. ChloroP uses a neural network method to identify chloroplast transit peptides and their cleavage sites based on a training set of proteins with known subcellular localisations.

In order to identify the localisation of the TXS protein, the TXS cDNA was amplified using a reverse primer designed to exclude the stop codon at the C-terminus of the TXS cDNA. The primers used for TXS-no stop codon (TXSns) amplification were TaxSynF (forward) and TaxSynRNS (reverse) (Appendix I), using the plasmid pBCTaxk carrying the full-length TXS cDNA as template, and a proof-reading enzyme. A 2.6 kb PCR product (Figure 4.1B) was obtained which was gel purified, subsequently cloned in the entry vector pCR8 by TOPO cloning and

transformed into *E. coli* strain DH5α as described previously (Section 2.12.2 and 2.12.3 respectively). The presence of the insert in the *E.coli* colonies which were resistant to spectinomycin selection was verified by PCR using the primers TaxSynF and TaxSynRns. A band of ~ 2600 bp was obtained from colonies 1 and 5 (Figure 4.1C) and the respective plasmids thus named TXSns::pCR8 were purified from the two colonies and sequenced with the primers M13Forward and M13Reverse (Figure 4.1A). The sequencing results confirmed that the stop codon had been deleted at the C-terminus of the TXS cDNA.

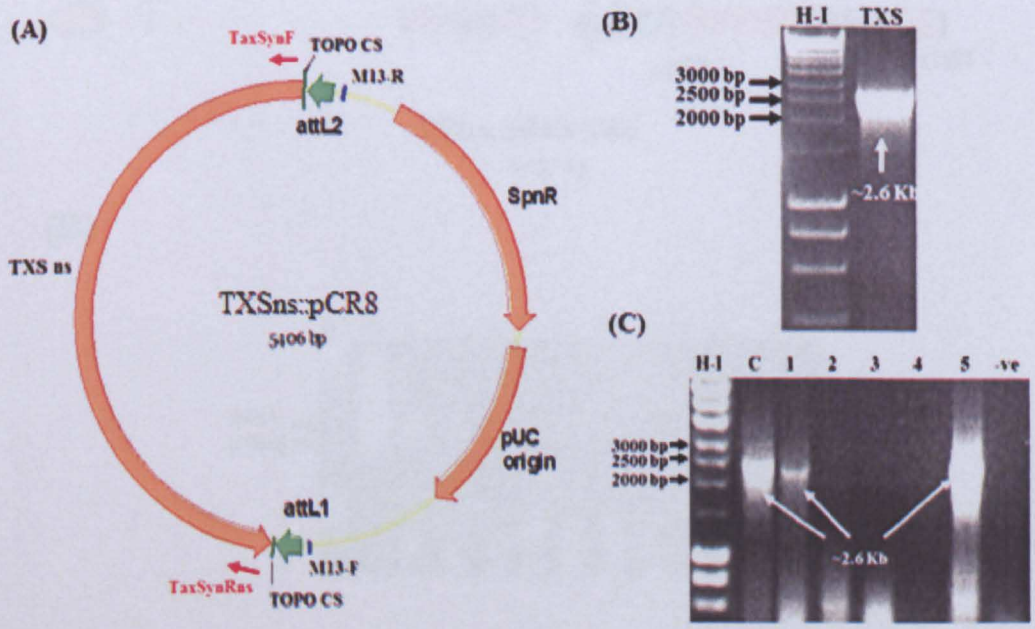


Figure 4.1 Clone pCR8::TXSns

(A) The schematic diagram of the construct generated by TOPO cloning carrying the TXS-ns cDNA between the TOPO cloning sites (TOPO CS). The positions of the primers TaxSynF and TaxSynRns used for PCR amplification of the cDNA are indicated. (B) PCR product showing a band of size ~2.6 Kb indicating the amplification of the TXSns. (C) Colonies 1 and 5 showing the presence of the correct size of TXSns (~2600 bp), where C is the positive control of the fragment prior to cloning and –ve is the negative water control.

The TXSns from plasmid 5 was fused in frame by LR recombination (Section 2.12.2) to the plant binary vector pGKPGWG (Zhong *et al.*, 2008). Following selection on kanamycin, the presence of the TXSns insert was verified by PCR on the colonies resistant to the antibiotic. The presence of the expected 2.6 Kb bands were observed in a few colonies (Figure 4.2 B), and the plasmid thus named TXSns::pGKPGWG (Figure 4.2 A) was purified.

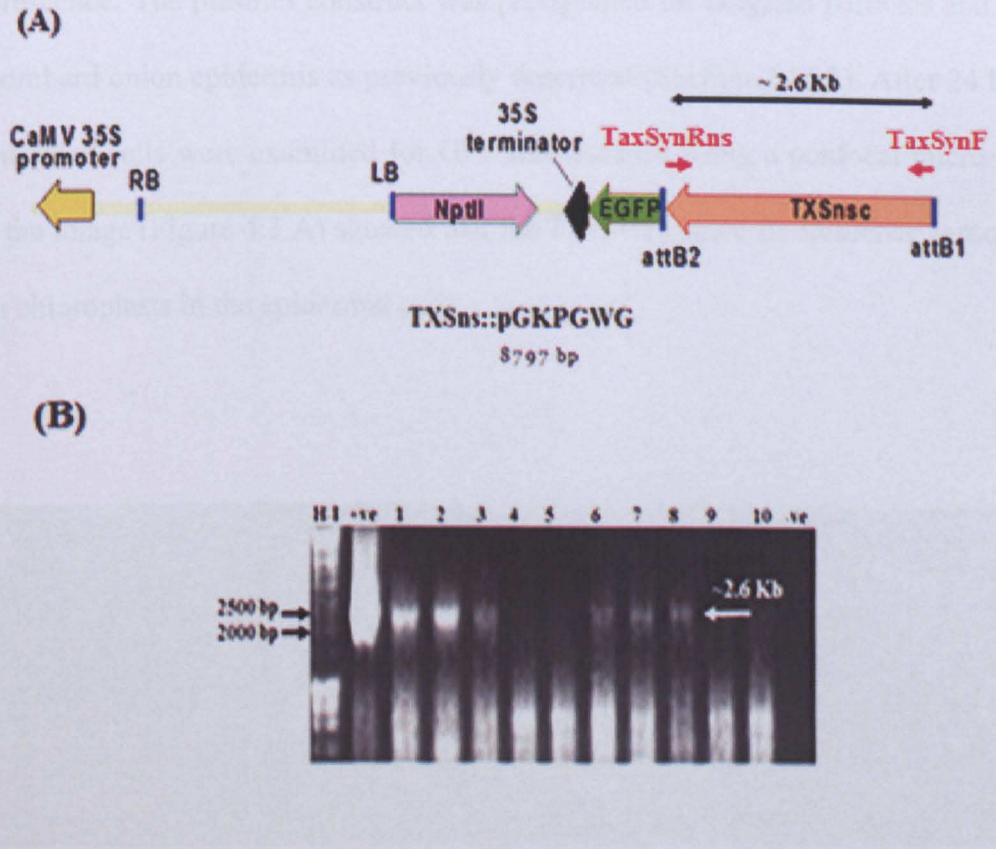


Figure 4.2 TXSns::pGKPGWG expression vector

(A) Schematic diagram of the linear display of the expression vector TXSns::pGKPGWG was generated by software VectorNTI (Invitrogen) and consists of TXSns cDNA driven by the constitutive 35S promoter. The primers TaxSynF and TaxSynRns are indicated which were used to check for the TXSns insert by PCR. NptII; Kanamycin resistance gene, EGFP; enhanced GFP, LB; Left border, RB; Right border. (B) Gel picture of PCR amplification of plasmids from *E.coli* colonies 1 to 10, using primers TaxSynF and TaxSynRns. The presence of a band size ~2600 bp in colonies 1,2,3,6,7 and 8 following LR recombination, indicates the presence of TXSns cDNA in the binary vector.

4.2.1.2 Transient Expression of Green Fluorescent Proteins

Prior to investigating the localisation of TXS in stable transgenic plants, a transient expression system was developed to test the CloroP predictions for subcellular localisation of TXS and also to verify whether the TXSns::pGKPGWG (TXS-GFP) construct was functional. Microprojectile bombardment of onion epidermis was chosen as the method to deliver the test transgene construct and onion epidermal cells were used as the recipient tissue due to their lack of intrinsic chlorophyll auto-fluorescence. The plasmid construct was precipitated on tungsten particles and used to bombard onion epidermis as previously described (Section 2.16.2). After 24 h, the bombarded cells were examined for GFP fluorescence using a confocal microscope and the image (Figure 4.3 A) showed that the TXS-GFP gave fluorescence associated with chloroplasts in the epidermal cells.

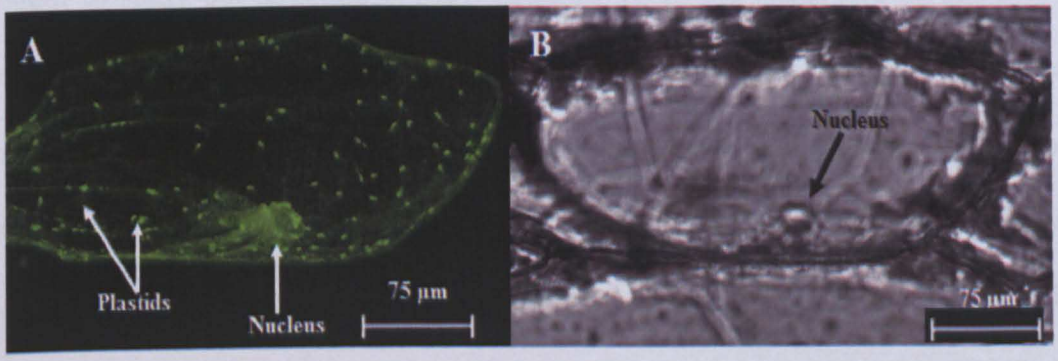


Figure 4.3 Confocal images of a single onion epidermal cell following microprojectile bombardment.

Images were taken 24 h after bombardment of onion epidermal tissue with tungsten particles coated with the TXSns-GFP plasmid construct under the control of the CaMV35S promoter. The cells expressing the GFP construct are epidermal cells and the green fluorescence is mainly found in the plastids which are indicated by the arrows. A) GFP image of the onion epidermal cell layer. B) Superimposed image on the bright-field image showing the epidermal cell image.

4.2.1.3 Sub-cellular localisation of TXS in stable transgenic plants

The TXS-GFP construct which was used for transient expression in onion epidermis, was transformed into *Agrobacterium* strain GV3101 as described previously (Section 2.12.7). One of the colonies which showed the presence of the insert was used for transformation of wild type *Arabidopsis thaliana* as described in section 2.5.1. The seeds obtained following *Agrobacterium* transformation were screened and selected on MSR3 medium containing 50 μg/ml kanamycin. The roots of the 1 week old transgenic lines were mounted on microscopic slides and observed under the confocal microscope. Figure 4.4 shows the expression of the GFP fusion protein mainly localised to the plastids of the transgenic *Arabidopsis* roots. Various parts of the lateral roots and root hairs of the transgenic plants were observed, which gave a consistent plastidial expression of the GFP.

The presence of chloroplast in the epidermal cells of tobacco makes it an ideal system to study the targeting of GFP fused to TXS. The main aim of this area of research was to generate a transgenic tobacco line expressing the three downstream enzymes TXS, T5OH and T5AT fused to a fluorescent protein respectively while using constructs which confer different antibiotic resistance genes. Wild-type tobacco leaves were transformed with the same TXS-GFP construct using *Agrobacterium*-mediated transformation as previously described (Section 2.12.7) and the tobacco transformants were selected on 70 mg/L kanamycin.

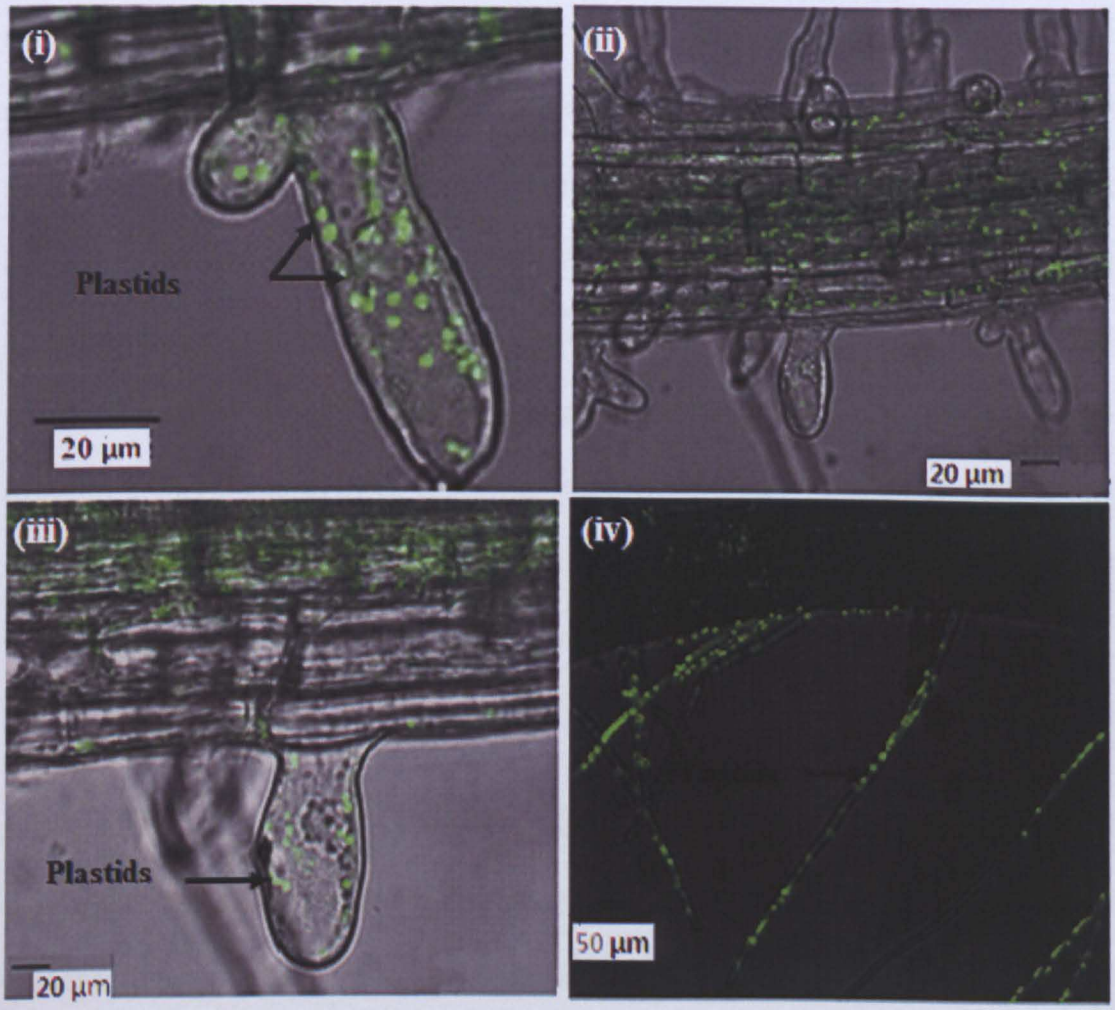


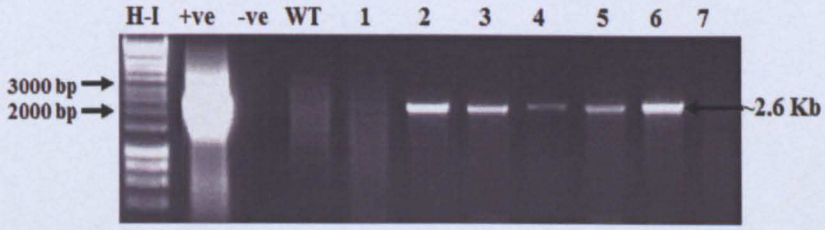
Figure 4.4 Confocal images from transgenic Arabidopsis expressing the TXSns-GFP construct.

(i-iv) show the TXS-GFP expression was predominantly observed in the plastids of the main and lateral roots as well as in the root hairs of the *Arabidopsis* plants. Scale bars of either 20 µm or 50 µm was used.

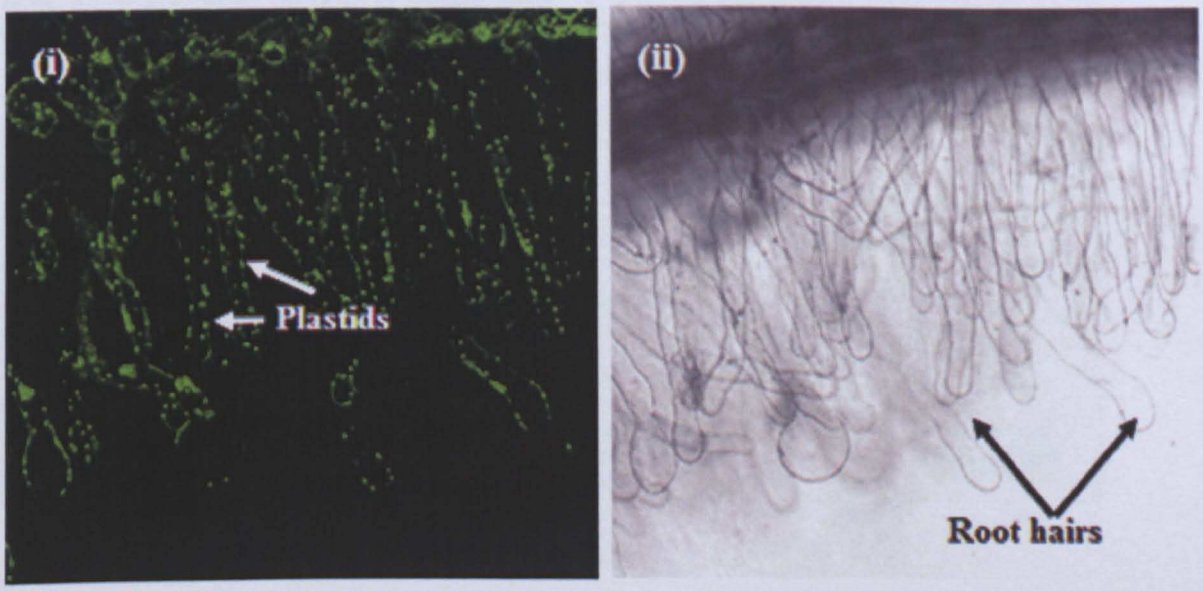
The presence of the TXS-ns insert in the plants which were resistant to kanamycin were verified by PCR using the primers TaxSynF and TaxSynRns to give an amplified fragment of ~2.6 Kb (Figure 4.5 A). The transgenic lines were transferred to the soil and allowed to grow for 3-4 weeks under controlled conditions. The lower surface of the epidermis and the roots of the transgenic plants were mounted on respective microscopic slides. The GFP fluorescence (green) and the chlorophyll auto-fluorescent (red) in the transgenic tobacco leaf epidermal cells were imaged simultaneously by using the confocal scanning microscope as previously mentioned (Section 2.16.1). The TXS-GFP protein was found to be present in plastids in the main root as well as root hairs of the tobacco plant (Figure 4.5B) and was associated with the chloroplasts of the tobacco epidermal cells (Figure 4.5C).

The red auto-fluorescence of the guard cell chloroplasts co-localised with the GFP signal suggested that the TXS was targeted to the chloroplasts. The localisation of the fusion protein in *Arabidopsis* was identical to that in tobacco cells. Thus, evidence from protein localisation using either stable transgenic plants or transient expression indicated that the TXS was targeted to the plastids. These data were in agreement to previous published reports and the software "predictprotein" as well as ChloroP analysis which predicted TXS to be a chloroplast protein.

(A)


H-I
+ve
-ve
WT
1
2
3
4
5
6
7
3000 bp
2000 bp
~2.6 Kb

(B)


(i)
Plastids
(ii)
Root hairs

(C)

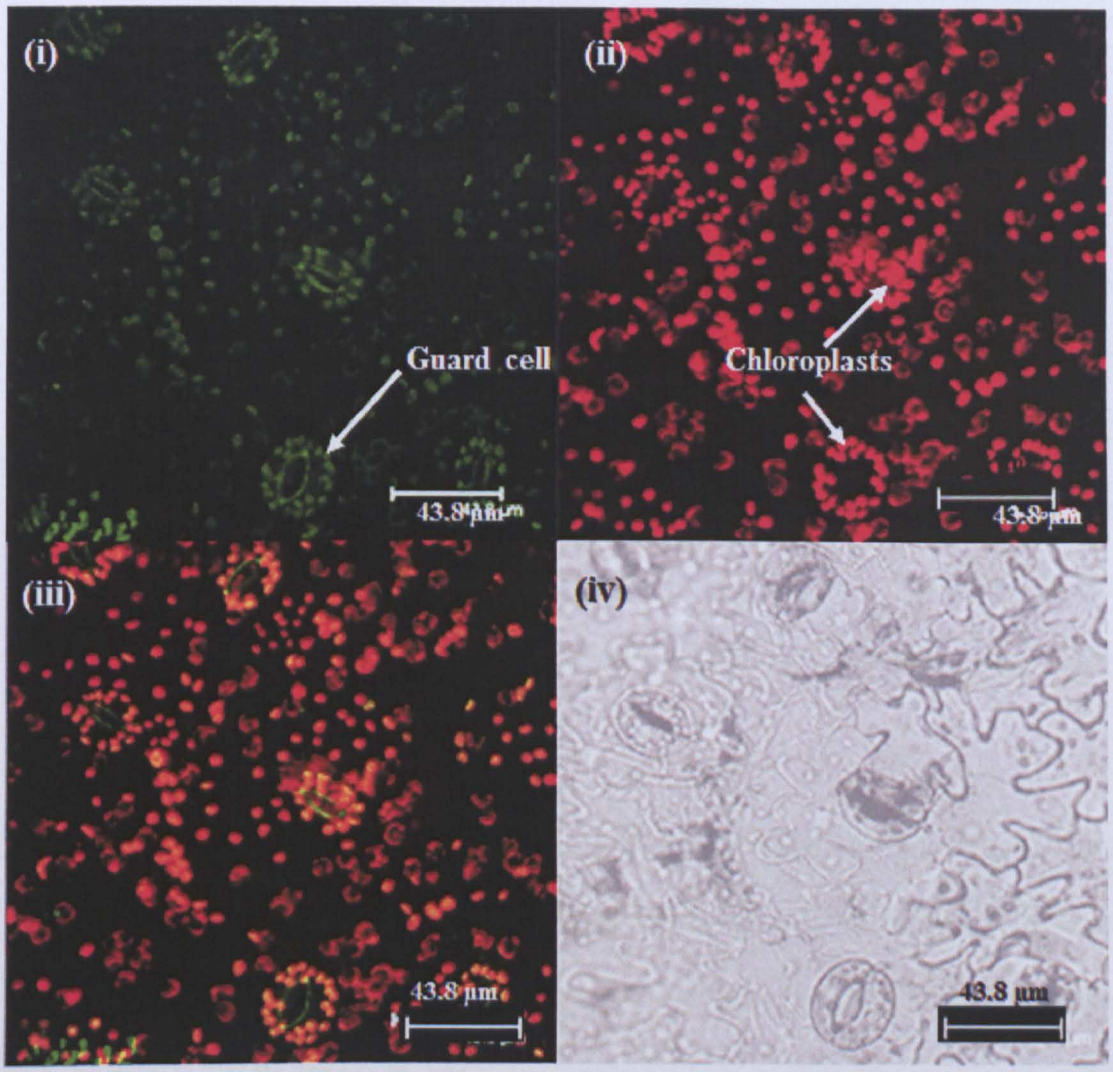


Figure 4.5 Localisation of TXSns-GFP in guard and roots cells

(A) Gel picture of PCR amplification of the TXSns::pGKPGWG transgene from the transgenic plants lines 1-7 using the primers TaxSynF and TaxSynRns. The +ve control used was 1µl of the plasmid TXSns::pGKPGWG and the –ve control was 1µl of water. A WT sample was also used as a negative control which indicates the absence of the 2.6 Kb TXS in WT tobacco plants. (B) The TXS-GFP expression was also predominantly observed in (i) the plastids of the root hairs of the tobacco plants. (ii) shows the phase contrasts of the tobacco roots. (C) The construct gave GFP fluorescence associated with the chloroplasts in the guard cells as well as in those of the surrounding epidermal leaf cells. The images show the separate GFP (i), chlorophyll (ii) channels, together with the merged channels (iii) and the phase contrast of the guard cells (iv). Where the chlorophyll and GFP signals overlap (iv), the image colour is shifted to orange. The chlorophyll signal from underlying mesophyll cells can also be seen; this signal is unavoidable due to the uneven nature of the leaf epidermis and the strength of the chlorophyll signal from the underlying mesophyll layer.

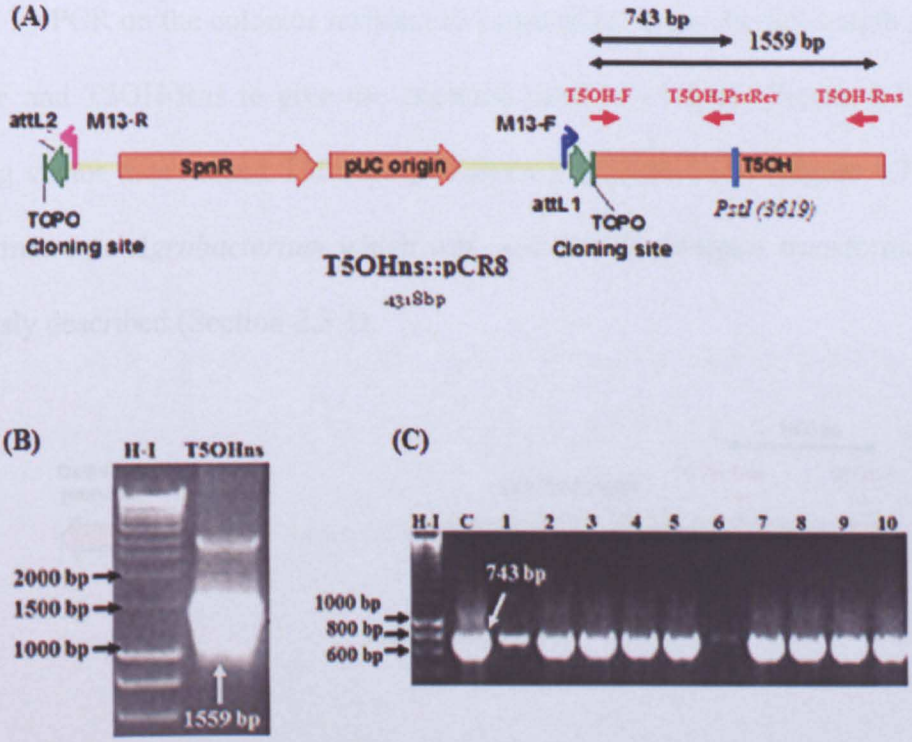
4.2.2 Taxadiene-5α-Hydroxylase is Targeted to the Plastid Envelope and ER

4.2.2.1 Preparation of the construct T5OHns-YFP

The CYP T5OH catalysing the first oxygenation of the Taxol biosynthetic pathway has been predicted to be located to the endoplasmic reticulum, consistent with the localisation of other plant CYPs. The sequence of T5OH cloned from *T. baccata* was analysed by ChloroP software and was predicted to not have a chloroplast transit peptide as would be expected for CYP enzymes. Taxadiene being a non-polar compound would need to be translocated across the plastid membrane to undergo oxygenation of its taxane core by the T5OH. It was reasoned that the T5OH should be spatially localised somewhere between the plastid and the ER to aid translocation of the taxadiene compound.

In order to verify the localisation of this mono-oxygenase, the cDNA of T5OH cloned from *T. baccata* mRNA pool was amplified using the T5OH::pCR8 (Section 3.2.2) as template, *Pfu* DNA polymerase, the forward primer T5OH-F and the reverse primer T5OH-Rns (see Appendix I) designed to exclude the stop codon at the C-terminal of the cDNA. An ~1.6 kb PCR product (Figure 4.6B) was obtained which was gel purified and subsequently cloned into the entry vector pCR8 by TOPO cloning and transformed into *E.coli strain* DH5α. The presence of the insert in the *E.coli* colonies which were resistant to spectinomycin was verified by PCR using the primers T5OH-F and T5OH-PstRev which amplified a fragment of 743 bp (Figure 4.6A & C). The purified plasmid from colonies 2, 4 and 7 were sequenced with the pCR8 primers M13Forward and M13Reverse (Figure 4.6A). Amino acid sequence of T5OH initially cloned from *T. baccata* was aligned with that obtained from T5OHns

(Figure 4.6D) to confirm that the stop codon has been deleted from the C-terminal of the T5OH cDNA.



(D)

```
T5OH      SCVGWEFSKMEILLFVHHFVKTFSSYTPVDPDEKISGDPLPPLPSKGFSIKLFPRP---- 499
T5OHNS    SCVGWEFSKMEILLFVHHFVKTFSSYTPVDPDEKISGDPLPPLPSKGFSIKLFPRPKGEF 539
```

Figure 4.6 Preparation of the T5OHns::pCR8 construct

(A) Schematic diagram of the construct generated by TOPO cloning carrying the cDNA T5OH without its stop codon. The positions of the primers T5OH-F and T5OH-Rns used for PCR amplification of the cDNA and T5OH-PstRev are indicated (B) PCR product showing a band of size ~1.6 Kb obtained using primers T5OH-F and T5OH-Rns indicating the presence of T5OHns cDNA (C) Colonies 1 to 10 showing the presence T5OH obtained by PCR with primers T5OH-F and T5OH-PstRev, where C is the positive control T5OH::pCR8. (D) Amino acid sequence alignment of T5OH which was used as template for the amplification of T5OHns and the amplified T5OHns without its stop codon. X= similar amino acid, X= part of pCR8 vector

Following sequencing which confirmed that the T5OHns has been amplified with a 100% match to the T5OH cloned from *T. baccata*, the T5OHns from colony 7 was fused in frame by LR recombination to YFP in the plant binary vector pGKPGWY (Zhong *et al.*, 2008) driven by a 35S promoter. The presence of the insert was verified by PCR on the colonies resistant to kanamycin, using the full-length primers T5OH-F and T5OH-Rns to give the expected band of ~1.6 kb (Figure 4.7B). The resulting vector thus named T5OHns::pGKPGWY (T5OH-YFP) (Figure 4.7A) was transformed into *Agrobacterium* which was used for *Arabidopsis* transformation as previously described (Section 2.5.1).

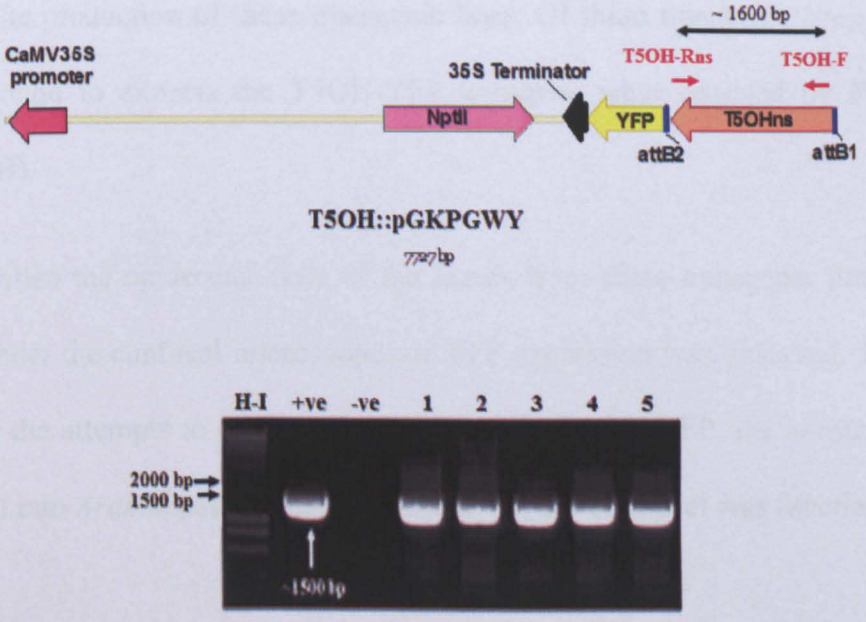


Figure 4.7 Presence of the insert T5OH-YFP

(A) Schematic diagram of the T5OH::pGKPGWY construct generated by VectorNTI (Invitrogen). The primers are indicated by the red arrows. The construct carries the yellow fluorescent protein and the nptII gene which confers resistance against kanamycin. (B) Gel picture following PCR amplification using primers T5OH-F and T5OH-Rns shows the presence of the ~1.6 Kb insert T5OH in the *E.coli* colonies 1-5 after LR recombination.

The initial attempts to transform the binary vector pGHPGWY carrying the T5OHns, a YFP gene and conferring hygromycin resistance in plants were unsuccessful in both *Arabidopsis* and tobacco. Consequently, the binary pGKPGWY conferring kanamycin resistance was used instead for *Arabidopsis* transformation.

Efforts to transform tobacco with the construct containing 35S::T5OH::YFP in many instances have returned only a few transgenic lines, where the T5OH transgene was not expressing when an RT-PCR was performed on the samples (data not shown). Seven transgenic lines carrying the *T5OH* gene (Figure 4.8A) were obtained after an elevated concentration of IAA was used in the regeneration medium, which facilitated the production of these transgenic lines. Of these transgenic lines, plants 1-4 were found to express the T5OH-YFP transgene when assayed by RT-PCR (Figure 4.8B).

However, when the epidermal cells of the leaves from these transgenic lines were observed under the confocal microscope, no YFP expression was observed. Thus, to accomplish the attempts to express and visualise the T5OH-YFP, the construct was transformed into *Arabidopsis* instead to test whether the construct was functional.

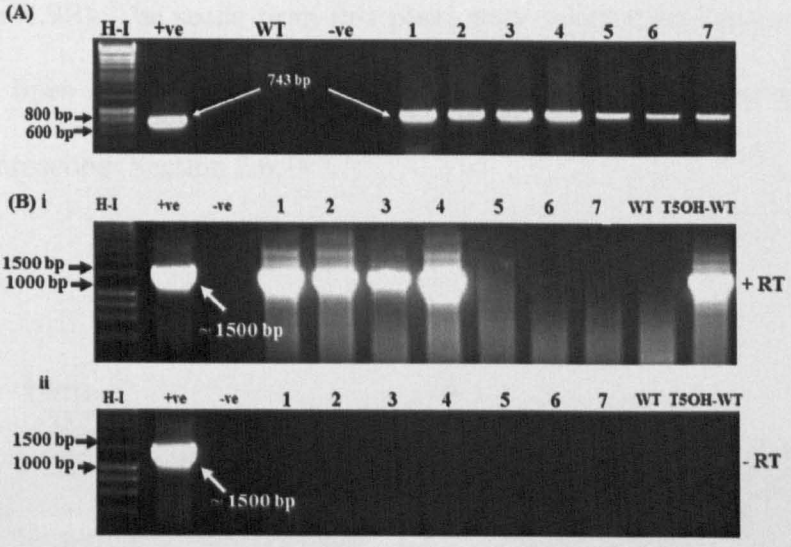


Figure 4.8 Analysis on transgenic tobacco plants

(A) PCR analysis was carried out on the gDNA extracted from tobacco plants obtained following transformation with the T5OH-YFP construct. The primers T5OH-F and T5OH-PstRev were used which amplified a PCR fragment of size ~700 bp (B) RT-PCR results following RNA extraction and DNase treatment of the RNAs from the transgenic lines 1-7. PCR performed with primers T5OH-F and T5OH-Rns to give a band of ~1500 bp, characteristic of T5OH DNA. (ii) A control was set up to check that the samples did not give positive results due to the presence of transgene DNA whereby the RT enzyme was not added during this reaction. +ve – positive control using T5OH::pCR8 plasmid; -ve – negative control using water, WT – wild type, T5OH-WT – transgenic line expressing the gene T5OH without fluorescent tag.

4.2.2.2 Sub-cellular localisation of T5OH in transgenic Arabidopsis

Following *Arabidopsis thaliana* (ecotype Columbia, Col-0) transformation with *Agrobacterium* carrying the 35S::T5OH::YFP construct, the seeds were screened and selected on kanamycin. The presence of the T5OH fragment was confirmed by PCR analysis (Figure 4.9A) on the genomic DNA (gDNA) extracted from the single *Arabidopsis* line which was resistant to kanamycin. The transgenic line was grown to maturity and the expression of the T5OH transgene mRNA was confirmed by RT-

PCR (Figure 4.9B). The seeds from this plant were selected on kanamycin and the roots of the lines resistant to the antibiotic were mounted and observed under the confocal microscope (Section 2.6.1).

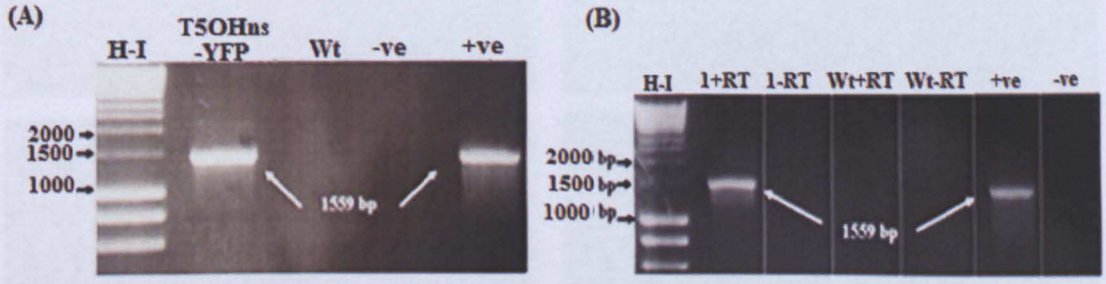


Figure 4.9 Analysis on transgenic Arabidopsis

(A) Gel picture of PCR analysis on the gDNA extracted from *Arabidopsis* line 1 (T5OHns+YFP) shows the presence of the ~1.6 Kb T5OH fragment. –ve = water, +ve = 35S::T5OH::YFP construct (B) RT-PCR was carried out using the total RNA extracted from the *Arabidopsis* leaf tissue. The gel picture shows the RT-PCR results obtained while using the primers T5OH-F and T5OH-Rns. The length (~1.6 Kb) of the fragment confirmed the expression of the T5OHns transgene mRNA in line 1. +RT = reverse transcriptase enzyme added, -RT = reverse transcriptase enzyme omitted during first strand cDNA synthesis, H-I – Hyperladder I as molecular marker.

Amino acid sequence analysis of T5OH as described in section 3.2.2 indicated a transmembrane helix at the N-terminal of the protein in agreement with the "predictprotein" results of T5OH being located to the ER. T5OH was not predicted to have a plastid transit peptide by ChloroP, as would be expected for CYP enzymes which are predominantly located in the ER. However, confocal images (Figure 4.10) of transgenic *Arabidopsis* roots showed that the T5OH-YFP was associated to the plastids (Figure 4.10 (iv) and (vi)) as well as localised to the ER network around the nucleus of the root cells (Figure 4.10 ii).

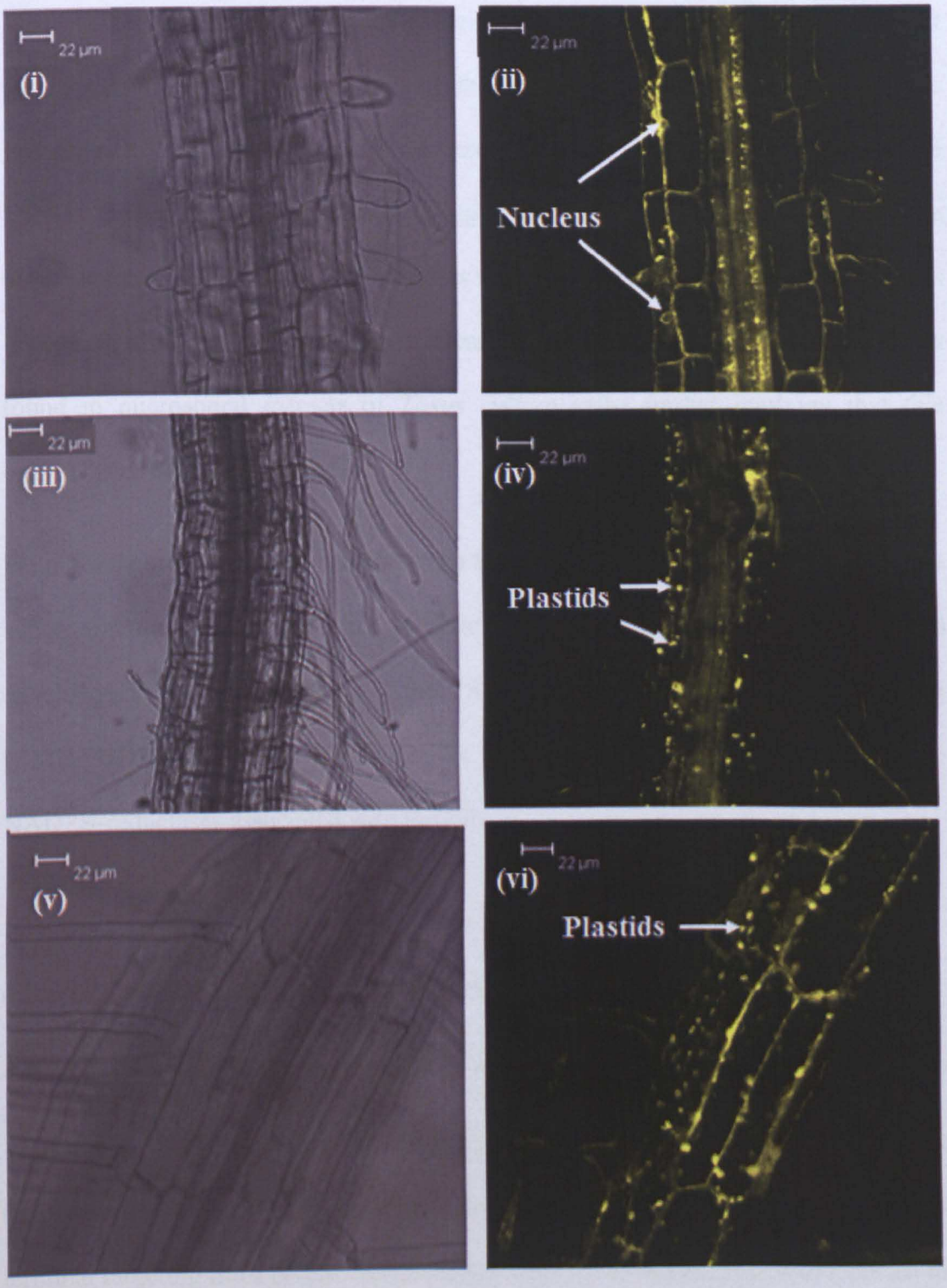


Figure 4.10 Localisation of T5OH-YFP.

Confocal scanning fluorescence microscope images of T5OH-YFP expressed in *Arabidopsis* roots showed the T5OH was targeted to the endoplasmic reticulum around the nucleus (i) and the plastid (iv) & (vi). (i), (iii) and (v) show the phase contrast images of the same roots.

Most nuclear-encoded chloroplast proteins have a cleavable N-terminal transit peptide, although proteins targeted to the chloroplast outer envelope membrane generally lack this transit peptide (Keegstra and Cline, 1999), as is the case of the T5OH. Overall, the confocal images suggested that T5OH might be targeted to the outer envelope membrane of the plastid as well as being associated with the ER. This is consistent with the previous reports (Jennewein *et al.*, 2004a) that the protein was found in microsomal extracts of *Taxus* culture cells, further implying that this enzyme is co-localised to the ER.

With T5OH spatially distributed between the plastid and the ER, this has led to the suggestion that this enzyme might attach itself to the taxadiene at the plastid envelope and help its movement to the ER, where the first hydroxylation at the C-5 position of the taxane core takes place by T5OH, in the presence of the cytochrome NADPH-reductases associated with the ER.

4.2.3 Taxadiene-5acetyl transferase (T5AT) is targeted to the ER

4.2.3.1 Sub-cellular localisation of T5AT in transgenic tobacco plants

Walker *et al.* (2000) reported that sequence analysis of the protein T5AT catalysing the third step of the Taxol biosynthetic pathway, did not appear to encode any N-terminal targeting information. However, the activity of this enzyme was reported to reside in the microsomal extracts, which led the authors to propose the protein to be localised to the ER. To study its localisation, T5AT cDNA was amplified without its stop codon using the T5AT::pCR8 as template and the primers 5Tax5Acyl (forward) and 3Acetyl T-NSC (reverse) (Appendix I) (work performed by former MSc student

Raj Kamal Mann). The amplified fragment was fused in frame to the plant binary vector pGBPGWC (Zhong *et al.*, 2008) to give the construct T5ATns::pGBPGWC (T5AT-CFP) (Figure 4.11A) which confers resistance against basta and carries a cyan fluorescent protein (CFP). This construct was transformed into wild type tobacco and the explants were selected on 10 mg/L Basta. The presence of the insert T5AT was verified by PCR and the transgenic plants were grown for 1 month.

A northern blot assay was carried out on the total RNA extracted from the leaf tissues of 1 month-old transgenic plants carrying the T5AT-YFP construct, in order to investigate the level of T5AT transcript expression in these RNA samples. The probe (Figure 4.11A) used for membrane hybridisation was prepared by PCR amplification of the T5AT using the primers 5Tax5Acyl and 3Acetyl T-NSC. The northern blot results (Figure 4.11B) indicated that the T5AT transgene was being expressed in all the transgenic lines, with line F having the lowest level of T5AT transcript expression.

The lower surface of the epidermis of the transgenic plants were mounted on respective microscopic slides and observed with the confocal microscope. The confocal images of the CFP fluorescence in the 35::T5AT::CFP transgenic tobacco cells (Figure 4.12) showed that the T5AT-CFP fusion protein was localised in the ER, and did not co-localise with the red chlorophyll auto-fluorescence of the chloroplasts. Sequence analysis of the T5AT by the "predictprotein" software indicated the presence of a transmembrane helix at the N-terminal of the protein, suggesting the protein to be localised to the ER, consistent with the confocal images of the transgenic tobacco leaf epidermal cells. However, some of the CFP

fluorescence were observed in the cytoplasm of the trichomes of the tobacco leaves (Figure 4.12; iii, iv and v). This cytoplasmic signal may have been in the endoplasmic reticulum of the trichomes, but the images obtained did not allow a conclusive location to be determined.

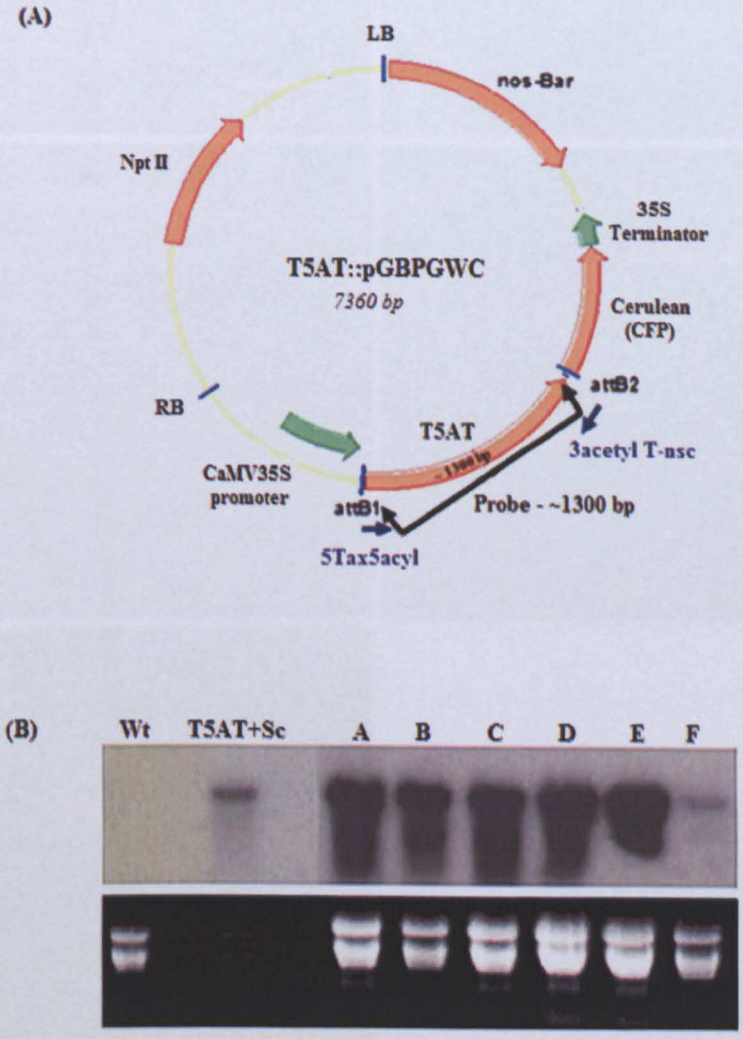


Figure 4.11 Expression of T5AT in tobacco leaves

(A) Schematic diagram of the T5AT::pGBPGWC vector generated by VectorNTI. Expression is driven by a constitutive 35S promoter and T5AT is fused at its C-terminus to a CFP gene. The construct also carries the nos-Bar gene which confers resistance to basta herbicide. (B) Northern blot analysis was carried out using 10 µg of total RNA from the tobacco leaves and hybridised to ^{32}P labelled cDNA probe which was generated by PCR using the T5AT::pGBPGWG as template. WT – RNA from wild type tobacco, T5AT+Sc – Tobacco line expressing the full-length T5AT cDNA.

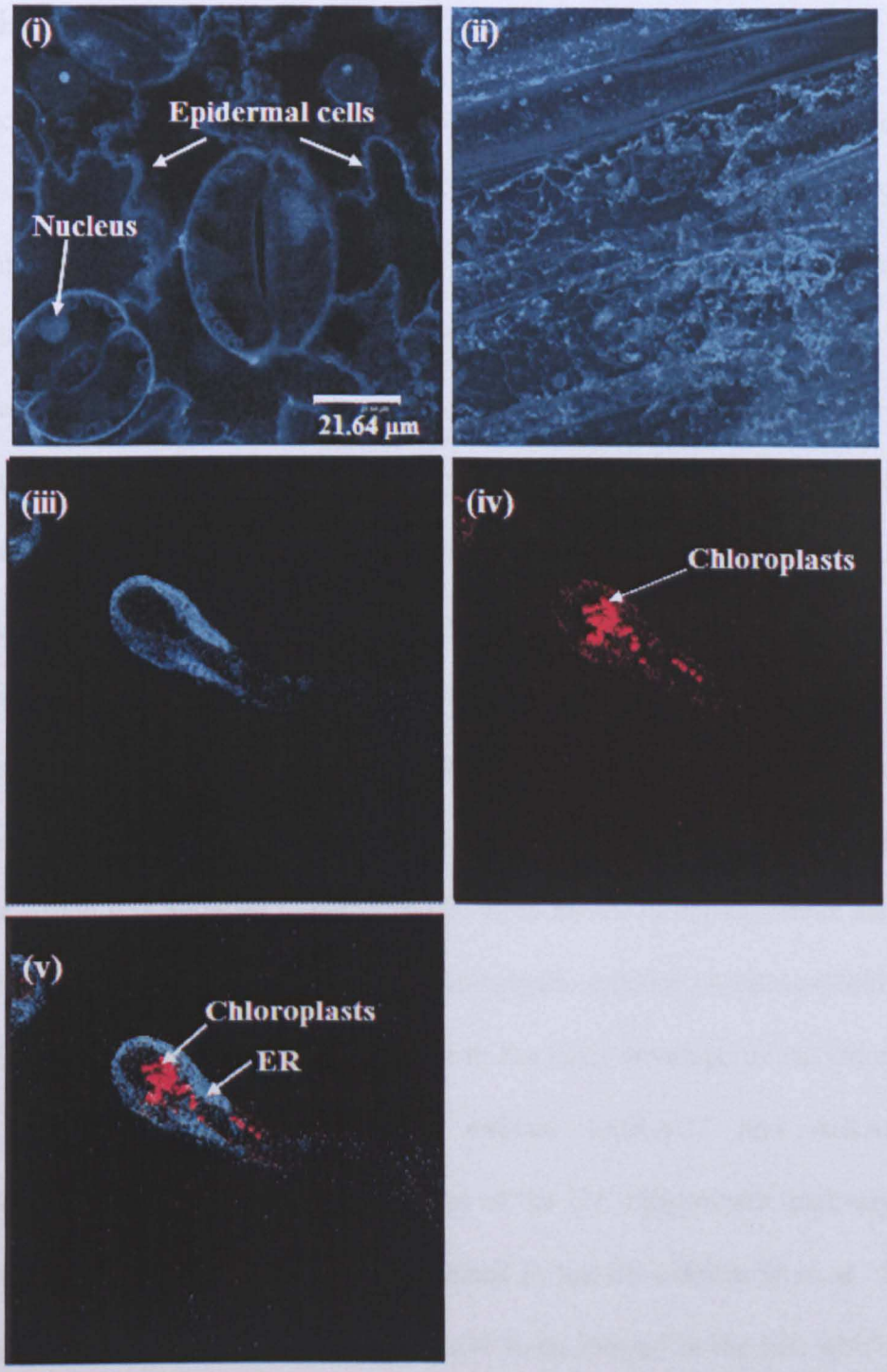


Figure 4.12 Localisation of T5AT-CFP

Confocal scanning fluorescence microscope images of T5AT-CFP expressed in tobacco epidermal cells showed the T5AT was targeted to the ER and possibly to the cytoplasm. (i) and (ii) indicate that the CFP fluorescence is predominantly localised to the ER around the nucleus and in the ER network respectively. (iii), (iv) and (v) show the CFP fluorescence, Chlorophyll auto-fluorescence and the CFP (blue) superimposed on the chlorophyll auto-fluorescence (red) respectively in the trichome.

4.3 Discussion

The sequence similarity and close clustering of the enzymes involved in the early steps of Taxol biosynthesis with those of the GA biosynthetic pathway (Section 3.3.3) might indicate that they might share a common evolutionary ancestry. In both GA and Taxol biosynthesis, these enzymes carry out the cyclisation of GGPP followed by CYP catalysed hydroxylations of the resulting skeleton. The genes encoding the first three steps of the gibberellins biosynthetic pathway in *Arabidopsis* have been isolated and their sub-cellular localisation previously studied by Helliwell *et al.* (2001b). The first two enzymes copalyl diphosphate synthase (AtCPS1) and *ent*-kaurene synthase (AtKS1) of the terpene cyclase, were both reported to catalyse the first two steps of GA biosynthesis from GGPP to *ent*-kaurene and were both localised to the plastid stroma. The first three oxidation steps from the non-polar *ent*-kaurene which was believed to partition into membranes, to *ent*-kaurenoic acid have been shown to be catalysed by the CYP *ent*-kaurene oxidase (AtKO1) (Helliwell *et al.*, 1999). This enzyme has been localised to the outer envelope of the chloroplast, while the enzyme *ent*-kaurenoic acid oxidase (AtKAO1 and AtKAO2 in *Arabidopsis*) catalysing the next three steps of the GA biosynthetic pathway, from ent-kaurenoic acid to GA_{12}, has been localised to the ER (Helliwell *et al.*, 2001a). The authors also reported some of the AtKO1 to be located in the ER, which led to the suggestion that *ent*-kaurene export from the plastid is linked to its further oxidation where it forms more water-soluble intermediates which are further metabolised by the ER-localised AtKO1 *en route* to GA.

It was reasoned that if the GA and Taxol biosynthetic enzymes share similar evolutionary origin and carry out similar kinds of reactions, the localisation of the early enzymes of Taxol and GA biosynthesis might also share similar sub-cellular localisation. Fusions of the Taxol biosynthesis enzymes TXS, T5OH and T5AT to fluorescent proteins were used to confirm the similar subcellular localisations of the early enzymes of the Taxol pathway to those of the GA pathway. Transgenic tobacco and *Arabidopsis* lines were constructed by *Agrobacterium* transformation with plasmid vectors containing a functional antibiotic resistance selectable marker and a fluorescent protein tag.

The N-terminal pre-sequences of the *T. baccata* terpene TXS encodes a plastidial peptide sequence which was shown to direct the protein to chloroplasts in transient assays following microprojectile bombardment of onion epidermis. The TXS-GFP fusion was also found to be imported by *A. thaliana* and *N. tabacum* chloroplasts. These results are consistent with the previous prediction of the TXS enzyme being localised to the plastids where it is involved in the cyclisation of the GGPP from the MEP pathway to yield taxadiene, the first compound of the Taxol biosynthetic pathway.

The first oxygenation of the Taxol pathway is catalysed by the cytochrome P450 mono-oxygenase T5OH which has been proposed to be associated to the ER. The N-terminal portion of T5OH carries a potential transmembrane signal sequence, suggesting that the enzyme is localised to the ER. However, the results of the confocal microscopy during this research indicated that a substantial proportion of YFP fused to the T5OH was localised to the plastids of *Arabidopsis*. Since the

sequence of this gene does not contain a cleavable transit peptide, it was speculated that the enzyme was localised on the outer face of the chloroplast envelope as well as being associated to the endoplasmic reticulum. Many proteins targeted to the chloroplast outer envelope membrane lack a cleavable transit peptide (Keegstra and Cline, 1999), as seems to be the case for T5OH and also for AtKO1 (Helliwell *et al.*, 2001b). The interesting point that was clarified in this area of research, was that the T5OH has more than one localisation in the cells. This is consistent with the suggestion of Werck-Reichhart and Feyereisen (2000) that some P450 enzymes have more than one subcellular localisation. T5OH and the *Arabidopsis ent*-kaurene oxidase have high sequence similarity indicative of a common evolutionary origin, both carrying out the first CYP-catalysed oxygenation step of a cyclised diterpene and both appear to have a dual localisation to both the plastid envelope and the ER (Helliwell *et al.*, 2001b; Section 4.2.2.2). The localisation of T5OH might suggest that this hydroxylase associated to the plastid-envelope is required for the export of taxadiene from the plastid to the ER as well as catalysing its conversion to taxadiene-5α-ol.

The same construct, T5OH-YFP, used for *Arabidopsis* transformation was also used for tobacco transformation. However, only seven transgenic lines were recovered and this was only possible after an elevated level of auxin was used. Of these plants, only four were found to be expressing the T5OH-YFP transgene. Plants expressing the full-length T5OH mRNA without the fluorescent tag were separately generated and were all found to be sterile, with non-viable pollen. This suggests that an active enzymatic form of T5OH may have substrates upon which it can act in tobacco. This could result in the pollen-lethality either by generating a new, toxic compound, or by

depleting an essential substrate. However, none of the transgenic lines expressing the T5OH-YFP construct were found to be sterile, which could lead to the suggestion that the T5OH-YFP mRNA was not efficiently translated, or that the protein produced was rapidly turned over, which would also explain the absence of YFP fluorescence in these lines.

Sequence analysis of T5AT indicated a transmembrane helix sequence, suggesting this enzyme to be localised to the ER. The prediction was shown to be correct for T5AT which was demonstrated to be localised to the ER of the epidermal cells of transgenic tobacco lines expressing the T5AT-GFP. This localisation data correlates with the previous report of this enzyme being present in *Taxus* microsomes (Wheeler *et al.*, 2001) which led the authors to suggest the localisation of this enzyme to the ER. Furthermore, there was no overlap of the CFP and chlorophyll signals from the transgenic cells for this construct confirming that T5AT is not associated with the chloroplasts, but is ER-located. The "predictprotein" software also predicted the T5AT to be a soluble enzyme localised to the cytoplasm, in agreement to the apparent localisation of T5AT-CFP in the cytoplasm of the transgenic tobacco trichomes. This cytoplasmic signal may have been in the endoplasmic reticulum of the trichomes, but the images obtained did not allow a conclusive location to be determined.

In conclusion, the confocal images of the translational fusion of the fluorescent proteins and the respective *T. baccata* genes have shown that TXS is targeted to the chloroplast, T5OH is associated to the outer envelope of the chloroplast as well as to the ER surrounding the nucleus, and T5AT is targeted to the ER. These results

confirmed that the stages of the Taxol biosynthetic pathway are compartmentalised within the plant cell, with T5OH potentially providing a crucial link between the plastid and ER-located steps of the Taxol biosynthetic pathway. These localisation data might provide an explanation to the previous hurdles encountered during the reconstitution of the Taxol biosynthetic pathway in *S. cerevisiae* (Dejong *et al.*, 2005) and in *E.coli* (Huang *et al.*, 1998; 2001). Both microbes lack the plastid organelles where TXS is targeted for the cyclisation of GGPP into taxadiene. The authors have reported the use of a truncated version of TXS where the plastidial sequence was deleted, to produce a more soluble form of the enzyme. However a maximum of only 0.5 mg/g DW of taxadiene was produced in yeast and 1.3 mg/L in *E.coli* which suggests a limited GGPP pool available in microbes for taxadiene synthesis. Dejong *et al* (2005) also reported a pathway restriction at the T5OH-hydroxylation step which led the authors to suggest that this might have been due to the low endogenous levels of NADPH-cytochrome P450 reductase (CPR) in yeast, required for redox coupling with the CYP T5OH. The localisation data of T5OH has indicated the association of this hydroxylase with the plastidial envelope and the ER, suggesting that in addition to being required for the synthesis of taxadiene-5α-ol, T5OH may also be needed for taxadiene translocation from the plastid to the ER where further modification of the taxane core takes place. The plastidial association of T5OH might provide an explanation of the pathway restriction in yeast which do not possess this particular organelle. Furthermore the low endogenous level of CPR in yeast might also restrict the action of the downstream CYP enzymes of the Taxol pathway which have all been suggested to be associated to the ER.

Taking into account these pathway restrictions previously encountered in these two microbes, it was reasoned that plants might provide a better system for the heterologous expression of the Taxol biosynthetic genes, where the presence of the large available pool of GGPP and the substantial number of plastids in plant cells would provide a better platform for taxadiene synthesis. Also, the high levels of endogenous CPR and the presence of endoplasmic reticulum in plants would allow further modifications of the taxane core by the CYP enzymes of the Taxol biosynthetic pathway which are responsible for nearly half of the steps of the pathway.

CHAPTER 5 : HETEROLOGOUS EXPRESSION OF THE TAXOL BIOSYNTHETIC GENES

5.1 Introduction

The slow growth of yew trees and the relatively low content (0.01-0.1% dry weight of bark) (Hezari *et al.*, 1997) of Taxol extracted limit the supply of the drug. The presence of unwanted analogues with very similar structures to that of Taxol makes its purification difficult. Transferring genes of the Taxol biosynthetic pathway into heterologous organisms for their overexpression may provide a method for increased production of the taxanes of interest whilst preventing the formation of side products, which would decrease the production cost of these compounds. It might also address the issue of separation and purification of the desired compound compared to its extraction from the native plant source as well as providing a route to novel taxanes with improved solubility and similar antitumour properties to Taxol.

Attempts for the reconstitution of the Taxol biosynthetic pathway in heterologous systems have been reported with mixed success in yeast, *E.coli, Arabidopsis thaliana* and *Nicotiana sylvestris*. The use of plant systems for the overexpression of the Taxol biosynthetic genes appears to be advantageous, due to the presence of plastids for the committed cyclisation of GGPP to taxadiene, and ER, where further modifications of the taxane core can take place. Also, the presence of high level of plant endogenous CPRs will be useful for the oxygenations of the taxadiene core by

the *Taxus* CYP mono-oxygenases. The other advantage of using heterologous plant species as expression hosts is the potential for cheap and easily extractable source of taxadiene and the other downstream taxanes.

A previous report on the constitutive expression of TXS in *Arabidopsis* has documented the production of taxadiene at level 20 ng/g DW (Besumbes *et al.*, 2004). Besides the low level of taxadiene produced in these *Arabidopsis* plants, the authors also reported pleiotropic effects associated with the constitutive expression of TXS; including reduced hypocotyl length, retardation of growth and flowering and pale, bleached phenotype. It was reasoned that the level of GGPP accumulation in *Arabidopsis* might be insignificant for its redirection for the production of taxanes in this plant system, without the adverse phenotypic effects.

The successful production of taxadiene in *yellow flesh* tomato mutant at a level of 339 μg/g DW in ripe fruits constitutively expressing TXS (Kovacs *et al.*, 2007) has led to the speculation that the GGPP in the easily manipulated *N. tabacum* could be redirected for the production of taxadiene and downstream taxanes in the plant. These taxanes could be used as starting material for the semi-synthesis of novel taxanes with better solubility than Taxol, and having the ability of modulating MDR in drug-resistant tumour cells.

5.2 Results

5.2.1 Analysis of taxadiene producing tobacco lines

T. baccata TXS cDNA was cloned into the plant binary vector pBC35, under the transcriptional control of the CaMV 35S and terminator cassette. The resulting plant binary vector thus named pBCTax has been used to transform wild type (Wt) tobacco plants (work previously done by Dr. Katalin Kovacs, The University of Nottingham, UK). The seeds from these taxadiene producing tobacco lines were provided and were grown for one month on MSR3 medium containing kanamycin antibiotic.

The crude hexane extracts from four independent lines constitutively expressing the TXS gene were analysed for the presence of taxadiene in the leaves. 1µl of the hexane extract from each leaf sample was subjected to chromatography-mass spectrometry (GC-MS) analysis. GC-MS analysis was performed on a GC 8000 gas chromatogram linked to an MD800 mass spectrometer (Fisons, Manchester UK). Samples were injected in split-less mode (injector 250°C) onto a 30m x 0.25 mm ID fused silica capillary column BP-5 (SGE, Milton Keynes UK) with a 1.0 µm film thickness. The initial oven temperature was 160°C, this was ramped to 300°C at 10°/min after a 1 min delay (using helium as the carrier gas).

All four lines analysed were found to give a major GC peak at retention time (Rt) 8.03 min and a smaller peak at Rt 7.20 min (Figure 5.1 i) which were both absent in the wildtype tobacco control. These peaks were scanned for the major ions at m/z 122 (100%), 121, 123 and 107, characteristic of taxadiene. The mass spectrum of the peak at Rt 8.03 was compared to that of taxadiene produced by the *r* tomato fruits (Figure 5.1 iii) (Kovacs *et al.*, 2007) and was found to be almost identical to that of

taxa-4(5),11(12)-diene. The mass spectrum at RT 7.20 min was not scanned but was speculated to be that of the taxa-4(20),11(12)-diene isomer which has been previously reported to have an RT very close to that of taxa-4(5),11(12)-diene (Williams *et al.*, 2000).

The line showing a more distinctive mass spectrum of taxadiene was transformed at many instances with a pGHPGWY (hygromycin resistant) into which T5OH was cloned, with the aim of generating transgenic lines co-expressing the two transgenes. Only a few transgenic lines were generated, however, the crude hexane extracts from these lines indicated no presence of taxadiene-5α-ol when subjected to GC-MS analysis (data not shown). As an alternative way of "stacking" the two transgenes TXS and T5OH in individual lines for the potential production of the downstream taxane, tobacco line producing taxadiene were to be cross-pollinated with lines expressing the T5OH transgene.

The seeds from the tobacco line showing the best mass spectrum of taxadiene were sown and crude hexane extracts from 21 of the progeny from this line were analysed by GC-MS under the following conditions: the initial oven temperature was 60°C, This was ramped to 400°C at 10°/min after a 1 min delay (using helium as the carrier gas). The 122 ion scan of the GC chromatograms indicated only three of the progenies (Line 1, 3 and 9) to show a peak at RT 36.58 min having mass spectrum similar to authentic taxadiene (Figure 5.2 i showing 122 ion scan for line 1) which was absent in the Wt tobacco hexane extract (Figure 5.2 iii (a & b)).

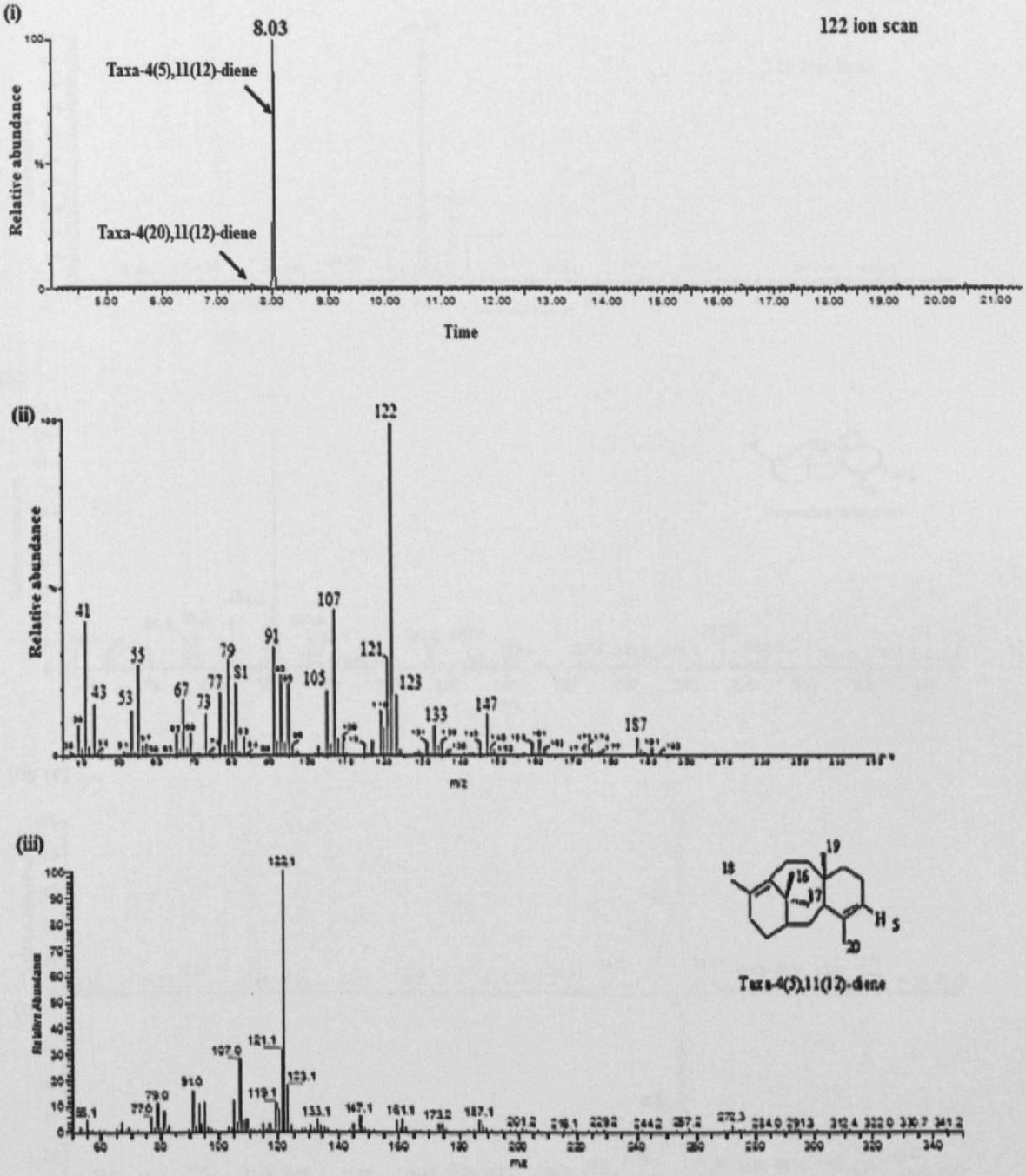


Figure 5.1 GC-MS analysis of crude hexane extracts from taxadiene synthase expressing tobacco plants.

GC-MS analysis of crude hexane extracts from TXS expressing tobacco plants (i) 122 ion scan of the GC chromatogram in extracts from transgenic leaf showing the peak at Rt 8.03 for taxa-4(5),11(12)-diene and the potential isomer taxa-4(20),11(12)-diene at Rt ~7.2 min. (ii) Mass spectrum of the peak with Rt 8.03 min from tobacco extracts is almost identical to that of (iii) taxa-4(5),11(12)-diene from transgenic tomato fruit expressing TXS.

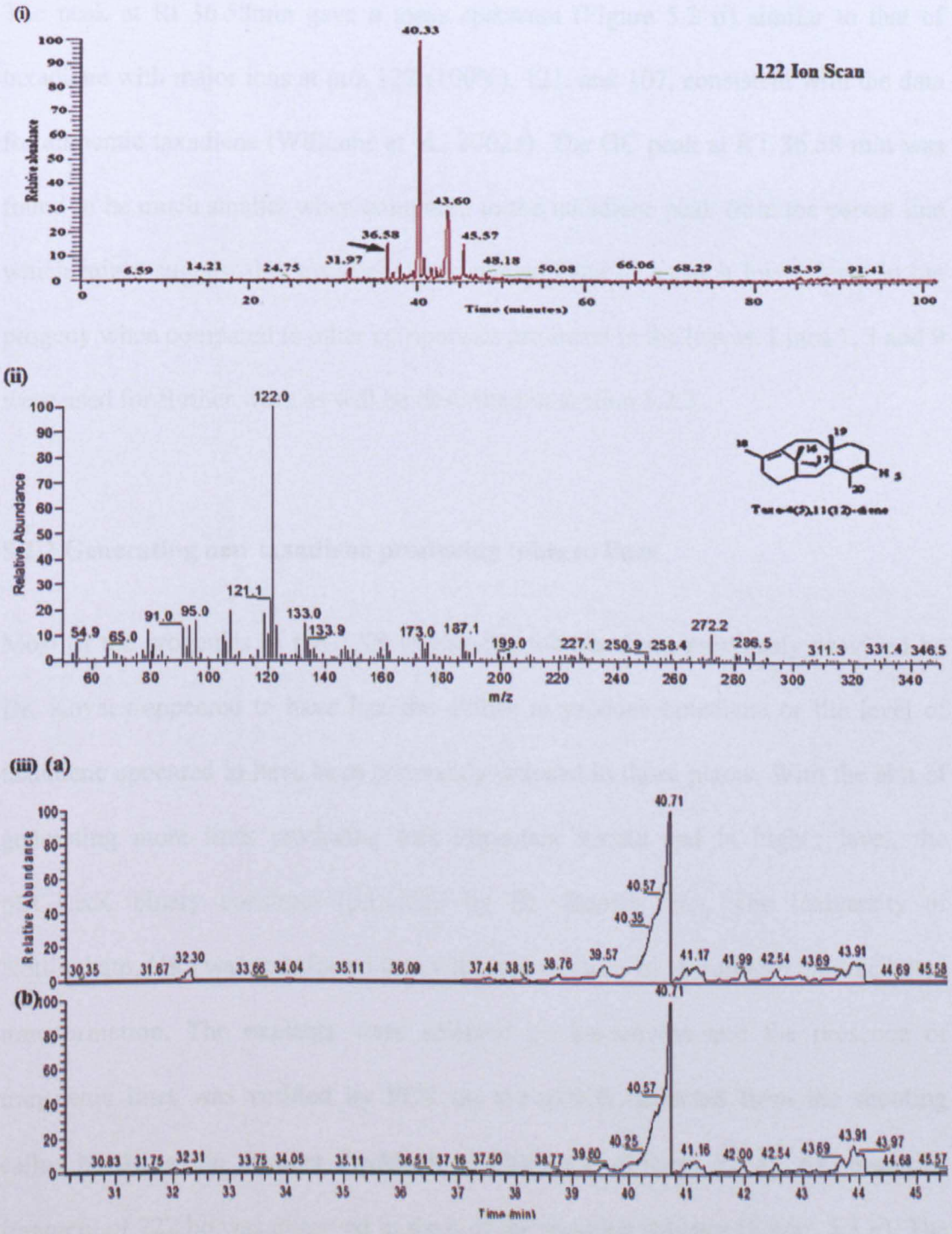


Figure 5.2 GC-MS analysis of crude hexane extracts from TXS Line 1.

The leaf exudates of TXS line 1 and Wt tobacco were subjected to GC-MS analysis. (i) GC profiles of 122 ion scan indicating the taxadiene peak at 36.58 minute. (ii) Mass spectral fragmentation of the peak at retention time 36.48 min having major ions characteristic of taxa-4(5),11(12)-diene. (iii) (a) TIC in extracts from wild type tobacco and (b) 122 ion scan showing the absence of the peak characteristic of taxadiene at 36.58 minutes under the same GC-MS conditions.

The peak at Rt 36.58min gave a mass spectrum (Figure 5.2 ii) similar to that of taxadiene with major ions at m/z 122 (100%), 121, and 107, consistent with the data for authentic taxadiene (Williams et al., 2002a). The GC peak at RT 36.58 min was found to be much smaller when compared to the taxadiene peak from the parent line which might suggest that taxadiene was accumulated to a much lower level in the progeny when compared to other compounds produced in the leaves. Lines 1, 3 and 9 were used for further work as will be described in section 5.2.3.

5.2.2 Generating new taxadiene producing tobacco lines

Most of the progenies of the TXS expressing tobacco lines previously provided by Dr. Kovacs appeared to have lost the ability to produce taxadiene or the level of taxadiene appeared to have been constantly reduced in those plants. With the aim of generating more lines producing this important taxane and in higher level, the pBCTaxK binary construct (provided by Dr. Rupert Fray, The University of Nottingham, UK) was transferred into wild type tobacco by *Agrobacterium*-mediated transformation. The explants were selected on kanamycin and the presence of transgenic lines was verified by PCR on the gDNA extracted from the shooting callus by using the primers TaxMid1 and Taxint35 (Figure 5.3 i). An amplified fragment of 722 bp was observed in most of the shooting calluses (Figure 5.3 ii). The transgenic plants were allowed to root in media containing auxin, before being transferred to soil and grown for one month.

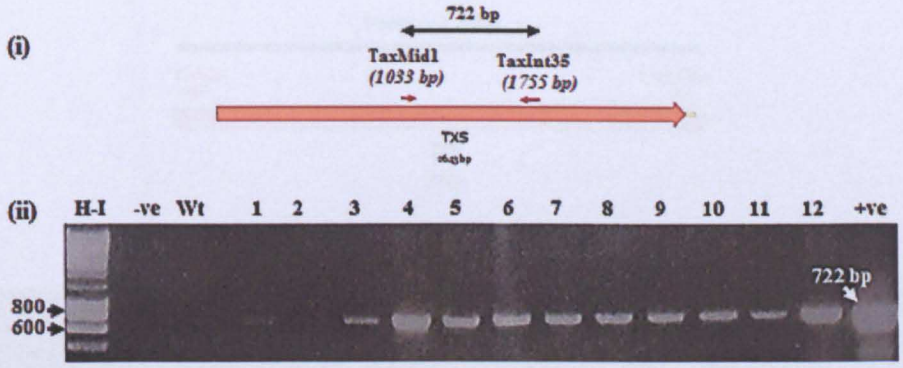


Figure 5.3 PCR analysis to check for TXS gene in shooting callus

(i) The schematic diagram of TXS cDNA and primers used for PCR analysis were generated by vector NTI (Invitrogen). (ii) Gel picture showing PCR amplification of leaf gDNA from lines 1-12 to check for the presence of TXS. PCR amplified fragment of 722 bp using the primers TaxMid1 and Taxint35 confirms the presence of the TXS transgene in most of the shooting callus. H-I refers to the HyperladderI molecular marker, Wt refers to wild type and –ve refers to the negative control where water was used instead of DNA.

A northern blot assay was carried out on the total RNA extracted from the transgenic leaves to investigate the level of expression of the TXS transgene mRNA in the transgenic plants. Line 4 which was found to express the highest level of TXS mRNA (Figure 5.4) died at a very young stage (~1 month old), thereby preventing any further analysis of this line. This led to the speculation that the high level accumulation of taxadiene in this line might have been toxic/lethal to the plant.

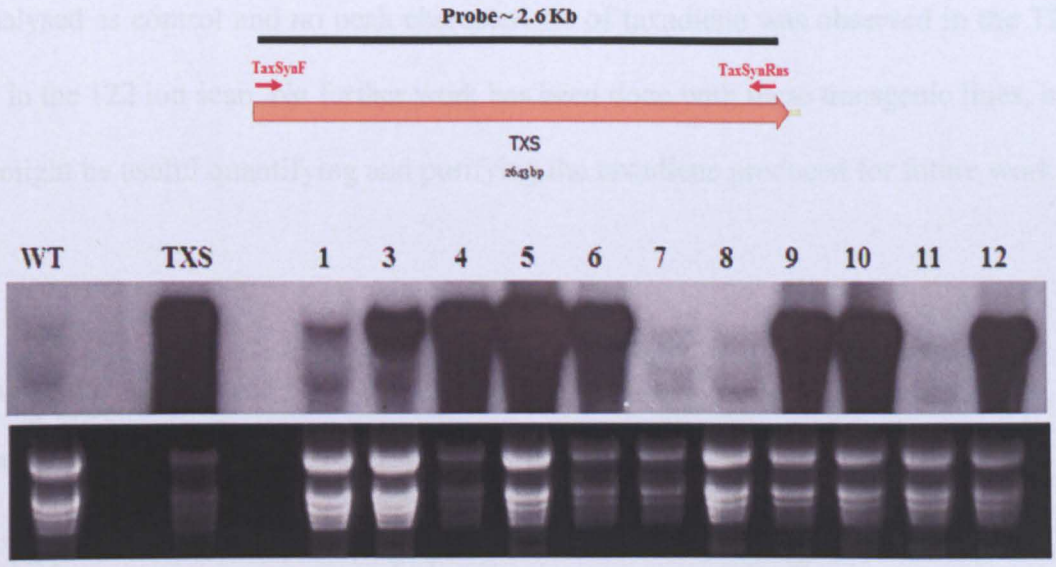


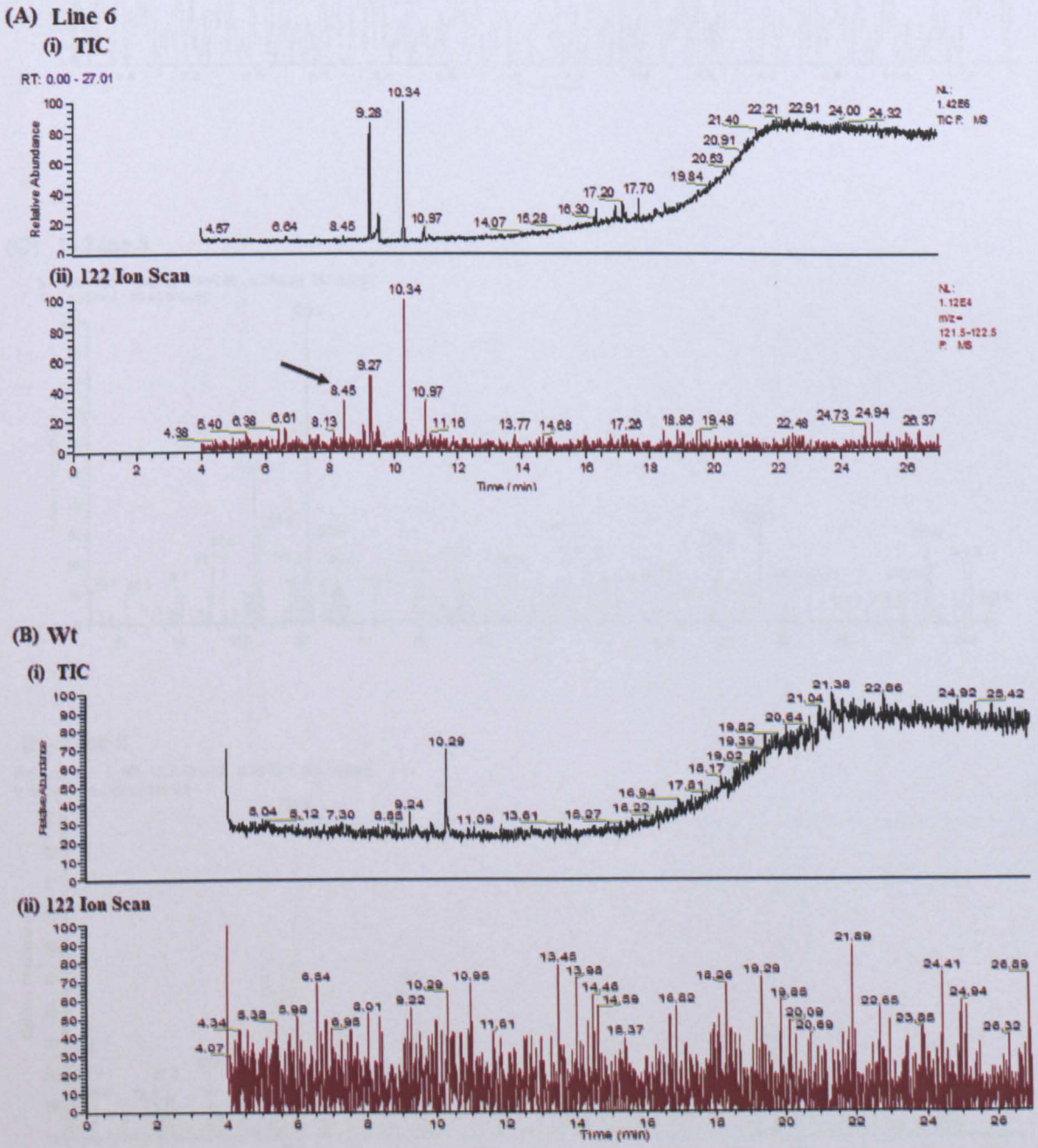
Figure 5.4 Expression of TXS mRNA in transgenic tobacco plants.

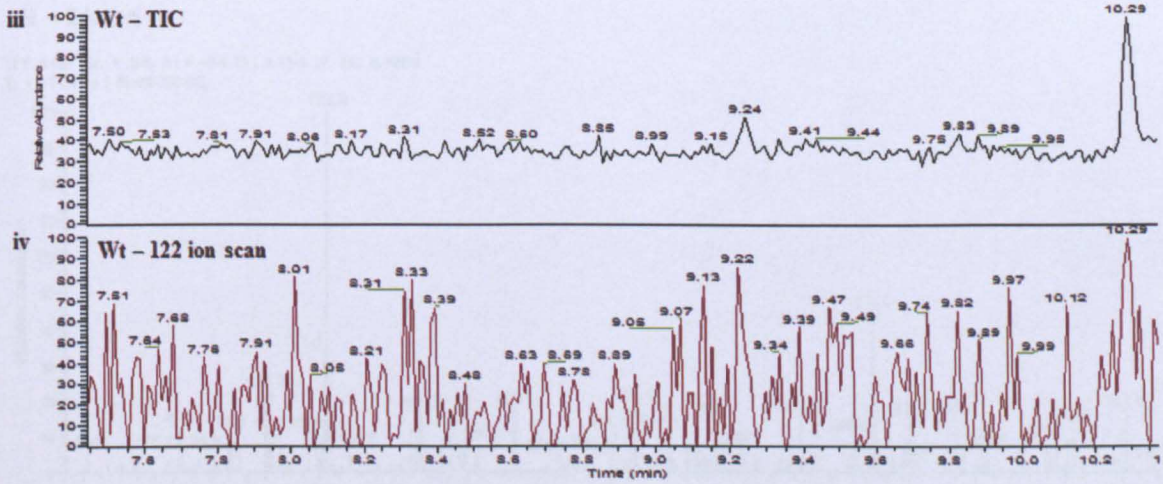
Northern analysis was performed using 10µg of total RNA extracted from transgenic tobacco lines. The northern blot membrane was hybridised with a full-length TXS cDNA probe generated by PCR amplification using the primers TaxSynF and TaxSynRns to give a 2.6 Kb fragment. The northern blot indicates varying levels of expression, with the highest TXS mRNA expression being in line 4 and 6.

Following confirmatory expression results from the northern blot, crude hexane extracts from lines 1, 3, 5, 6, 9, 10 and 12 were prepared and subjected to GC-MS analysis under the following conditions: the initial oven temperature was 160°C, this was ramped to 300°C at 10°/min after a 1 min hold (using Helium as carrier gas).

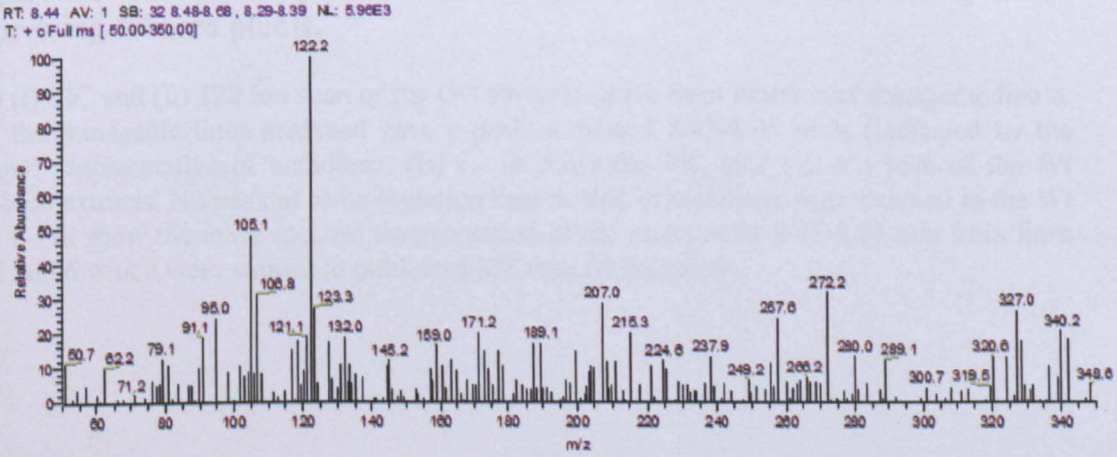
The 122 ion scan of the GC chromatogram from the generated TXS expressing lines showed a peak at around 8.43 - 8.45 mins in the extracts from the transgenic lines, having a mass spectrum very similar to that of taxadiene (Figure 5.5 A (i & ii)). The mass spectrum of this peak gave major ions at m/z 122 (100%), 121, 123, 107, 272, (Figure 5.5 C (i – iii)) consistent with the publised data for taxadiene (Kovacs *et al.*, 2007). Hexane extracts from wild type tobacco (Figure 5.5 b (i-iv)) was also

analysed as control and no peak characteristic of taxadiene was observed in the TIC or in the 122 ion scan. No further work has been done with these transgenic lines, but it might be useful quantifying and purifying the taxadiene produced for future work.

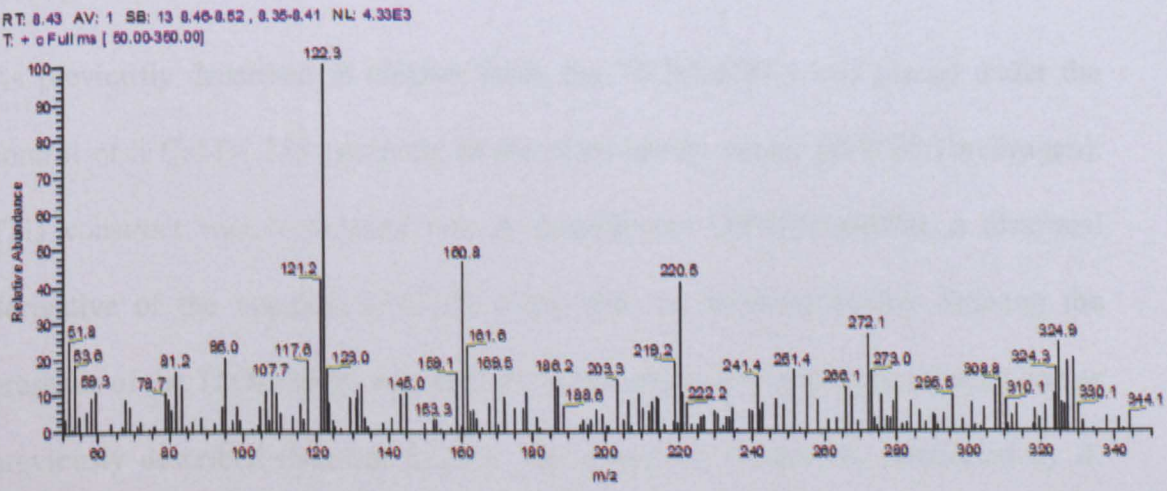




(C) i - Line 3



ii - Line 5



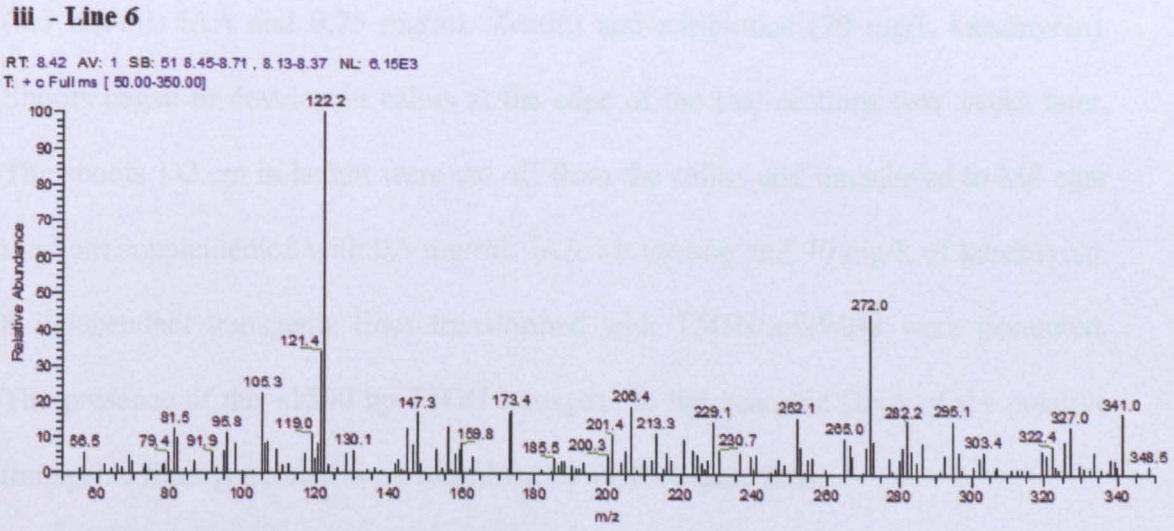


Figure 5.5 GC-MS analysis of crude hexane extracts from taxadiene synthase expressing tobacco plants.

(A) (i) TIC and (ii) 122 ion scan of the GC chromatogram from extratcs of transgenic line 6. All the transgenic lines analysed gave a peak at around 8.43-8.45 mins (indicated by the arrow) representative of taxadiene. (B) i - iv show the TIC and 122 ion scan of the Wt tobacco extracts. No peak at same retention time as that of taxadiene was observed in the Wt (C) i – iii show the mass spectral fragmentation of the peaks at Rt 8.43-8.45 min from lines 3, 5 and 6 which were similar to published MS data for taxadiene.

5.2.3 Generating tobacco lines expressing the taxadiene-5α-hydroxylase transgene

As previously described in chapter three, the T5OH cDNA was placed under the control of a CaMV 35S promoter in the plant binary vector pGWB8 (Invitrogen). This construct was introduced into *A. tumefaciens* GV3101/pMP90, a disarmed derivative of the nopaline GV3101 strain and the bacterial colony showing the presence of the T5OH insert was used for transformation of wild type tobacco line as previously described (Section 2.12.7). The tobacco leaf-sections transfected by *A. tumefaciens* GV3101 were placed on selection medium of MS agar with hormones

(0.5 mg/mL IAA and 0.75 mg/mL Zeatin) and antibiotics (70 mg/L kanamycin). Shoots began to develop in callus at the edge of the leaf-sections two weeks later. The shoots 1-2 cm in height were cut off from the callus and transferred to MS agar medium supplemented with 0.5 mg/mL IAA for rooting and 70 mg/L of kanamycin. 8 independent transgenic lines transformed with T5OH::pGWB8 were generated. The presence of the ~1500 bp T5OH transgene in the genomic DNA of the putative transgenic lines generated were examined by PCR (Figure 5.6).

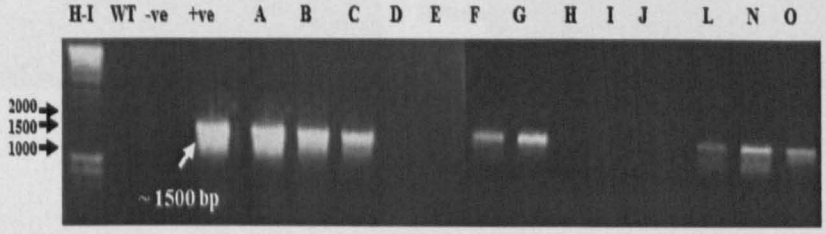


Figure 5.6 PCR analysis on shooting callus for the presence of T5OH fragment.

Gel picture showing the presence of PCR fragments of ~ 1500 bp in size. The PCR was performed using the primers 5Tax5alph and 3Tax5alph, amplifying the full-length of the T5OH cDNA from the gDNA of shooting callus. +ve = positive control (T5OH::pGWB8 plasmid), -ve = negative control (water).

To verify the expression of the T5OH transgene mRNA in the transgenic lines, a nothern blot analysis was carried out on the RNA extracted from the lines A, B, G, L, N and O (lines C and F died prior to further analysis). The northern blot results (Figure 5.7 ii) indicated that the transgenic lines L, N and O expressed the highest level of T5OH mRNA transcripts. An RT-PCR was performed on the RNA from Line O to verify the size of the mRNA transcribed, and the full-length PCR fragment of size ~1500 bp (Figure 5.7 iii) confirmed that no post-transcriptional modifications

such as "wrong" splicing had occurred *in planta*. Line L, N and O expressing high levels of T5OH mRNA were used for further work.

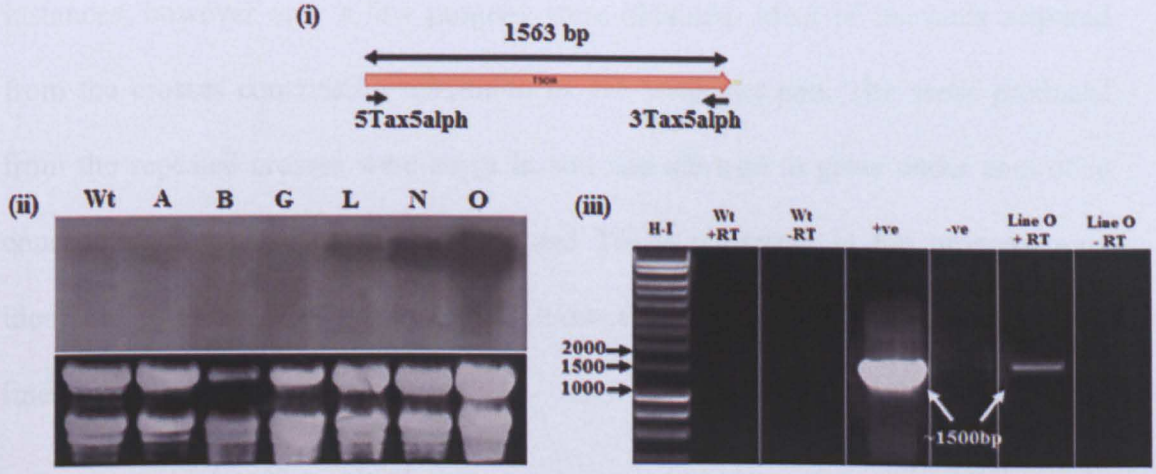


Figure 5.7 Expression analysis of T5OH transgene mRNA.

Northern analysis was performed using 10µg of total RNA extracted from the transgenic and Wt tobacco lines. The northern blot membrane was hybridised to cDNA probe as shown in (i) generated by PCR amplification of the full-length T5OH cDNA by the indicated primers. (ii) Northern blot results indicating the highest level of expression of the T5OH mRNA transcipt in lines N and O. Wt tobacco RNA was used as negative control. (iii) RT-PCR results using the primers 5Tax5alph and 3Tax5alph confirming the correct size of the expressed T5OH mRNA in line O. +RT = reverse transcriptase enzyme added, -RT = reverse transcriptase omitted during first strand cDNA synthesis, +ve = T5OH::pGWB8 as positive control.

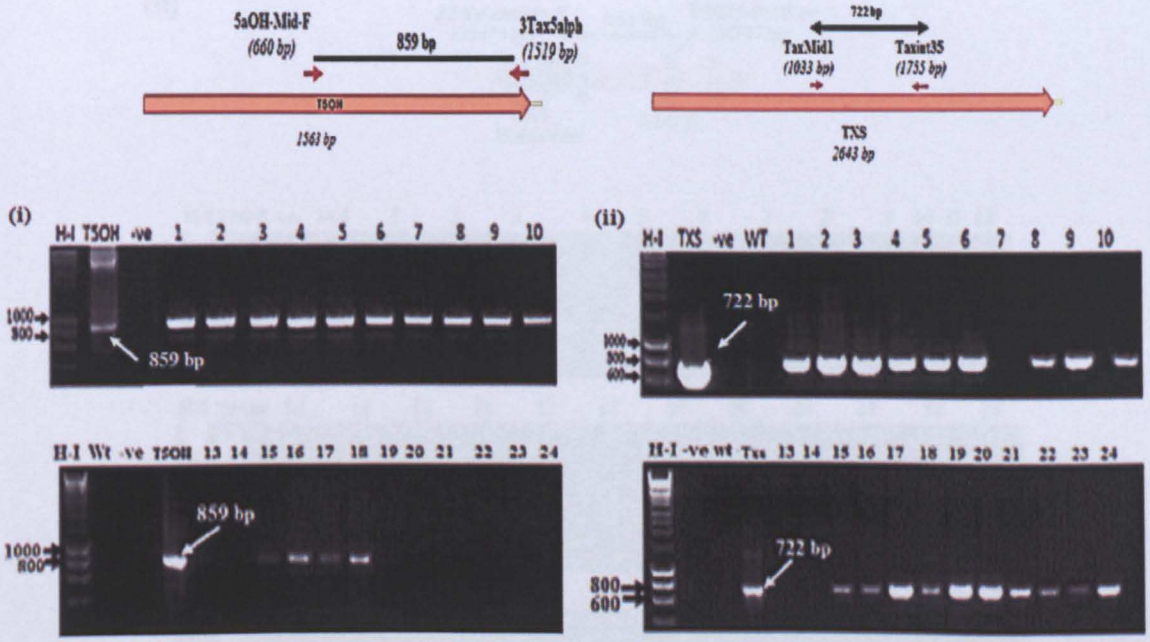
5.2.3 Analysis of progeny from crosses between T5OH and TXS expressing lines

With the aim of producing the first oxygenated taxane *en route* to Taxol, transgenic tobacco plants expressing the TXS transgene (Lines 1, 3 and 9) (Section 5.2.1) were cross-pollinated to lines L and O expressing the T5OH transgene. These T5OH transgene expressing lines however demonstrated a male and possible female sterile

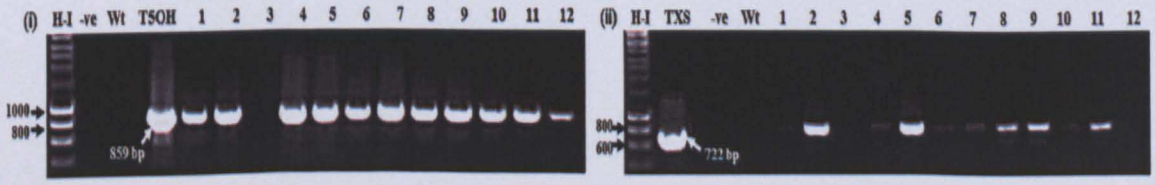
phenotype and typically showed no seed production, hence making crossing difficult. Taking this phenotype into account, the pollen from the taxadiene producing tobacco lines were used as the male gamete and gently rubbed on the stigma of the flowers from the T5OH expressing tobacco lines. These lines were crossed at multiple instances, however only a few progeny were obtained. Most of the pods acquired from the crosses contained a maximum of 3-4 seeds per pod. The seeds produced from the repeated crosses were sown in soil and allowed to grow under controlled conditions. The presence of the TXS and T5OH transgene in the progeny were identified by PCR on the genomic DNA extracted from the individual 2-3 weeks old lines.

PCR analysis showed the presence of both TXS and T5OH transgenes in most progeny obtained from the crosses between TXS 3♂ + T5OH L♀(Figure 5.8A). Only lines 2, 5, 8, 9 and 11 from crosses between TXS 9♂ + T5OH L♀(Figure 5.8 B) were found to be transgenic for both TXS and T5OH trangene. For the progenies obtained from crosses between TXS 1♂ + T5OH O♀(Figure 5.8C), all the lines were shown to carry the TXS transgene while 29 out of the 36 progeny tested showed the presence of the T5OH transgene. PCR analysis with the primers TaxMid1 and Taxint35 amplified a TXS fragment of size 722 bp in the progeny transgenic for TXS. The presence of T5OH in progeny from crosses between TXS3 + T5OH-L and TXS9 + T5OH-L were identified using the primers 5aOH-Mid-F and 3Tax5alph (Figure 5.8Ai & Bi) to give an amplified fragment of 859 bp. To check for the presence of T5OH in progenies from crosses between TXS1 + T5OH-O, primers CaMV35 S promoter and T5OH-R-Pst primer (Appendix I) were used to amplify a T5OH fragment of 864 bp (Figure 5.8C (ii)).

(A) Progenies from crosses between TXS 3 ♂ and T5OH L ♀



(B) Progenies from crosses between TXS 9 ♂ and T5OH L ♀



(C) Progenies from crosses between TXS 1 ♂ and T5OH O ♀

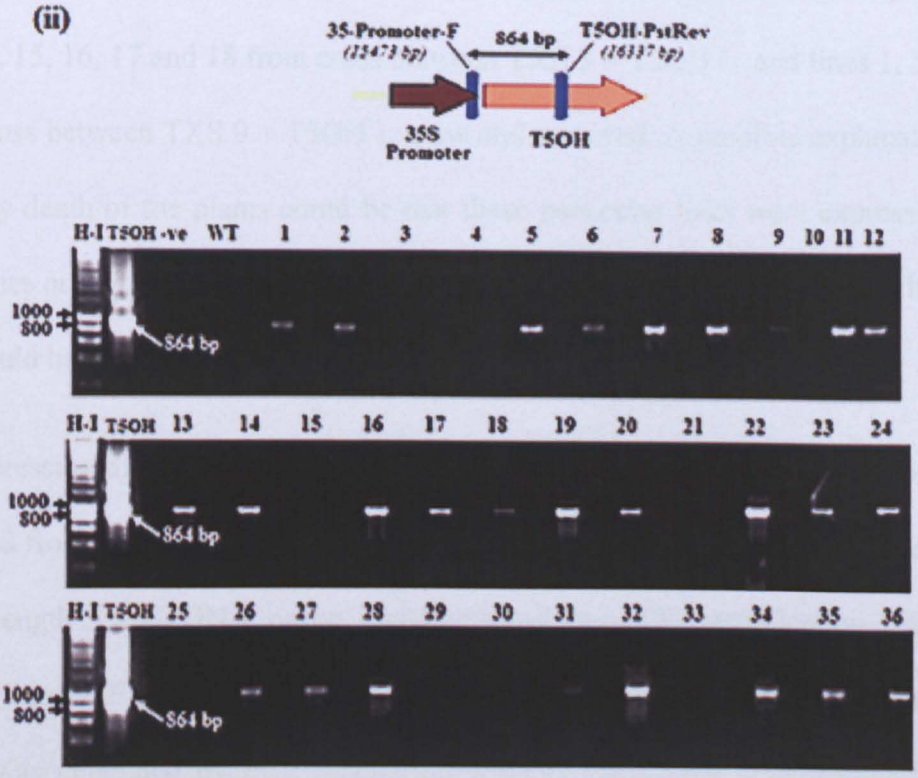


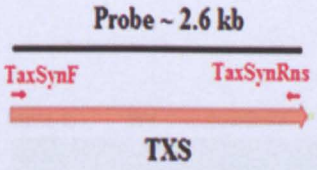
Figure 5.8 PCR analysis on progenies from crosses between T5OH and TXS expressing lines.

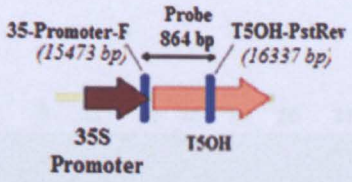
The positions of the primers used for each PCR and the expected fragment sizes were generated by Vector NTI. A(ii), B(ii) and C(i) show the result of PCR amplification of TXS transgene in gDNA extracted from progenies, using primers TaxMid1 and Taxint35 to give a fragment of 722 bp. A(i) and B(i) show the gel pictures of PCR amplification of the T5OH transgene from the gDNA in the progenies using primers 5aOH-MidF and 3Tax5alph to give a fragment of size 859 bp. C(ii) shows the PCR amplified fragments of 864 bp using the primers 35S-promoter-F and T5OH-PstRev. H1 = Molecular marker Hyperladder I, TXS = pBCTaxK plasmid DNA, T5OH = T5OH::pCR8 plasmid DNA, -ve = water

Following confirmation by PCR analysis, the lines carrying both TXS and T5OH transgenes were transferred to soil and allowed to grow for one month. Only lines 2, 3, 7, 10, 15, 16, 17 and 18 from cross between TXS 3 + T5OH L and lines 1, 5, 9, 11 from cross between TXS 9 + T5OH L grew and survived. A possible explanation for the early death of the plants could be that these particular lines were expressing the transgenes at high level or the accumulation of a novel compound produced in these lines could have been toxic to the plants.

The expression of the transgenes was examined by a northern blot assay on the RNA extracted from the leaves of each progeny carrying both TXS and T5OH transgenes. A full-length TXS cDNA probe was generated by PCR amplification using the primers TaxSynF and TaxSynRns and pBCTaxK as template. The cDNA probe for T5OH was generated by PCR amplification using the CaMV 35S-promoter-F and T5OH-PstRev primers and T5OH::pGWB8 as template. The results of the northern blot analysis as well as the positions of the primers used to generate the probes are shown in Figure 5.9.

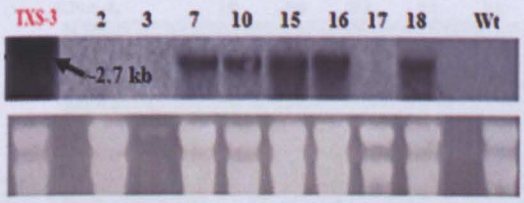
The results indicated that only line 18 from crosses TXS 3 + T5OH L (Figure 5.9 A) expressed both TXS and T5OH mRNA transcript while no progenies from crosses TXS9 + T5OH L (Figure 5.9 B) expressed the T5OH mRNA. The progenies from crosses between TXS 1+ T5OH O showed varying levels of expression of both TXS and T5OH transgenes in most of the lines. The size of the bands obtained from the progenies corresponded to those of the parent lines used for crossing, however the smeared RNA ladder (Figure 5.9 C) did not permit the band sizes to be confirmed.



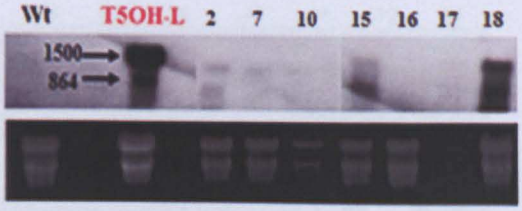


(A) - TXS 3 + T5OH L

(i)



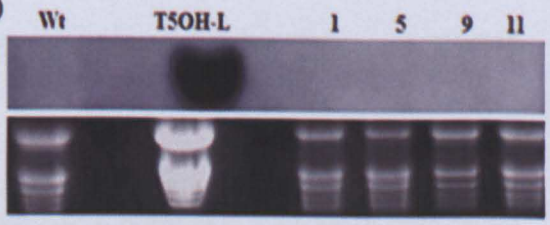
(ii)



(B) - TXS 9 + T5OH L

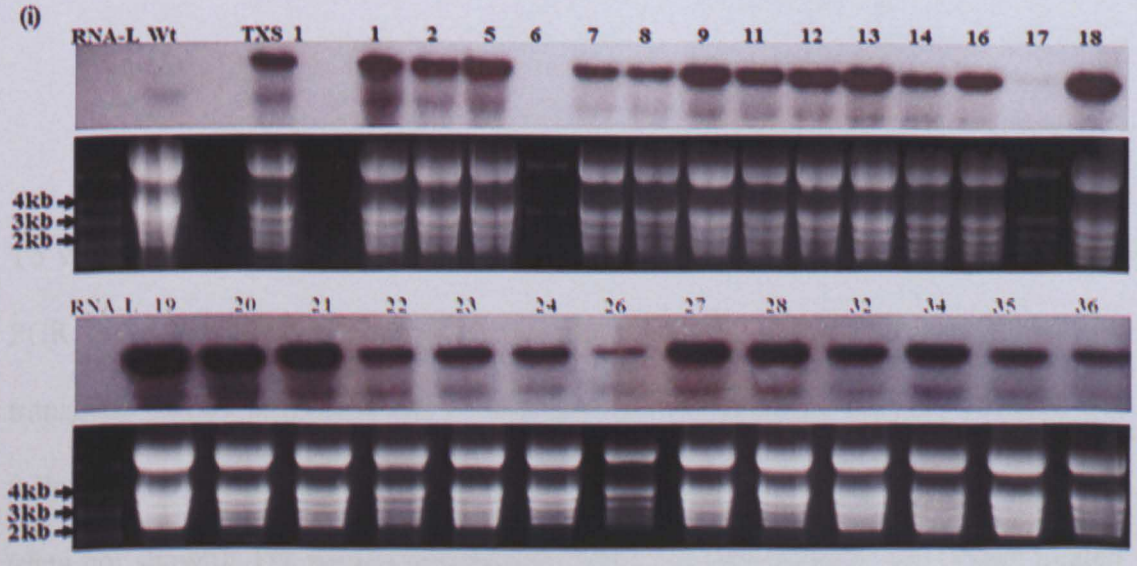
(i)

(ii)



(C) - TXS-1 + T5OH-O

(i)



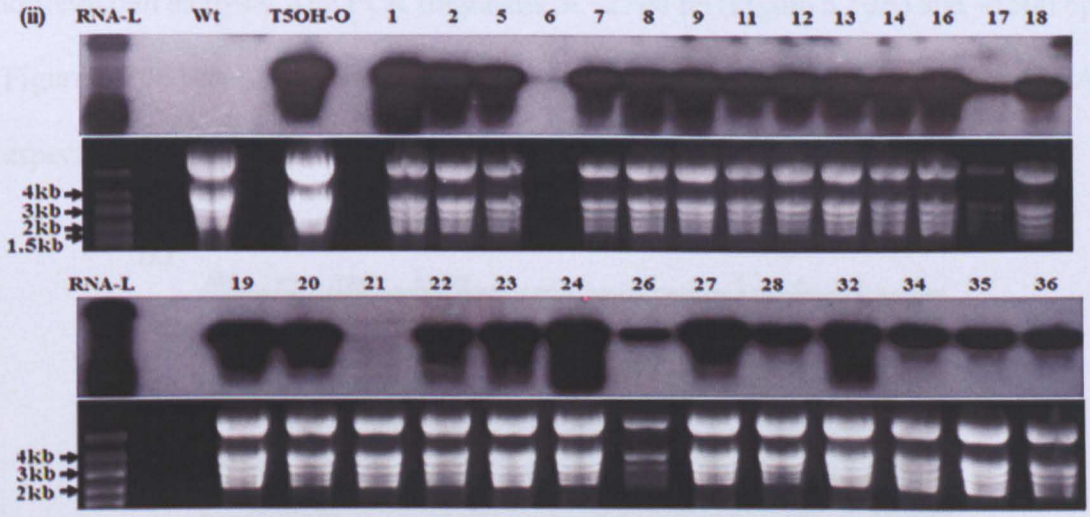


Figure 5.9 Northern blot analysis on progenies carrying the TXS and T5OH transgenes.

The expression of the T5OH and TXS mRNAs was verified by northern blot analysis. 10μg of RNA was used for each sample, wild type tobacco RNA was used as the control and the parent line used for each cross were used also as control. A schematic diagram of each probe used for membrane hybridisation is shown which are of sizes ~2.6 kb and 864 bp for TXS and T5OH respectively. (A) Line 18 from progenies obtained from crosses between TXS 3+ T5OH-L showed the expression of both transgenes while none of the progenies from crosses TXS9 + T5OH-L (B) showed expression of the T5OH mRNA. (C) Most of the progenies obtained from crosses between TXS1 + T5OH-O showed the expression of both transgenes.

To verify the band sizes of TXS and T5OH obtained from these progenies, an RT-PCR was carried out on the RNA samples using the full-length primers for the transgene cDNA amplification. A control RT-PCR reaction was also carried out following DNase treatment where the reverse transcriptase enzyme was not added (data not shown). The results indicated no DNA contamination in the DNase treated

samples. The RT-PCR results shown in Figure 5.10 correlated with those from the northern blot analysis. The PCR fragments of ~2700 bp (Figure 5.10A) and ~1500 bp (Figure 5.10B) amplified, confirmed the expression of the TXS and T5OH mRNA respectively.

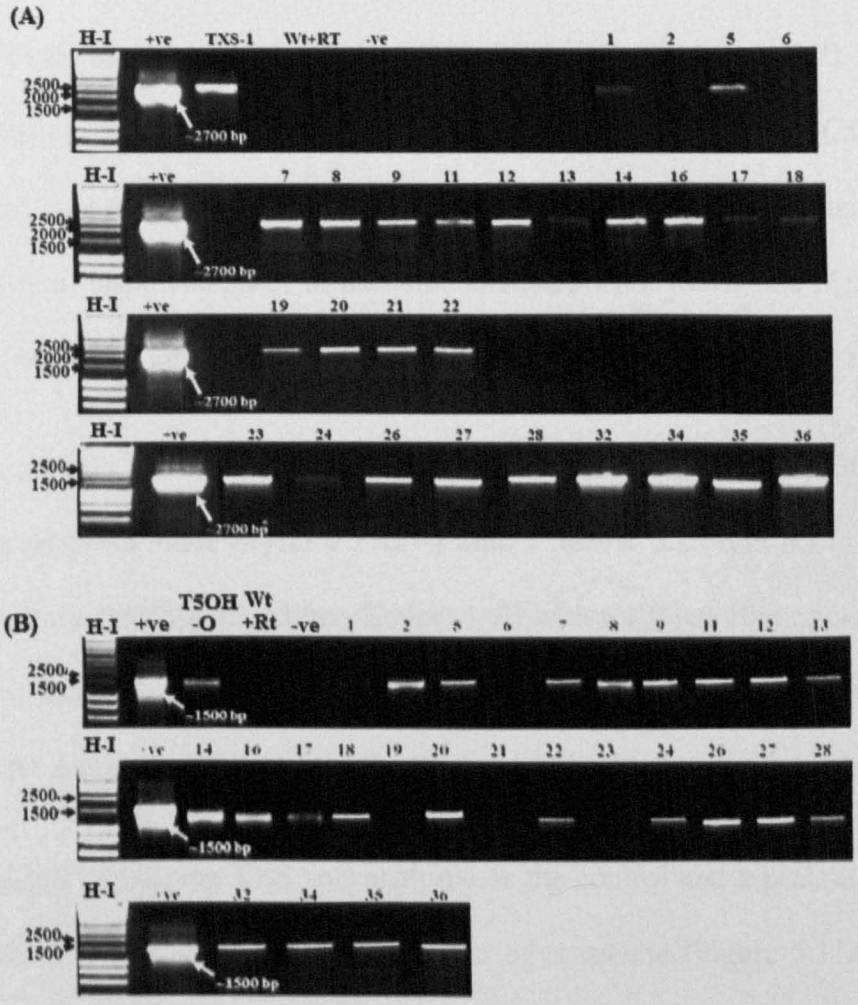


Figure 5.10 RT-PCR analysis of transgenic lines

The size of the TXS and T5OH mRNAs expressed in each of the transgenic lines were confirmed by RT-PCR using the full-length primers 5Tax5alph and 3Tax5alph for T5OH cDNA amplification and Txs-F and TxsRns for TXS cDNA amplification. A PCR fragment of size ~1500 bp was observed in the lines expressing the T5OH transgene and a PCR fragment of size ~2700 bp characteristic of the full-length TXS cDNA was observed in lines expressing the TXS. The RNA from parent lines TXS-1 and T5OH-O used for crossing was used as positive control. +ve = plasmid carrying the cDNA, -ve = water

5.2.4 GC-MS analysis of crude hexane extracts from lines expressing TXS and T5OH

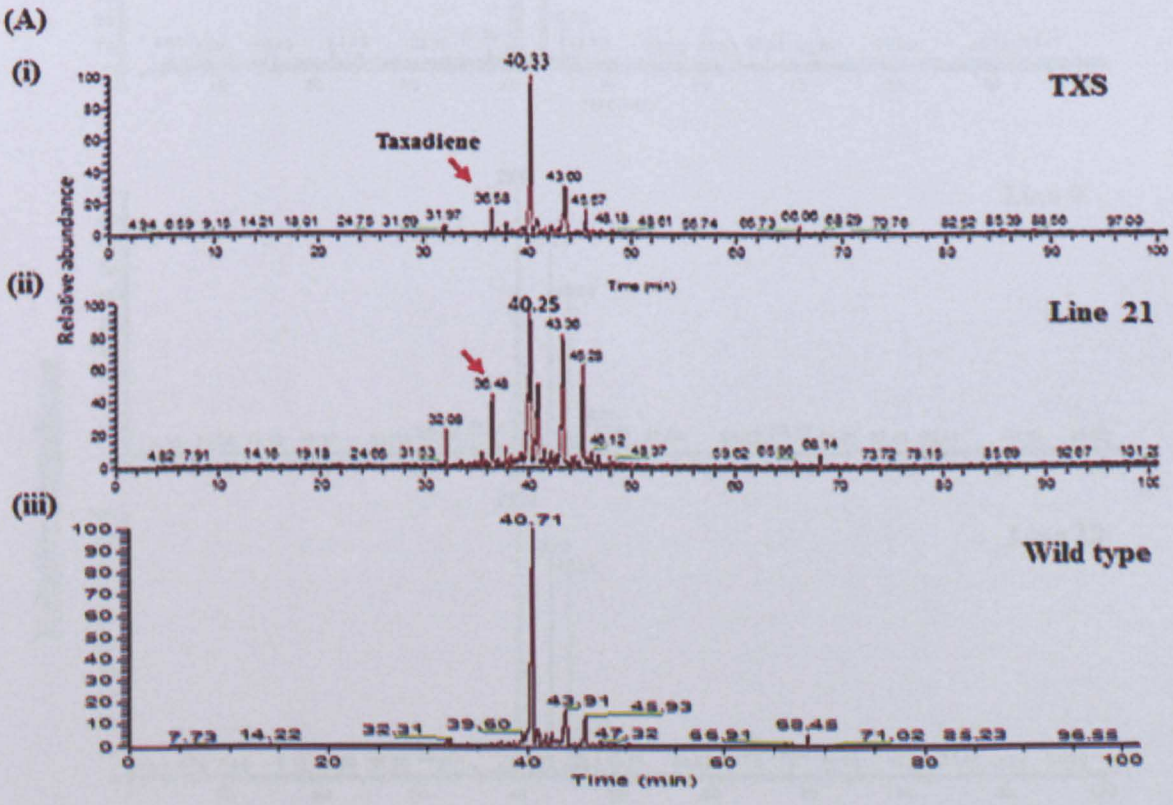
Following northern blot analysis and RT-PCR confirming the expression of the TXS and T5OH transgenes in the progenies obtained from the crosses, the leaves from these lines were extracted in hexane as described previously (Section 2.15.1) and 1 µl of each crude extract was subjected to GC-MS analysis (Section 2.15.2). The GC-MS conditions were slightly modified from those used by Rontein *et al.* (2008) who reported the presence of an oxidised taxane 5(12)-Oxa-3(11)-cyclotaxane (OCT) OCT instead of taxadiene-5α-ol in trichome exudates from transgenic *N. sylvestris* expressing both TXS and T5OH transgenes.

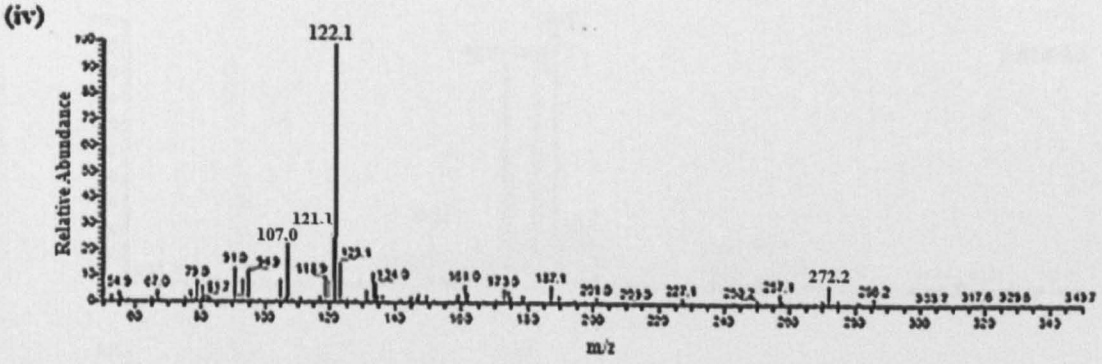
The crude extracts from the transgenic plants were injected with an AS3000 auto-sampler in split-less mode (injector 250°C) onto a 30m x 0.25 mm ID fused silica capillary column BP-5 (SGE, Milton Keynes UK) with a 1.0 µm film thickness. The initial oven temperature was 40°C, this was ramped to 100°C at 10°C/min (using helium as the carrier gas) followed by 3 °C/min to 360 °C and a 5-min hold.

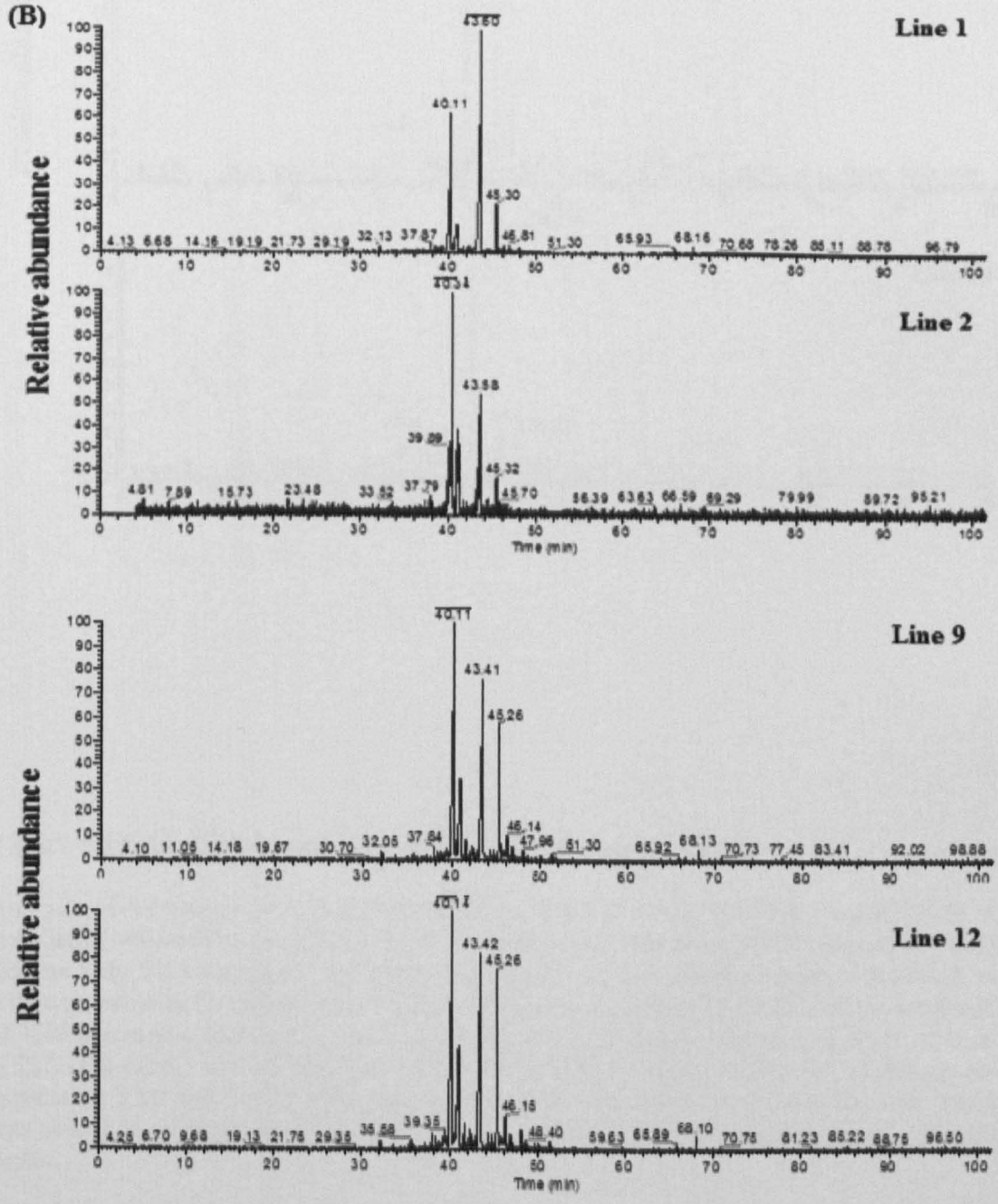
The parent line expressing TXS was analysed as the control and a peak at Rt 36.58 min (Figure 5.11A i) having a mass spectrum of taxadiene (Figure 5.11A iv) was observed which correlates with the Rt for taxadiene (36.86 min) from previously published data by Rontein *et al* (2008). No new product was detected in the parent line expressing only T5OH. The major ions 95, 105, 120 and 288, characteristic of taxadiene-5α-ol (Hefner *et al.,* 1996) as well as the ions 191, 149 and 288 for OCT were scanned for in the GC chromatogram obtained from each TXS+T5OH expressing plants. In transgenic lines co-expressing the TXS and T5OH transgenes,

the taxadiene was absent from the GC-MS profile (Figure 5.11 B), but taxadiene-5α-ol could not be detected either. Similar observations were previously reported in *N. sylvestris* coexpressing TXS and T5OH, whereby taxadiene was absent in the lines expressing both transgenes (Rontein *et al.*, 2008).

A hexane extract from the line 21 carrying both transgene TXS and T5OH but expressing only the TXS mRNA demonstrated the presence of taxadiene at Rt 36.48 min (Figure 5.11 ii). This data indicated that the taxadiene peak disappeared only when T5OH was being expressed, where taxadiene could have been used up to produce novel compound(s). It could also be postulated that the taxadiene-5α-ol was produced but rapidly turned over, explaining why it was not detected.




(iv)
122.1
121.1
107.0
272.2
Relative Abundance
m/z


(B)
Line 1
Line 2
Line 9
Line 12
Relative abundance
Relative abundance
Time (min)
Time (min)

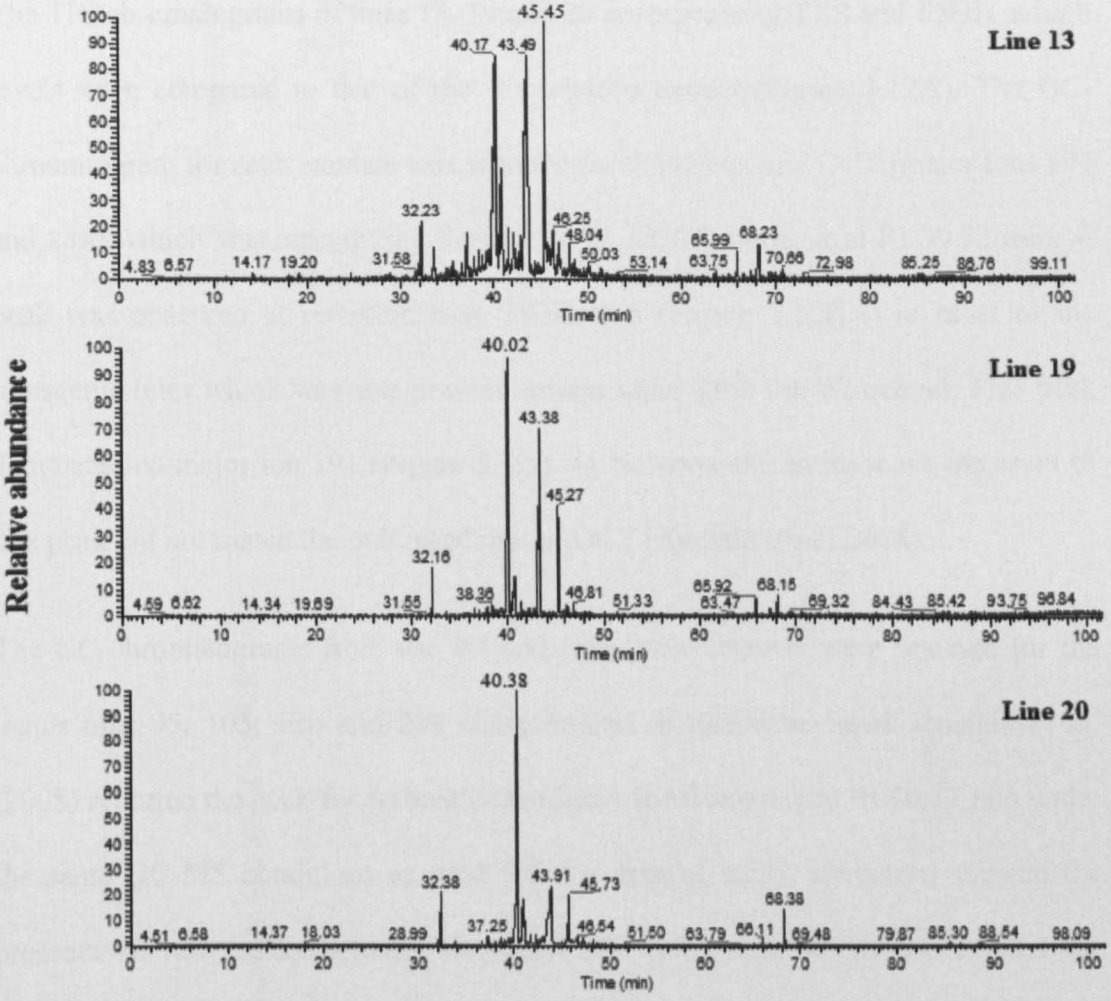


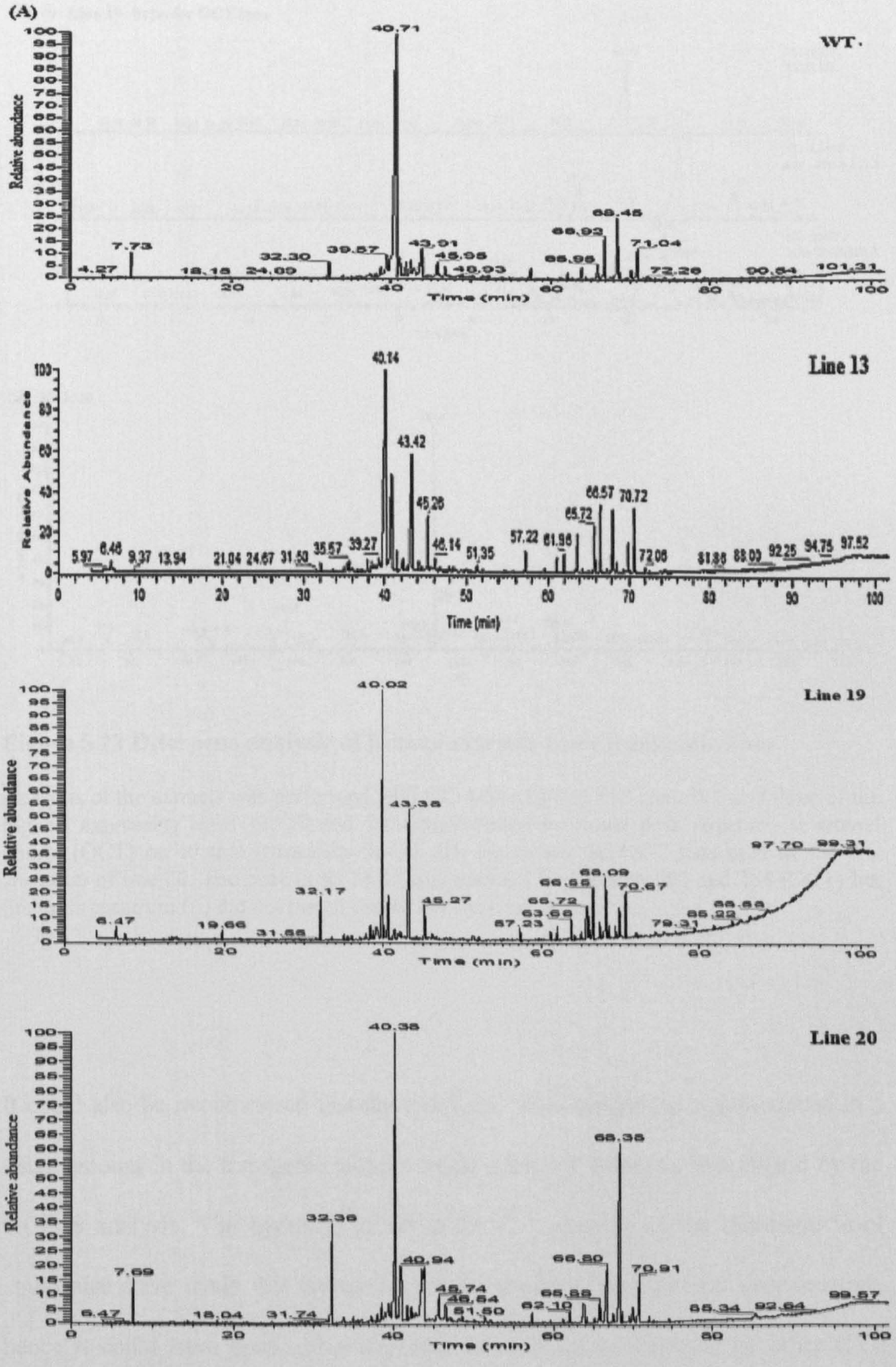
Figure 5.11 Analysis of hexane extracts from transgenic lines.

Analysis of the extracts was performed with GC-MS. (A) i & ii show the retention time of taxadiene (indicated by the arrows in i & ii) in the parent line and Line 21 respectively both express only TXS transgene. (iii) shows the absence of the taxadiene peak at Rt 36.58 in wild type extracts. (iv) Mass spectral fragmentation of the peak at Rt 36.48 min from extracts of TXS parent line and Line 21 expressing only TXS show major ions characteristic of taxa-4(5),11(12)-diene. (B) 122 ion scan of the GC-spectrum from exudates of progenies co-expressing TXS and T5OH. The peak characteristic of taxadiene with retention time 36.48 was absent in all these lines. The line name is indicated on the top right hand side of each figure.

The TIC chromatograms of lines 13, 19 and 20 co-expressing TXS and T5OH at high levels were compared to that of the Wt tobacco extract (Figure 5.12A). The GC-chromatogram for each exudate was scanned for the compound OCT (major ions 191 and 288), which was reported by Rontein *et al.* (2008) to occur at Rt 39.52 min. A peak was observed at retention time 39.40 min (Figure 5.12B i) in most of the transgenic lines which was also present around same Rt in the Wt extract. This peak demonstrated major ion 191 (Figure 5.12B ii), however the entire mass spectrum of this peak did not match the published one of OCT (Rontein *et al.*, 2008).

The GC-chromatograms from the Wt and transgenic extracts were scanned for the major ions 95, 105, 120 and 288 characteristic of taxadiene-5α-ol. Rontein *et al.* (2008) reported the peak for authentic taxadiene-5α-ol to occur at Rt 40.72 min under the same GC-MS conditions as used for this area of study. Ions scan showed the presence of new peaks between Rt 40.16 and 40.89 min, which was around the expected Rt for taxadiene-5α-ol. These peaks were scanned for the major ions of taxadiene-5α-ol, but this strategy failed to detect the compound. The peak for this oxidised taxane could have been overlaid by the major peaks at retention time 40.17 – 43.49 min (See appendix III).

The GC scans for ions of taxadiene-5α-ol for each transgenic line are shown in Appendix III. A few peaks present in the transgenic but absent in the Wt tobacco extracts were consistent in the extracts of most transgenic lines, however, none of them demonstrated mass spectrum of either taxadiene-5α-ol or OCT.


(A)
WT.
Relative abundance
Time (min)
40.71
7.73
4.27
18.16
24.09
32.30
39.57
43.91
45.95
49.93
65.95
68.93
68.45
71.04
72.29
90.64
101.31
Line 13
Relative Abundance
Time (min)
40.14
43.42
45.26
46.14
51.35
57.22
61.96
65.72
66.57
70.72
72.06
5.97
6.46
9.37
13.94
21.04
24.67
31.50
35.57
39.27
81.86
88.09
92.25
94.75
97.52
Line 19
Relative abundance
Time (min)
40.02
43.38
45.27
32.17
6.47
19.66
31.66
57.23
63.66
66.72
66.95
68.09
70.67
79.31
85.22
88.88
97.70
99.31
Line 20
Relative abundance
Time (min)
40.38
68.35
32.36
7.69
6.47
18.04
31.74
40.94
45.74
46.54
51.60
62.10
66.80
66.88
70.91
85.34
92.64
99.57

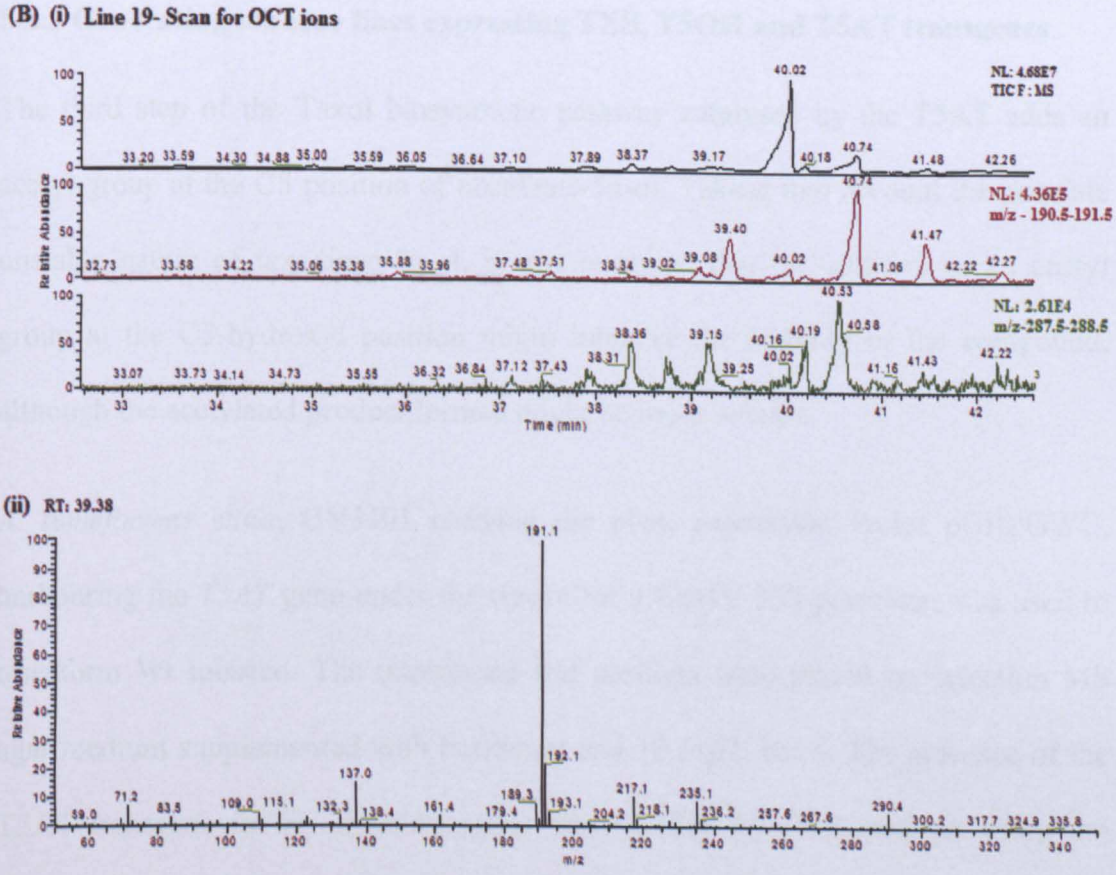


Figure 5.12 Diterpene analysis of hexane extracts from transgenic lines.

Analysis of the extracts was performed with GC-MS. (A) The TIC from WT and three of the highest expressing lines 13, 19 and 20, demonstrated no novel peak expected at around 36min (OCT) or 40 min (taxadiene-5α-ol). (B) (i) shows the OCT ions scan of the GC spectrum of line 20. The peak at Rt 35.87 was scanned for the ions 191 and 288 (OCT) but the mass spectrum (ii) did not match that of the OCT compound.

It could also be hypothesised that the taxadiene-5α-ol compound was produced in a small amount in the transgenic tobacco lines, hence its detection was missed by the GC-MS analysis. The hydroxyl group at the C-5 position of the taxadiene-5α-ol could also have made this compound highly unstable, volatile and very reactive, hence it could have been either degraded or used up as substrate by other CYP enzymes in the plants.

5.2.5 Generating tobacco lines expressing TXS, T5OH and T5AT transgenes

The third step of the Taxol biosynthetic pathway catalysed by the T5AT adds an acetyl group at the C5 position of taxadiene-5α-ol. Taking into account the possible unstable nature of taxadiene-5α-ol, it was reasoned that the addition of an acetyl group at the C5-hydroxyl position might improve the stability of the compound, although the acetylated product formed might be more volatile.

A. tumefaciens strain GV3101 carrying the plant expression vector pGBPGWC, harbouring the *T5AT* gene under the control of a CaMV 35S promoter, was used to transform Wt tobacco. The transfected leaf sections were placed on selection MS agar medium supplemented with hormones and 10 mg/L basta. The presence of the T5AT transgene in the shooting callus was verified by PCR analysis using the primers 5Tax5acyl and Acetyl-Rns (Appendix I), the positions of which are indicated in Figure 5.13 (A). Only one transgenic line was generated (Figure 5.12B) and this was transferred to soil and allowed to grow for one month.

A northern blot assay was carried out on the total RNA extracted from the transgenic line 1, to confirm the expression of T5AT mRNA. The probe used for membrane hybridisation was generated by PCR amplification using the same primers as indicated in Figure 5.13A. The results of the northern blot indicated that T5AT mRNA was expressed at a relatively low level (Figure 5.13C) in Line 1 when compared to the transgenic tobacco line expressing T5AT-ns previously generated for localisation of the protein.

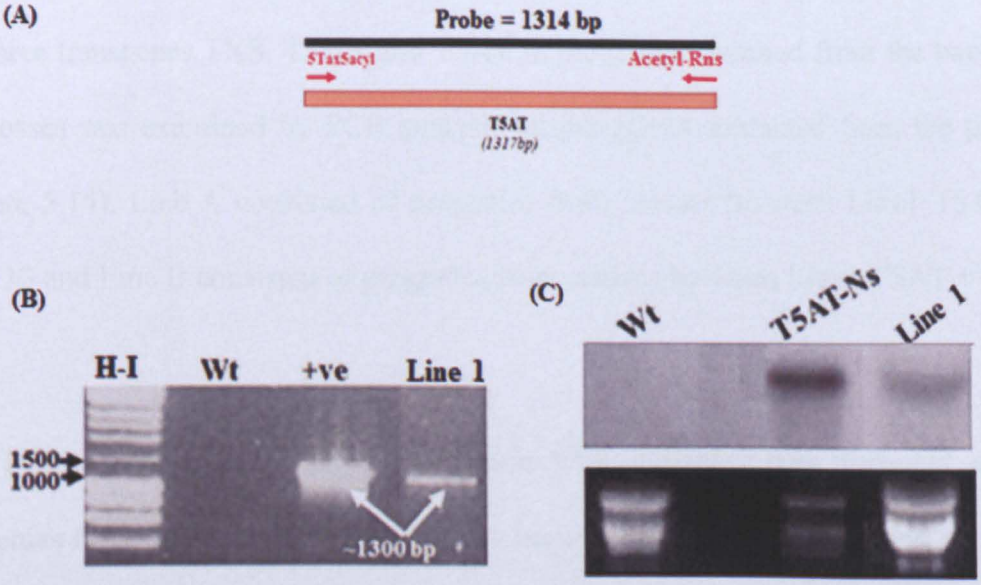


Figure 5.13 Generating transgenic tobacco plant expressing T5AT transgene.

(A) Schematic diagram showing the size of the probe used for membrane hybridisation during northern blot assay. The positions of the primers 5Tax5acyl and Acetyl-Rns used for PCR amplification are also indicated (B) Gel picture of PCR fragments of 1300 bp showing the presence of the T5AT cDNA in gDNA of line 1 while using the above primers. +ve = T5AT::pGBPGWC as template,. (C) Northern analysis was performed using 10µg of total RNA extracted from transgenic and Wt tobacco lines. The northern blot membrane was hybridised to cDNA probe (A). The first lane shows no expression in Wt tobacco line, second lane is the RNA from plants expressing the T5ATns::pGBPGWC and the third lane indicates the expression of T5AT in transgenic tobacco line 1.

In order to generate plants expressing the first three genes of the Taxol biosynthetic pathway, line 1 expressing T5AT mRNA transcripts was cross-pollinated to lines 13 and 20 co-expressing high levels of TXS and T5OH transcript mRNA. A male sterility phenotype was observed in lines 13 and 20, therefore the reciprocal crosses

were not possible. It was again observed that the number of seeds per pod were less than 2-3, and very few seed pods were obtained following crossing. The presence of the three transgenes TXS, T5OH and T5AT in progenies obtained from the two sets of crosses was examined by PCR analysis on the gDNA extracted from the plants (Figure 5.14). Line A consisted of progenies from crosses between Line1 T5AT + Line 13 and Line B consisted of progenies from crosses between Line1 T5AT + Line 20.

The PCR analysis results shown in Figure 5.14, indicated that three out of 10 progenies from line A and 20 out of the 31 lines tested from Line B carried all three transgenes TXS, T5OH and T5AT. In some cases, the amplified bands following PCR were very faint as in the case of line A6 which showed a pale band for T5OH transgene (Figure 5.14 B). The lines showing the presence of all three transgenes (indicated by arrows in Figure 5.14) were used for further analysis.

A northern blot assay was carried out on the total RNA extracted from the transgenic lines carrying all three transgenes. The primers used for the preparation of probes used for each membrane hybridization and the sizes of each probe are indicated in Figure 5.15. The results of the northern blot showed that progenies 1 and 10 from Line A and 18, 23 and 19 from Line B expressed all the three transgenes, with line B19 expressing the highest level of the TXS, T5OH and T5AT mRNA transcript (Figure 5.15).

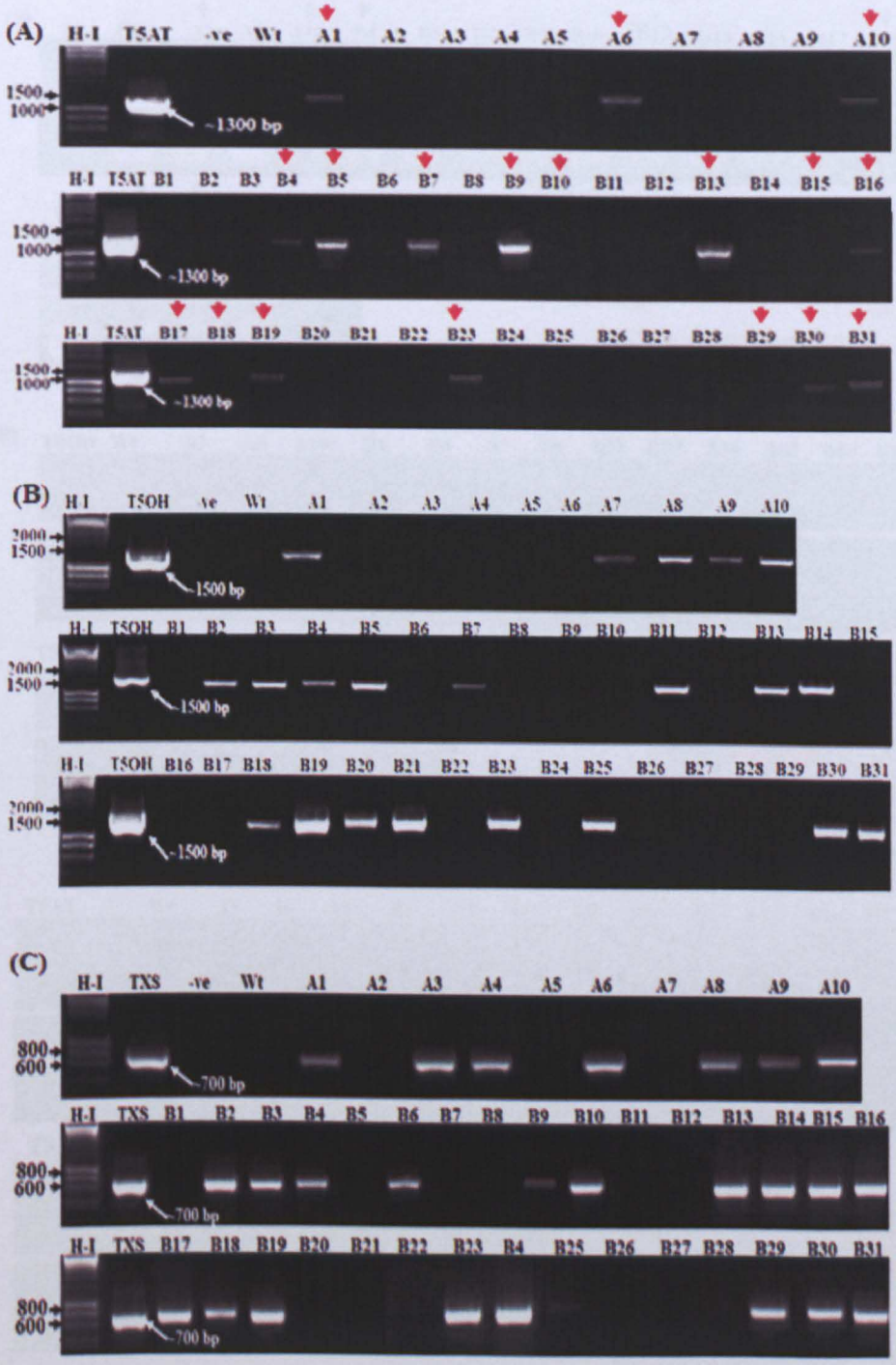


Figure 5.14 PCR analysis on progenies from lines A and B.

(A), (B) and (C) show the presence of the DNA fragments amplified to show the presence of the transgenes T5AT, T5OH and TXS respectively. The primers used were as follows: 5tax5acyl and 3tax5acyl for amplification of full-length T5AT cDNA of ~1300 bp, 5tax5alph and 3tax5alph for amplification of full-length T5OH cDNA of ~ 1500 bp and internal primers TaxMid1 and Taxint35 for amplification of fragment size ~700 bp showing the presence of the TXS transgene. The lines carrying all three transgenes are indicated by the red arrows. The positive controls T5AT, T5OH and TXS indicate the PCR fragments obtained from the parent line in each case, –ve = water as negative control and H-I = Hyperladder I as marker.

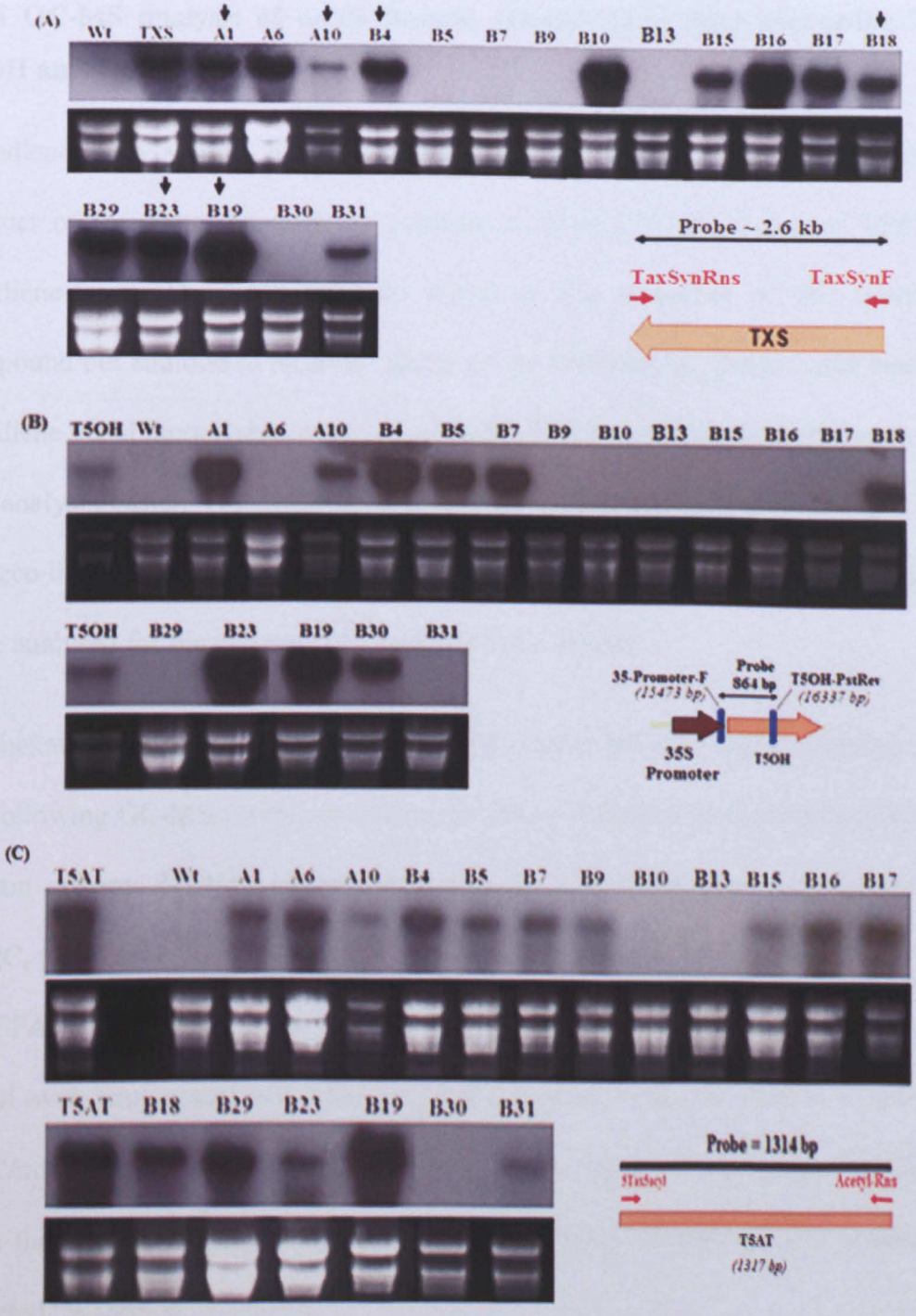


Figure 5.15 Northern blot assays on progenies from crosses between Line 13 and T5AT and Line 20 and T5AT.

The northern assay was carried out on 10µg of RNA. The full-length cDNA probes were used in the cases of TXS and T5AT while for T5OH probing, the CaMV35S promoter - T5OH-PstRev PCR amplified probe was used for membrane hybridisation. The primers used for the preparation of each probe and the size of the probes are shown next to each respective northern blot. The lines expressing the three transgenes are indicated by the black arrows. The RNA from the parent lines used for crossing was used as a positive control for the respective northern blot and Wt tobacco RNA used as negative control.

5.2.6 GC-MS analysis of crude hexane extracts from lines expressing TXS, T5OH and T5AT.

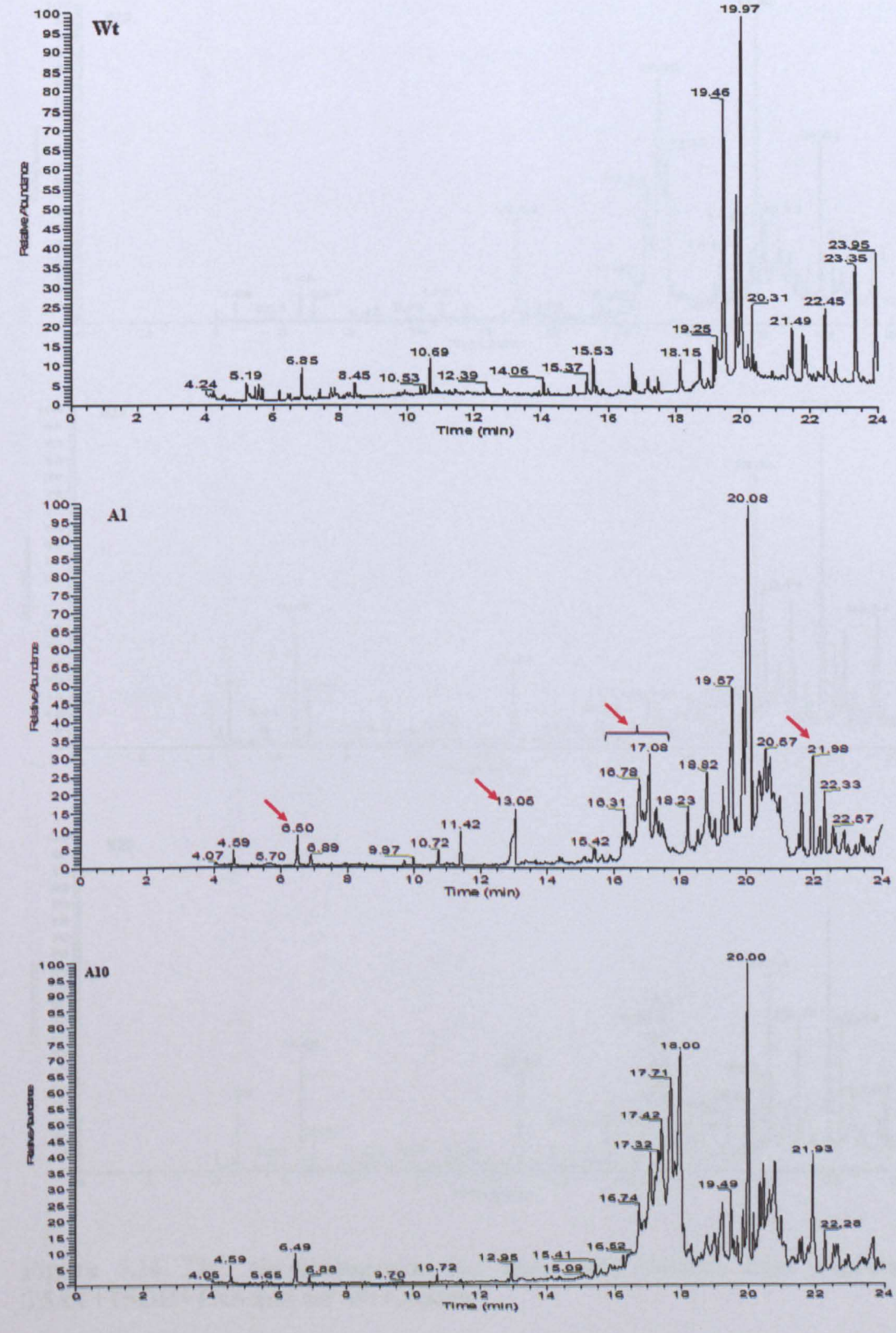
Taxadiene-5α-yl acetate, the first acetylated compound and the third downstream product of the Taxol biosynthetic pathway is formed by the action of T5AT on taxadiene-5α-ol. There has been no report on the properties of this acetylated compound but addition of an acetyl group on the C5-hydroxyl group could make the taxadiene-5α-ol more stable, hence its accumulation might make its detection by GC-MS analysis easier. The crude hexane extracts (Section 2.15.1) from the leaves of tobacco lines A1, A10, B18, B19 and B23 expressing TXS, T5OH and T5AT mRNA were analysed for the presence of taxadiene-5α-yl acetate.

The hexane extracts were analysed using both a polar and a semi-polar column under the following GC-MS conditions: Using the 30m x 0.25 mm ID fused silica capillary Zebron column ZB-1HT (Semi-polar column): The initial oven temperature was 100°C, this was ramped to 330°C at 10°C/min using helium as carrier gas. Using the ZB-FFAP (polar) column the samples were run after the following conditions: The initial oven temperature was 50°C with a 2 min hold. This was ramped to 250°C at 10°C/min with a 3 min hold. The latter GC-MS conditions were slightly modified from the previously published conditions under which taxadiene-5α-yl acetate was detected (Walker *et al.*, 1999).

The TIC obtained from the samples under each of the GC-MS conditions were compared to that of the wild type tobacco extracts, the results obtained using the polar column ZB-FFAP are shown in Figure 5.16B. A few new peaks (indicated by red arrows in Figure 5.16B) which were absent in the Wt tobacco extract were

consistently found in each transgenic line. The GC-chromatogram obtained from each transgenic was scanned for the major ions 105, 120, 255 and 270, characteristic of taxadiene-5α-yl acetate (Walker *et al.*, 2000), however, the mass spectrum obtained from these new peaks did not correspond to that of the acetylated taxane.

It was therefore speculated that the taxadiene-5α-yl acetate was not present or could have been lost during hexane extraction due to its possibly high volatility. The failure to detect this compound could also mean that the substrate taxadiene-5α-ol was not present initially or that it was quickly turned over or converted to other compounds by the plants endogenous enzymes, hence the acetylated compound could not be detected in the extracts. The peaks present in the transgenics but absent from the Wt extracts might indicate an accumulation of an increased level of compounds in the plants as response to stress following the expression of the three transgenes.


Wt
Relative Abundance
Time (min)
19.97
19.46
23.95
23.35
20.31
22.45
21.49
19.25
18.15
15.53
10.69
6.85
5.19
4.24
8.45
10.53
12.39
14.06
15.37
A1
20.08
19.57
17.08
20.57
21.98
18.82
16.78
22.33
13.06
16.31
18.23
22.67
11.42
6.50
4.59
6.89
9.97
10.72
4.07
5.70
15.42
A10
20.00
18.00
17.71
17.42
17.32
21.93
19.49
16.74
22.28
6.49
4.59
4.05
5.65
6.88
9.70
10.72
12.95
15.41
15.09
16.62

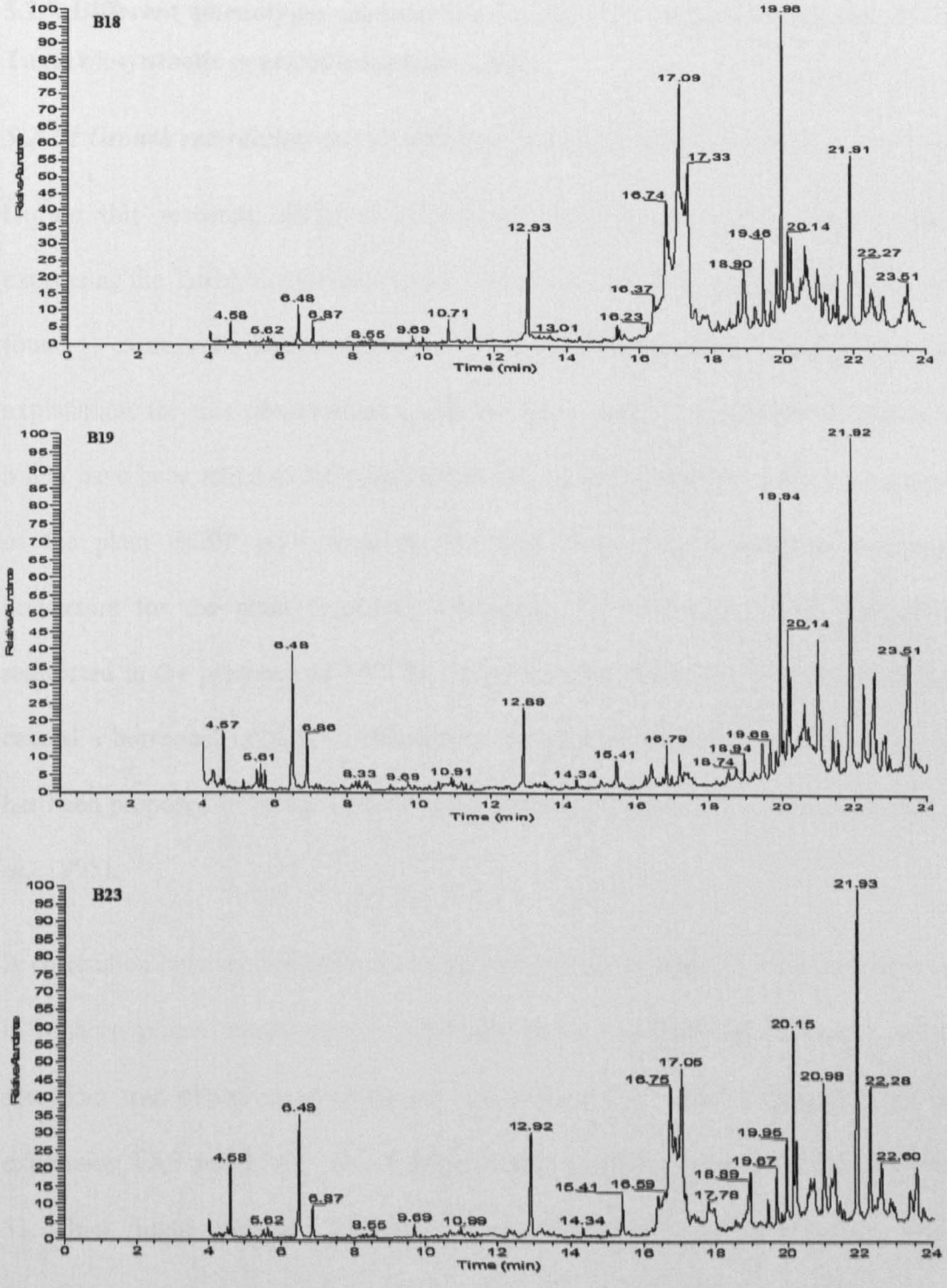


Figure 5.16 TIC chromatograms for transgenic tobacco lines expressing T5AT+T5OH+TXS and for Wt tobacco.

Analysis of the extracts was performed with GC-MS. TIC obtained from the hexane extracts from each line were compared. The peaks indicated by the red arrows were absent or very much reduced in the WT sample but present in most of the transgenics expressing the three transgenes.

5.2.7 Different phenotypes associated with the heterologous expression of the Taxol biosynthetic genes in transgenic plants.

5.2.7.1 Growth retardation and dwarfism in transgenic tobacco plants.

During this research, different phenotypes were observed in transgenic lines expressing the Taxol biosynthetic genes. Line 4 (Section 5.2.2) generated which was found to express the highest level of TXS died at a very young stage. A possible explanation for this observation would be that a high accumulation of taxadiene might have been lethal to the plant. It can also be postulated that a large percentage of the plant GGPP pool required for other isoprenoid biosynthesis producing precursors for the plant regulators cytokines, GA and ABA, might have been redirected in the presence of TXS for the production of taxadiene. This might have caused a hormonal imbalance obstructing the normal development of the plant, as has been proposed to be the case when phytoene synthase is over-expressed (Fray *et al.*, 1995).

A correlation between reduction in height and transgene mRNA levels was observed in tobacco plants constitutively expressing the Taxol biosynthetic genes. Severe dwarfism was observed in transgenic tobacco lines 15 and 31 (Figure 5.17) co-expressing TXS and T5AT. This reduced stature was much more pronounced in line 31, which might suggest a high level of taxadiene production in this plant. These dwarf phenotypes could have also been the result of a hormonal imbalance in these plants due to the redirection by TXS of the GGPP normally used for the biosynthesis of the plant hormones GA and BRs involved in stem elongation and cell division. Line B30 (Figure 5.17) expressing only the T5OH mRNA was also found to have reduced internode lengths when compared to the wild type. A possible explanation

for this phenotype could be that the constitutive production of an active enzyme form of T5OH in this line might have been acting upon substrates which are involved in the GA or BR biosynthetic pathway, thereby disrupting these biosynthetic pathways. As a result, less GA or/ and BRs are formed, affecting cell division and stem elongation in the plant. Abnormal pigmentation was also observed in leaves from most transgenic lines expressing the Taxol biosynthetic genes, which could have been attributed to a reduced level of available GGPP for the synthesis of the carotenoids or of the phytol side chain of chlorophyll in the plants. Further work would be needed to test these hypotheses.

Figure 5.17 Altered plant growth with constitutive expression of the Taxol biosynthetic genes.

Reduced internode lengths were observed in lines B15, B30 and B31. The wild type (Wt) tobacco is shown as control. Line B19 constitutively expressing T5OH and TXS were found to be sterile and possessed narrower leaves when compared to those of the Wt. The expression of the respective mRNAs are indicated above each plant, for instance Line B15 co-expresses TXS and T5AT mRNA.

5.2.7.2 Pollen viability and germination test from plants expressing the Taxol biosynthetic genes

As mentioned in section 5.2.3, the T5OH expressing lines were all found to be sterile and were almost male-infertile due to reduced pollen number, viability, and impaired pollen release in most cases. The purpose of this area of study was to determine the viability of the tobacco and tomato pollen of lines expressing T5OH and TXS respectively. To test pollen viability in these lines, the pollen grains were stained with Alexander's solution (see Appendix II) for around 10 minutes before being observed under a light microscope. Photographs of the pollen from each plant are presented in Figure 5.18 and Figure 5.19 all of which demonstrate clear differentiation of aborted and non-aborted pollen grains. In all tested samples, non-aborted pollen grains stained magenta-red and aborted pollen grains stained pale blue-green.

It was found that nearly 99% of the pollen grains from Wt tobacco were viable, while the percentage of viable pollen grains from the parent line expressing T5OH mRNA as well as the progenies expressing T5OH, was decreased to less than 25% (Figure 5.18, Column A). The percentage of pollen in taxadiene producing tobacco plants were reduced to ~92%, however, unlike the severe male-sterile T5OH lines, the pollen grains from the taxadiene producing plant were still mostly viable. It was also noticed that the anthers from the T5OH expressing lines usually contained fewer pollen grains than that of the wild type tobacco plant and only produced about 20% of the pollen per anther of the wild type. This dramatically reduced number of pollen grains might account for much of the reduced male fertility in these T5OH expressing lines.

The pollen tube elongation was examined by growing the pollen on MS media overnight as previously described (Section 2.17). The pollen grains were incubated with Alexander's stain for 10 mins and observed under a light microscope. Most of the wild-type pollen tubes grew while the germination rate of T5OH expressing plants was considerably less. < 20% of the viable pollen from these transgenic lines were found to have germinated (Figure 5.18, Column B), indicating that pollen tube growth is very likely to have contributed to the male sterility in the T5OH expressing lines.

The redirection of GGPP in *yellow flesh* tomato mutant for the production of taxadiene (Kovacs *et al.*, 2007) has resulted into severely reduced seed formation, generally producing parthenocarpic fruits, or fruits with less than ten viable seeds. As part of this research, a pollen viability and germination test was carried out on the pollen from these *r* tomato plants producing taxadiene, referred to as pBCTaxK in Figure 5.19. The number of viable pollen per anther were considerably less in the transgenic lines when compared to wild type tomato plants (Figure 5.19, column A). The germination rate of the pollen tubes from these transgenic tomato lines was found to be almost nil, however the presence of seemingly viable pollen was observed (Figure 5.19, column B). This may suggest that the pollen development was affected at the pollen tube development stage which is regulated by the phytohormone brassinosteroid (Clouse and Sasse 1998). Furthermore, previous crosses of Wt pollen to taxadiene producing tomato lines have led to the production of infrequent or parthenocarpic fruits. This might suggest that taxadiene accumulation in these transgenic lines might have been toxic to the plant, thereby affecting the female reproductive organs as well as the pollen formation.

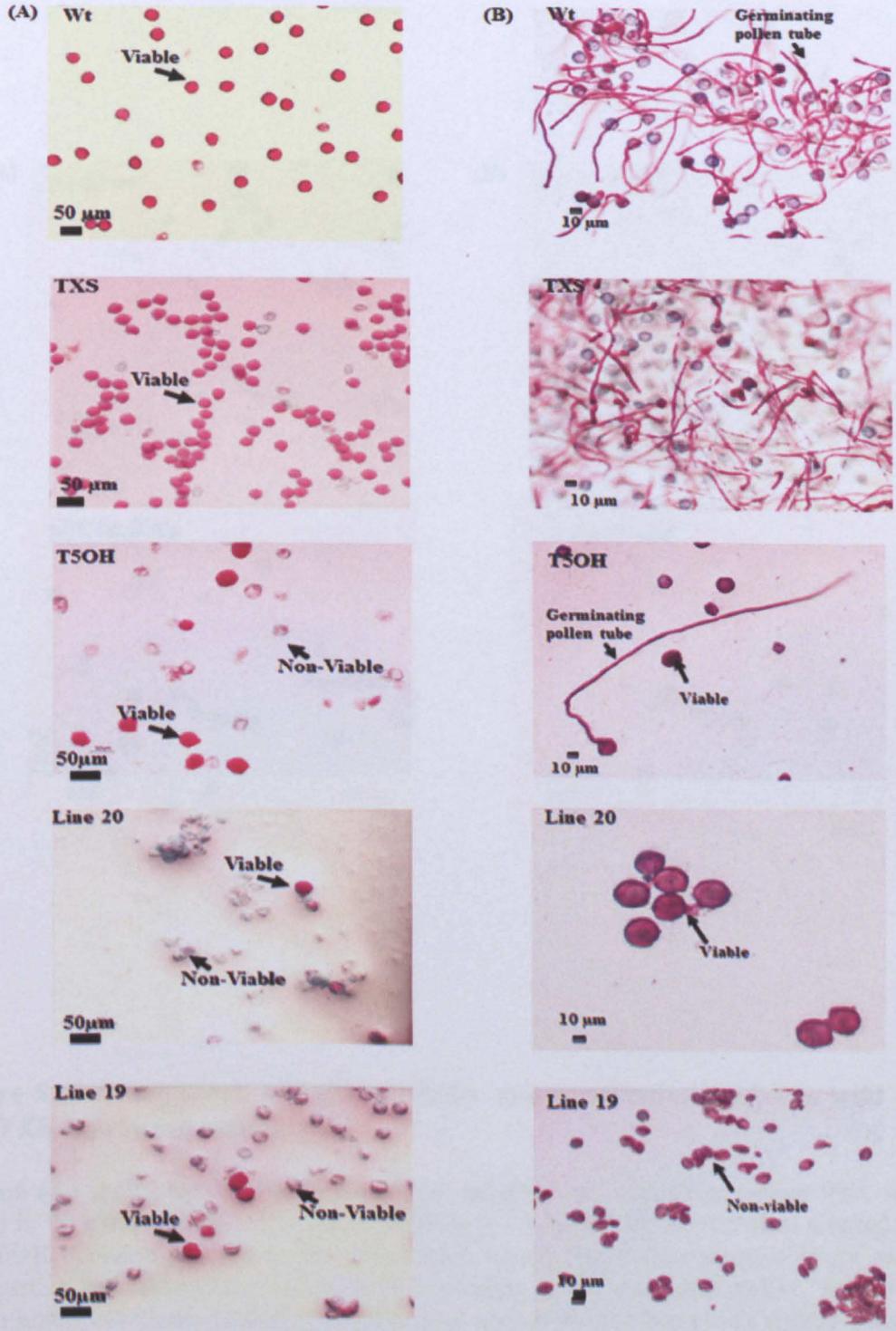


Figure 5.18 Comparison of pollen among the transgenics and wild-type tobacco.

Column A show the results of the pollen viability test and column B shows the results of the pollen germination test. The viable and non-viable pollen are indicated. Alexander's stain coloured the aborted pollen grains blue-green, and non-aborted pollen grains stain magenta-red. Lines 19 and 20 expressed both TXS and T5OH transgenes, Wt refers to wild type tobacco pollen, TXS refers to pollen from TXS expressing plants and T5OH refers to pollen from T5OH expressing plants. Scale bar 10µm, 50µm.

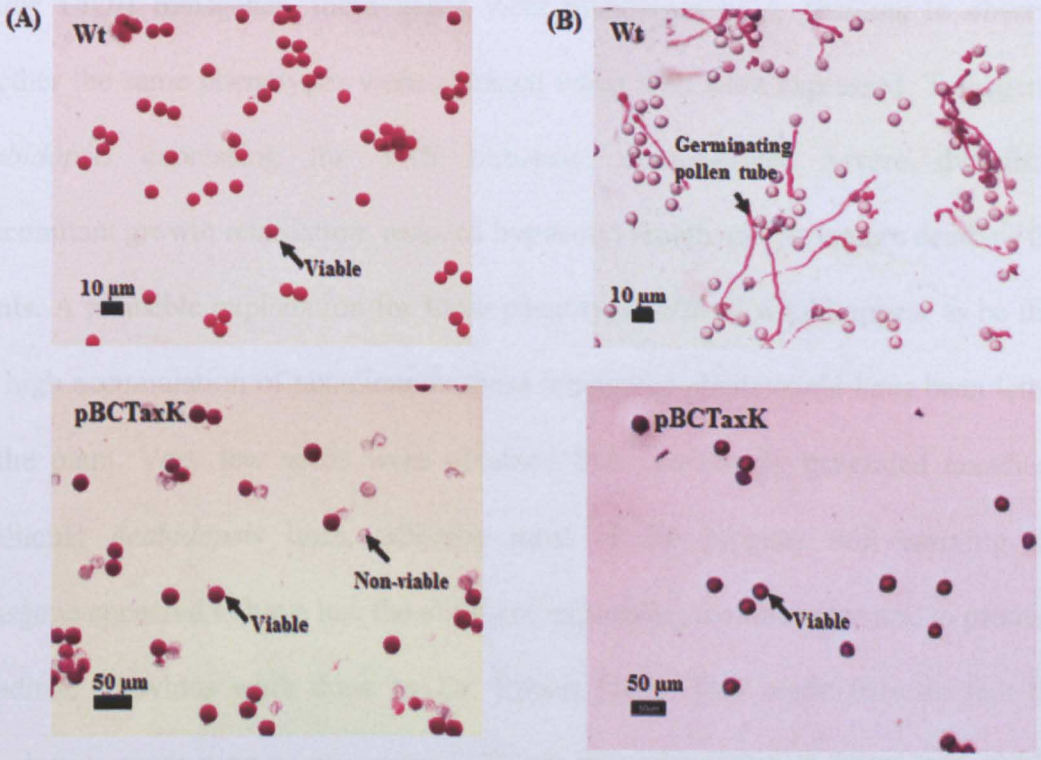


Figure 5.19 Comparison of pollen viability and germination between wild type and TXS expressing tomato.

Column (A) shows the results of the pollen viability test, indicating around 99% viable pollen in Wt tomato while ~50% viable pollen in TXS expressing tomato plant referred to as pBCTaxK. Column (B) shows the germination among the pollen where a larger rate of germination had taken place within the Wt but none in the pBCTaxK pollen. The aborted pollen grains have been stained blue-green, and non-aborted pollen grains stained magenta-red by Alexander's stain. Scale bar, 10µm and 50µm.

5.2.7.3 Generating transgenic Arabidopsis carrying the Taxol biosynthetic genes.

Taking into account the phenotypes observed in tobacco plants expressing the TXS and/or T5OH transgenes, these genes were transferred to *A. thaliana* to observe whether the same phenotypes were obtained when they were expressed. Transgenic *Arabidopsis* expressing the TXS transgene demonstrated severe dwarfism, concomitant growth retardation, reduced hypocotyl length and premature death of the plants. A plausible explanation for these phenotypic effects would appear to be that the high accumulation of taxadiene in these transgenic plants could have been lethal to the plant. Very few seeds were obtained from previously generated taxadiene producing *Arabidopsis* lines, whereby most of the progeny still carrying the transgene appeared to have lost the ability of expressing the transgene and to produce taxadiene (Previous work done by Dr. Rupert Fray). This might indicate that the constitutive production of an active TXS enzyme might have redirected the GGPP required for phytohormone biosynthesis and carotenoids pathway, for the formation of taxadiene. As a result, this could have caused a hormonal imbalance, thereby affecting pollen and seed development. It could also be hypothesised that the high level accumulation of taxadiene in these progeny might have been toxic to the plant, hence only lines expressing the lowest level of TXS or those not producing taxadiene survived.

Transgenic *Arabidopsis* expressing the T5OH transgene mRNA were generated (data not shown), however these lines showed no phenotypic effect as observed in tobacco lines expressing this transgene. This could suggest that T5OH enzyme might be acting upon a substrate present in tobacco but not in *Arabidopsis*. As part of this area of study, more Wt *Arabidopsis* plants have been transformed with the constructs

carrying either T10BOH or T5AT and transgenics were generated, however no further work or analysis have been carried out at this time.

5.2.8 Yellow flesh tomato as an alternative and better system for the production of taxanes.

5.2.8.1 GC-MS analysis of the crude hexane extracts from r mutant tomato expressing TXS

The level of taxadiene accumulation in transgenic tobacco lines generated during this study was found to be relatively low when compared to other compounds accumulated in the plant. The presence of major peaks representative of plant isoprenoids (e.g phytol, cholesterol and other tobacco steroids) in the GC-MS chromatograms from these lines indicated that other terpenes might have dissolved in the extraction buffer used, which would have made the purification of taxadiene difficult. Furthermore, most of the progenies obtained from the parent line previously provided, appeared to have lost their ability to produce taxadiene.

Yellow flesh tomato mutants were considered as an alternative method to co-express TXS and T5OH to produce taxadiene-5α-ol. A mutation in the *Psy* gene in this variety of tomato makes it an attractive system in which to express TXS, where the GGPP normally used for carotenoid synthesis has been previously demonstrated to be rerouted for the production of substantial amount of taxadiene in these mutants (Kovacs *et al.*, 2007). Another advantage of using the *r* mutant lines is that the low carotenoid level in these tomatoes would allow the facile extraction of any taxane produced from these fruits due to reduced carotenoid level dissolving in the

extraction buffer. The use of *r* tomato mutant as a heterologous system for taxadiene production might also provide a cheap and easily extractable source of this important compound.

The GC-MS results from *yellow flesh* tomato lines constitutively expressing TXS are shown in Figure 5.20. The 122 ion scan confirms the presence of the taxa-4(5),11(12)-diene at Rt 12.33 min and its isomer taxa-4(20),11(12)-diene at Rt 11.90 min. The mass spectrum of both taxadiene isomers (Figure 5.20 ii and iii) correspond to the previously published data (Williams *et al.*, 2000a) with taxa-4(5),11(12) being principally produced at Rt 12.33 min and a very low-level coproduction of taxa-4(20),11(12)-diene detected at Rt 11.90 min.

Efforts have been made to transform the tomato cotyledons from the taxadiene producing plants with a hygromycin resistant construct carrying the T5OH cDNA (pGHPGWY), but no transgenic lines were generated. The elite lines producing the highest amount of taxadiene have been previously crossed with Wt *yellow flesh* tomato to generate more tomato line producing taxadiene, however fruits from these crosses were infrequent and parthenocarpic, which can be partially explained by the failure of pollen tube growth as described in the previous section. It was therefore reasoned that generating *r* tomato plants expressing T5OH for crossing with the taxadiene producing plants might not be successful, due to low fertility issues.

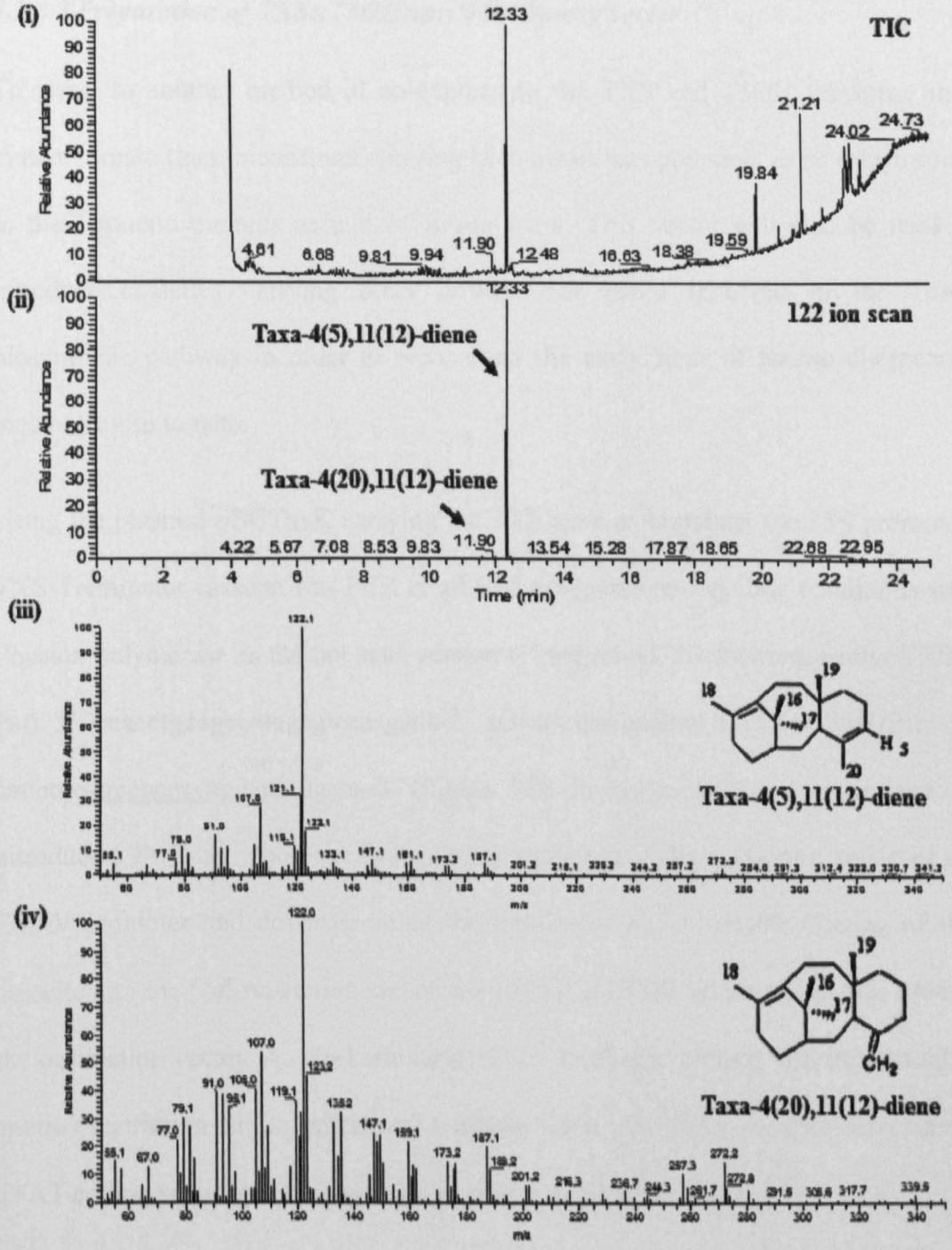


Figure 5.20 GC-MS analysis of crude hexane extracts from *r* tomato mutant expressing TXS.

(i) GC profile of TIC in extracts from transgenic fruits showing a prominent peak at Rt 12.33 min. (ii) 122 ion scan of the TIC showing the two peaks at Rt 11.90 min and 12.33 min. (iii) Mass spectrum of the peak observed at Rt 12.33 min shows the major ions characteristic of taxa-4(5),11(12)-diene and (iv) mass spectrum of the peak at 11.90 min showing the presence of the isomer taxa-4(20),11(12)-diene, these data match the published spectrum for the respective taxadiene isomers.

5.2.8.2 Preparation of TXS::T5OH::pGWB8 binary vector

To resort to another method of co-expressing the TXS and T5OH transgene in *r* mutant tomato lines, a construct carrying both genes was prepared, to be transformed in these tomato mutants as part of future work. This vector will also be used to introduce cassettes carrying other downstream genes involved in the Taxol biosynthetic pathway in order to reconstruct the early steps of taxane diterpenoid metabolism in tomato.

Using the plasmid pBCTaxK carrying the *TXS* gene as template, the 35S promoter-TXS-Terminator cassette was PCR amplified using standard cycling conditions with Phusion polymerase as the hot start version (Finnzymes), the forward primer TXSF-Pst1 5′- ccac**ctgcag**ccatggagtcaaagattc-3′ and reverse primer TXSRev-Sbf1/Pst1 5′-catt**ctgcaggcgcgcc**aggaaacagctat-3′ (Figure 5.21 i). These primers were designed to introduce a ***Pst*I** restriction site immediately upstream of the ATG start codon of the CaMV promoter and downstream of the terminator end to enable cloning of this cassette into the *Sbf*I restriction site of the T5OH::pGWB8 vector which was used as the destination vector. An *Asc*I site (underlined in reverse primer) was introduced as another restriction site to introduce, at a later stage, the T5AT cassette carrying the T5AT coding sequence along with its promoter and terminator.

The ~3400 bp PCR fragment (Figure 5.21 ii) carrying the 35S+TXS+Terminator cassette obtained following PCR amplification was gel purified using a MiniElute Kit (QIAGEN) and the resulting amplicon was digested with *Pst*I restriction enzymes overnight at 37°C. The digested sample was separated on an agarose gel (Figure 5.21 iii) to give the expected bands of 3193 bp and 245 bp. The bigger fragment of 3193

bp carrying the 35S::Txs and 245 bp PCR fragment being the terminator were each gel extracted and purified.

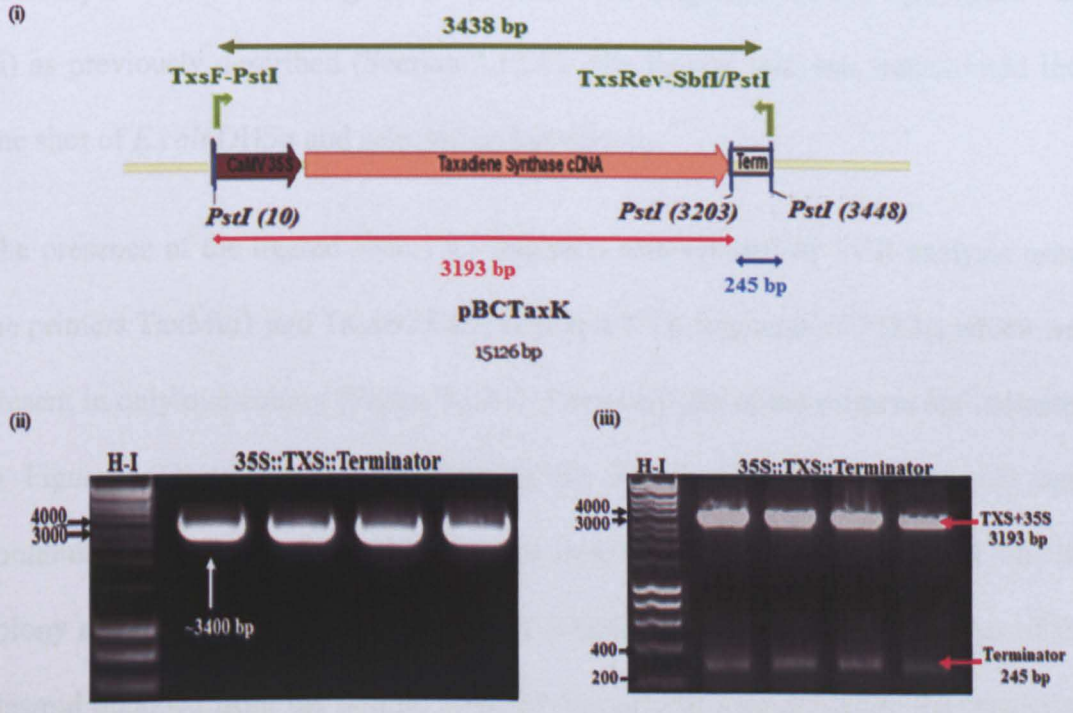


Figure 5.21 Preparation of the 35S+Txs and terminator fragments.

The schematic diagram of the pBCTaxK vector generated by Vector NTI (Invitrogen) is shown in (i) and include the position of the primers TxsF-PstI and TxsRev-SbfI/PstI used for the PCR amplification of 35S promoter-TXS-Terminator cassette. The *Pst*I restriction sites within this cassette and fragment sizes are also indicated (ii) A PCR fragment of around 3400 bp was amplified using the indicated primers confirming the amplification of the 35S-TXS-Term cassette. (iii) Fragments obtained following *Pst*I digest of the 35S-TXS-Term cassette gave a band of ~ 3400 bp carrying the 35S+TXS cDNA and a band of ~240 bp for the terminator.

Large scale preparation of the plasmid T5OH::pGWB8 was carried out following the protocol described in Section 2.12.5. The plasmid T5OH::PGWB8 carrying an *Sbf*I restriction site upstream of the 35S promoter, was linearised by a restriction digest reaction using the *Sbf*I enzyme. The digested product was then dephosphorylated as previously described (Section 2.12.1) and gel purified. The linearised T5OH::pGWB8 was then ligated to the 35S+TXS fragment of 3193 bp (Figure 5.21 iii) as previously described (Section 2.12.1). The ligated mix was transformed into one shot of *E.coli* DH5α and selected on kanamycin.

The presence of the ligated 35S::TXS fragment was verified by PCR analysis using the primers TaxMid1 and Taxint35 amplifying a TXS fragment of 722 bp which was present in only one colony (Figure 5.22 i). The positions of the primers are indicated in Figure 5.22 ii. This colony carrying the insert was restreaked on LB agar containing kanamycin. After 24 hours, a rapid "boil prep" was prepared on this colony as previously described in Section 2.12.6. Following RNase treatment of the plasmid obtained from the boiling prep, the vector was digested using the restriction enzymes *Hin*dIII and *Pst*I to verify the presence and orientation of the 35S::TXS fragment in the pGWB8 vector. The expected band sizes following the respective digests are shown in Figure 5.22ii, and the results of the restriction digests (Figure 5.22 iii) matched the expected fragment sizes indicating that the 35S+TXS was inserted in the correct orientation in the pGWB8 vector.

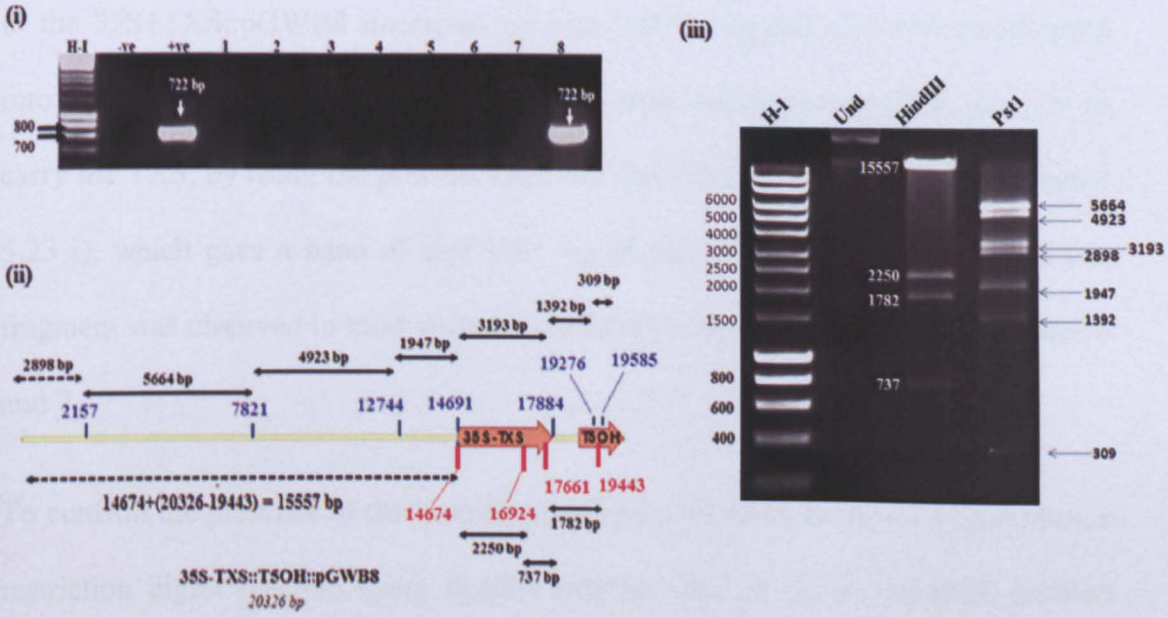


Figure 5.22 Preparation of 35S+TXS::pGWB8 vector

(i) Gel picture showing the presence of the TXS insert in colony 8 following PCR amplification using the primers TaxMid1 and Taxint35. pBCTaxK plasmid DNA was used as +ve control and water as the –ve control (ii) The Schematic diagram of the 35S+TXS::pGWB8 construct was generated by Vector NTI. The *Hin*dIII and *Pst*I restriction sites are indicated in red and blue respectively. The fragment sizes produced following restriction digests are indicated above the arrows (iii) The results of the *Hin*dIII and *Pst*I restriction digests of plasmid 8 are shown to match the expected fragment sizes, indicating that the 35S+TXS fragment was inserted in the correct orientation in the pGWB8 plasmid. The undigested plasmid 8 (Und) was used as control, and H-I refers to the marker Hyperladder I.

Following confirmatory results from the digests, the 35S+TXS::pGWB8 plasmid from colony 8 was purified and linearised by *Sbf*I digestion, which cuts the vector downstream of the TXS coding sequence. The terminator of size 245 bp was ligated to the 35S+TXS::pGWB8 linearised plasmid and the ligated mix was transformed into *E.coli*. Kanamycin resistant *E.coli* cells were confirmed by PCR analysis to carry the TXS, by using the primers TaxMid2 and TxsTev-SbfI/PstI primers (Figure 5.23 i), which gave a band of size 1281 bp (Figure 5.23 ii). The presence of this fragment was observed in most colonies identified with strongest bands in colonies 6 and 7.

To confirm the presence of the terminator and its orientation in these two colonies, a restriction digest reaction using *Hin*dIII enzyme was set up on plasmids purified from these two colonies. *Hin*dIII enzyme cuts within the terminator at 18084 bp and towards the end of the 35S-TXS cassette at 17661 bp to give a fragment of 423 bp (Figure 5.23 i). The plasmid from colony 8 carrying the 35S+TXS without its terminator was also digested using the *Hin*dIII enzyme to be used as control. Figure 5.23 iii shows the fragments obtained following the restriction digest whereby the ~400 bp terminator fragment (Figure 5.23iii) was present in plasmids 6 and 7 but absent in colony 8 (35S+TXS) lacking the terminator.

The plasmid 7 was sequenced using the primers shown in Figure 5.24 and the sequencing results indicated that a new *Sbf*I site was created next to the *Asc*I site downstream the TXS-terminator.

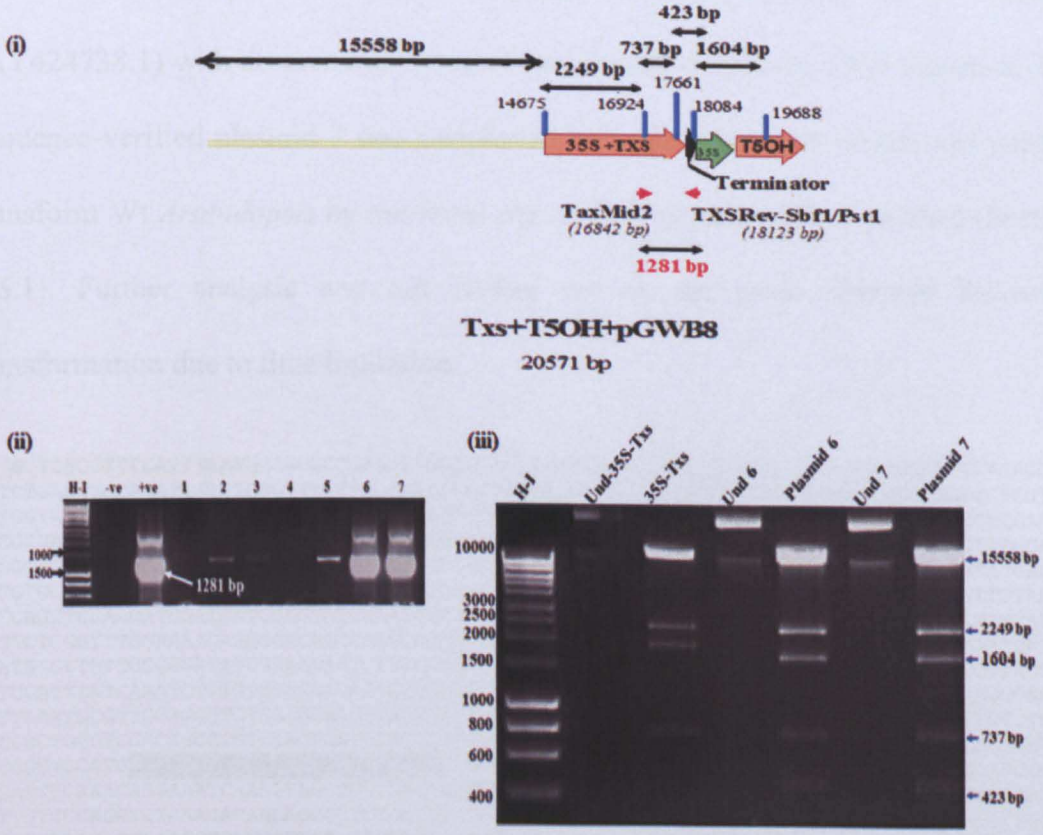


Figure 5.23 Preparation of the TXS::T5OH::pGWB8 construct

(i) Schematic diagram of the pGWB8 contruct carrying 35S+TXS+terminator cassette was generated by Vector NTI. The *Hind*III restriction sites are indicated in blue, the primer sites of TaxMid2 and TXSRev-SbfI/PstI are indicated and the PCR fragment of 1281 bp generated using these primers is also shown in red (ii) Gel picture of PCR amplification using the above mentioned primers gave a more intense band of 1281 bp in colonies 6 and 7. The plasmid pBCTaxK DNA was used as +ve control and water as –ve control (iii) Results of *Hind*III restriction digests of plasmids 6 and 7 matched the expected fragment sizes as in (i). The plasmid carrying the 35S-TXS without the terminator was used as control and each respective undigested plasmid labelled as Und was also run alongside the digested ones as control.

A BLAST search (http://www.ncbi.nlm.nih.gov/BLAST/) of the TXS sequence cloned into the pGWB8 vector indicated a 100% identity to *T. baccata* (AY424738.1) with no premature stop codon present within the TXS sequence. The sequence-verified plasmid 7 was transferred into *Agrobacterium* which was used to transform Wt *Arabidopsis* by the floral dip method as previously described (Section 2.5.1). Further analysis was not carried out on the seeds obtained following transformation due to time limitation.

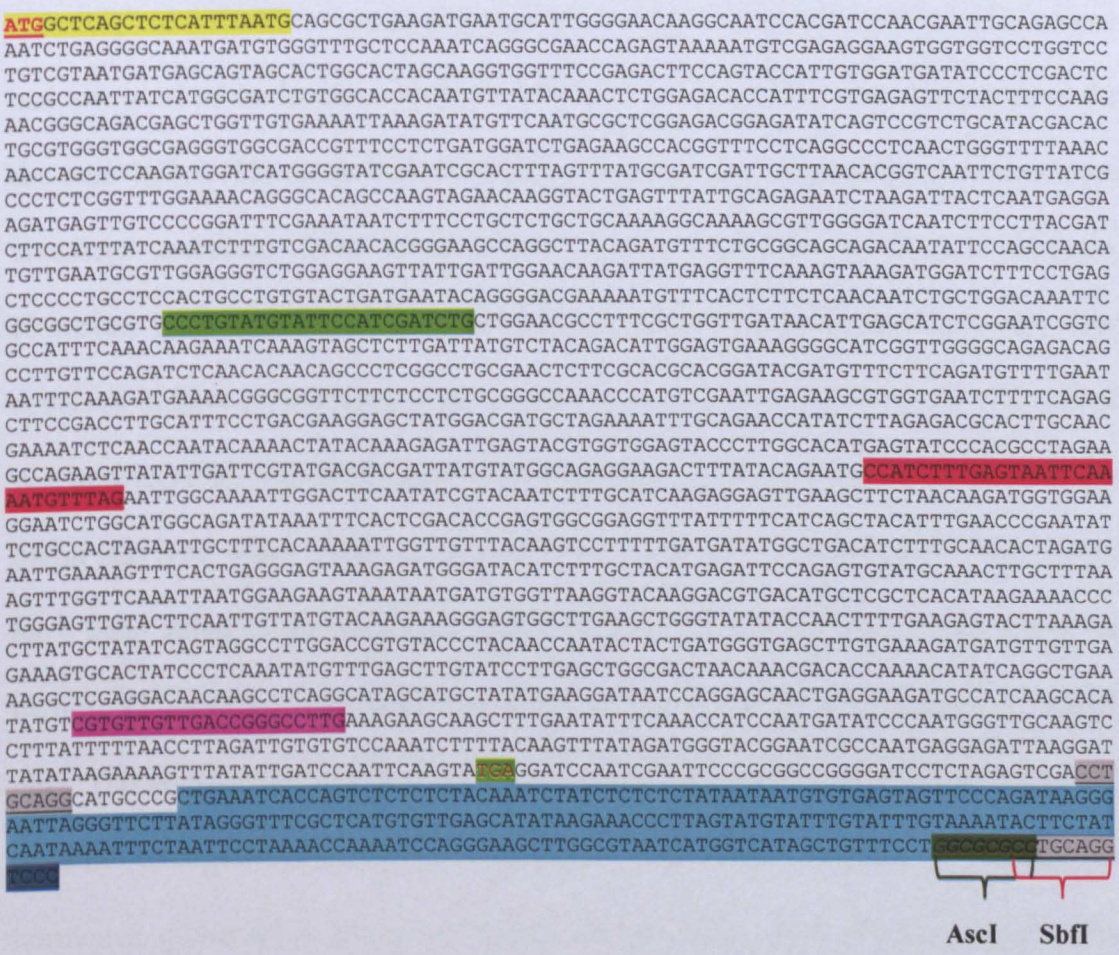


Figure 5.24 Nucleotide sequence of TXS+Terminator.

Nucleotide sequence obtained from sequencing of the TXS:T5OH:pGWB8 vector. The primers used for sequencing were: TaxSynF (X), TaxMid1 (X), TaxMid2 (X) and Tax-Terminator (X). The two PstI/SbfI sites (X) are shown to have been recreated following ligation of the 35S+TXS to the Terminator (X) and to the pGWB8 vector (X). X = *Asc*I site, X = Start codon, X = Stop codon.

5.3 Discussion

5.3.1 Novel taxanes were not detected in transgenic plants expressing the Taxol biosynthetic genes

Previous work on the constitutive expression of TXS in *yellow flesh* tomato mutants (Kovacs *et al.*, 2007) has resulted in a high level accumulation of taxadiene in both the fruit and leaves of the transgenic lines. Tobacco plants expressing the same pBCTaxK construct carrying the TXS gene initially provided at the beginning of the research, were found to accumulate a relatively high level of taxa-4(5),11(12)-diene and low level of the isomer taxa-4(20),11(12)-diene. The mass spectrum obtained for the taxa-4(5),11(12)-diene from transgenic tobacco was very similar to that of pure taxadiene extracted from tomato fruits, when subjected to GC-MS analysis. However, taxadiene accumulation level in the progeny from this tobacco parent line appeared to have considerably reduced, and in most of the plants, the taxadiene peak was undetectable. It could be possible that progeny producing the lowest level of this olefin survived while a high accumulation of taxadiene in homozygous lines might have been lethal to the plants, as appeared to be the case in line 4 expressing the highest TXS mRNA level, which went on to die at a very young stage (~ 1 month old). The constitutive production of the active TXS enzyme might have interfered with the endogenous production of plastidial isoprenoids, such as the diterpenoid hormones gibberellins and carotenoids and the side chain of chlorophylls, thus resulting in hormonal imbalance affecting the plant development. The large array of isoprenoids/terpenes produced in tobacco plants could also suggest a competition for the GGPP precursor, hence the recombinant TXS might not have been channeling significant amounts of GGPP away from the biosynthetic pathway of these plastidial

isoprenoids. This might account for the relatively smaller peak characteristic of taxadiene in the tobacco leaves when compared to that produced in tomato fruits. Taxadiene has been recently reported to be volatile (Ajukumar *et al.*, 2010) which might also explain its low accumulation and hence detection level in the crude hexane extracts from the transgenic tobacco lines.

With the aim of constructing the next step *en route* to Taxol in tobacco, transgenic plants co-expressing the TXS and T5OH transgenes were successfully generated. The GC-MS conditions under which these samples were analysed were slightly modified from the published GC-MS data under which an oxygenated taxane (OCT) was reported (Rontein *et al.*, 2008). Under these conditions, the authors reported the presence of the taxadiene peak at Rt 36.86 min, in transgenic *N. sylvestris* expressing the TXS gene. A peak was observed at Rt 36.58 in the hexane extracts from transgenic *N. tabacum* line expressing the TXS transgene, having the similar mass spectrum to that of authentic taxadiene, which suggested that the peak observed in this parent line used for crossing, was indeed that of taxadiene. Consistent with the previous findings (Rontein *et al.*, 2008; Ajikumar *et al.*, 2010) the taxa-4(5),11(12)-diene peak was absent from the GC-MS profile in transgenic *N. tabacum* lines when TXS and T5OH were co-expressed. If this is due to the conversion of this olefin in these lines, this would indicate that the cytochrome P450 reductase from tobacco interacts well with the yew CYP enzymes and that the NADPH is an available and non-limiting co-substrate in these cells.

The first oxygenated step *en route* to Taxol has been proposed to be the slow step of the pathway due to the low detection of the T5OH mRNA in methyl jasmonate elicited *Taxus* cell cultures and to the rarely detectable level of taxadiene-5α-ol in these cultures (Croteau *et al.*, 2006). In agreement with these observations, Ajikumar *et al* (2010) reported the production of only 60 mg of taxadien-5α-ol generated from 1g of taxadiene by the *E. coli* strain co-expressing TXS and T5OH transgenes. This ~20 fold less in the amount of taxadiene-5α-ol produced might be partially due to the volatile nature of this compound. The retention time for taxadiene-5α-ol under the GC-MS conditions (Rontein *et al.*, 2008) used was expected around 40.72 min, however, the GC chromatogram of the hexane extracts from tobacco lines co-expressing TXS and T5OH failed to show the presence of this oxygenated compound. The GC chromatogram indicated major peaks around the same retention time as taxadiene-5α-ol which could have overlaid that of the taxadiene-5α-ol if present in very low level, hence making its detection impossible. The failure of detecting taxadiene-5α-ol in the trichomes of transgenic *N. sylvestris* co-expressing TXS and T5OH by Rontein *et al.* (2008) could also be due to the large surface area of the trichomes, thereby allowing the escape of the volatile taxadiene-5α-ol, as could have been the case in the *N. tabacum* expressing these two transgenes. The presence of other CYPs in tobacco could also suggest that the taxadiene-5α-ol was used as substrate by these mono-oxygenases and converted to other compounds in the plant. The stability of taxa-4(20),11(12)-diene-5α-ol is unknown, hence it could be postulated that the product might have been synthesised in the transgenic plants but degraded prior to GC-MS analysis of the transgenic *N. tabaccum.*

The presence of the C5-hydroxyl group of taxadiene-5α-ol could also suggest that this compound is unstable and more susceptible to hydroxylation or acylation reactions, thus was rapidly modified in the transgenic *N. tabacum*. The complex structure of the taxadiene-5α-ol by-product OCT (Rontein *et al.*, 2008) displays a globular shape in which the oxygen atom of the ether bond formed between C5-C12 of the taxane skeleton is partially hidden inside the structure. This has led to the suggestion that the structure of this compound makes it non-volatile in standard conditions and insensitive to air oxidation, hence making it stable enough to be detected by GC-MS analysis. Recently, OCT was also reported at similar accumulation levels to that of taxadiene-5α-ol produced in transgenic *E. coli* co-expressing the TXS and T5OH (Ajukumar *et al.*, 2010). Williams *et al.* (2000b) previously documented the presence of the isomer taxa-3(4),11(12)-diene along with the other taxadiene isomers in *E. coli* expressing a "pseudomature" form of TXS gene. This taxa-3(4),11(12)-diene isomer appears to carry a C3=C4 double bond on the taxane skeleton and could be the potential substrate for the formation of OCT recently observed in the transgenic *N. sylvestris* trichomes and *E. coli*.

The hexane extracts from *N. tabacum* co-expressing the T5OH and TXS were also scanned for the ions characteristic of OCT but no such compound was identified in any of the samples. It could be postulated that the taxa-3(4),11(12)-diene isomer was modified to compounds which were used up as substrates in the leaves, hence would have not been available for the synthesis of OCT. It can also be hypothesised that OCT was produced in the transgenic plants but rapidly degraded by other cytochrome P450 enzymes present in the tobacco leaves but not in the trichomes of

the tobacco plants, explaining why the OCT was previously reported in the transgenic trichomes.

Taking into account the possible unstable nature of taxadiene-5α-ol which suggests that it might have been rapidly modified to novel compounds, plants expressing the three downstream genes TXS, T5OH and T5AT of the Taxol biosynthetic pathway were generated. The presence of an acetyl group at the C5-position of taxadiene-5α-ol was reasoned to probably stabilise this compound, and also prevent further modifications of the hydroxyl group at the C5 position of the taxadiene-5α-ol, thus allowing the accumulation of the acetylated compound in the transgenic tobacco. However, no peak characteristic of taxadiene-5α-yl acetate was observed, suggesting that the taxadiene-5α-ol might have been degraded by other CYPs associated to the endoplasmic reticulum, prior to being acetylated by the T5AT enzyme. Addition of an acetyl group to a compound might have also increased its volatility which might explain why the GC-MS analysis failed to detect this acetylated taxane.

5.3.2 Heterologous expression of Taxol biosynthetic genes in plants causes growth defects

GGPP is the immediate precursor of all isoprenoids and plays primary roles in the formation of the phytol moiety of chlorophylls, synthesis of carotenoids and phytohormones such as gibberellins and abscisic acid as well as for the synthesis of other isoprenoids.

The constitutive expression of the full length TXS in *A. thaliana* plants led to concomitant growth retardation, severe dwarfism and a lack of viable seed production, in agreement with the phenotypes previously reported in transgenic *Arabidopsis* producing taxadiene (Besumbes *et al.*, 2006). These observations suggest a possible toxic effect of taxadiene and/or a hormonal imbalance due to redirection of GGPP by TXS for the production of taxadiene. As a result, the lower accumulation of endogenous plastid isoprenoid products such as GA which is responsible for cell elongation and seed production could account for this dwarf phenotype.

Growth rate was also affected in transgenic tobacco lines expressing TXS, T5OH and T5AT. Line B30 expressing only the T5OH mRNA, and lines B15 and B31 co-expressing TXS and T5AT were found to have reduced internode lengths when compared to the wild type. Brassinosteroids are involved in growth-promotion and previous reports have demonstrated that disruptions of the BR biosynthetic pathway gives rise to dwarf plants (Reviewed in Clouse and Sasse, 1998). It can therefore be hypothesised that the dwarf phenotypes observed are due to the altered flux through the isoprenoid pathway, with the synthesis of high levels of taxadiene taking place in lines expressing the TXS transgene. In Line B30, it was postulated that an active enzyme form of T5OH might have acted on substrates involved in the GA/BRs biosynthetic pathways. As a result, this would have compromised the synthesis of these phytohormones which are crucial for plant development and stem elongation. A similar effect was previously seen in transgenic tomato plants engineered to constitutively over-express the phytoene synthase (*psy1*) gene (Fray *et al.*, 1995) where the authors suggested that the reduced stature and pale nature of the *psy1*

expressing plants were attributed to reduced availability of GGPP for the GA and phytol biosynthetic pathways.

Besides the dwarf phenotypes and slow growth, other phenotypes were observed as a result of the constitutive expression of TXS in the transgenic *r* tomato mutant. Most of the fruits producing high level of taxadiene were parthenocarpic or with less than ten viable seeds per fruit. A viability and germination test was performed on these transgenic lines and only around 50% of viable pollen grains were observed which all failed to germinate in the time period where ~50% of the wild type tomato pollen grains germinated. The high level expression of the TXS in these plants appeared to have affected the fertility of the plants, possibly due to the reduced accumulation of GA and/or BR. These hormones are involved in pollen development and pollen tube growth respectively, and the GGPP normally used for their biosynthesis could have been redirected by TXS for the production of taxadiene, at the expense of these phytohormones.

A male sterility phenotype was also observed in tobacco lines constitutively expressing the T5OH transgene, which was associated with abnormal seed formation and pollen infertility. The failure of germination of the viable pollen from lines expressing T5OH could be the result of a reduced level of production of brassinosteroid involved in pollen tube germination. Both gibberellins and brassinosteroids have been previously reported to be essential for the development of stamens and are involved in pollen fertility. This phenotype observed in these lines might suggest T5OH to be competing for a substrate involved in the GA or BRs biosynthesis pathway, leading to depleted levels of the respective bioactive hormone,

thereby affecting normal pollen development and germination. It could also be postulated that an active enzyme form of T5OH could have acted upon substrates in the transgenic lines for the synthesis of compound(s) which might have been toxic to the pollen and hence affected male sterility. When TXS expressing tobacco lines were crossed to the transgenic line expressing T5OH, reduced number of seeds per pod was obtained. This could suggest that the production of the T5OH enzyme might have affected the female reproductive organ as well as the production of pollen. Similar observation was made when wild type tomato lines were crossed to transgenic lines producing taxadiene, which could have been due to the toxic effect of the high accumulation level of this taxane, thereby affecting female sterility.

In conclusion, taking into account the low accumulation level of taxadiene compared to the other compounds synthesised and the many phenotypes associated with the expression of the Taxol biosynthetic genes in transgenic tobacco, it was reasoned that this particular plant might not be suitable for the heterologous expression of the Taxol biosynthetic genes. The *r* tomato mutant were considered to be a better alternative system in which to express TXS, where it has previously been shown to be possible to redirect the GGPP normally used for carotenoid biosynthesis, for the production of significant level of taxadiene. Another advantage of using the *r* mutant lines is that the low carotenoid level in these tomatoes would allow the facile extraction of any taxane produced from these fruits due to less carotenoid dissolving in the extraction buffer. The new TXS+T5OH::pGWB8 vector could be transformed into *r* mutant tomatoes, where a high level of taxadiene accumulation following the TXS expression should provide more substrate for the T5OH to catalyse into the taxadiene-5α-ol.

CHAPTER 6 : CONCLUSIONS AND FUTURE WORK

6.1 Discussion and Conclusion

The low accumulation level of Taxol in the yew trees and the intense activity in developing new uses of the drug in cancer chemotherapy and in the treatment of other human maladies has driven an extensive search for new sources to replace the yew tree for the continued interest in the supply and cost of this valuable natural product. A range of studies have revealed several of the early and late steps of the Taxol biosynthetic pathway, however, the sequence of reactions and the enzymes involved in the mid-section of the pathway which includes the formation of the oxetane ring are still uncertain. To date, 14 of the 20 genes required for the biosynthesis of Taxol have been cloned and their sequences are available. This has opened up the possibility of metabolic engineering for the improvement of taxoid synthesis by transferring the entire pathway to a heterologous surrogate.

The genes encoding the enzymes involved in the early steps of the pathway *en route* to Taxol were successfully cloned from *T. baccata* mRNA during this study. The sequence alignment of the hydroxylase genes T5OH, T13OH, T10BOH and T7OH with their respective *Taxus* homologues revealed a high level of conservation of the homologous taxoid hydroxylases among the different *Taxus* species. Minor genetic variation in the DNA sequences of the hydroxylases and acetyl transferase cloned

was observed when compared to their respective *Taxus* homologues, which could have been the result of single nucleotide polymorphism. BLAST database search indicated that these *Taxus* CYP oxygenases resembled the CYP families involved in gibberellic acid, abscisic acid and brassinosteroid biosynthetic pathways. Sequence analysis indicated that the *Taxus* hydroxylases share high similarity to the GA biosynthetic genes, indicating that these two pathways might share a common evolutionary ancestry. In both Taxol and GA biosynthesis, the early steps involve the cyclisation of GGPP in plastids followed by CYP-catalysed hydroxylation of the skeleton. It was reasoned that due to their sequence relatedness and possibly common ancestry, the enzymes of the early steps of GA and Taxol biosynthesis might be targeted to similar sub-cellular compartments. Localisation data obtained by translational fusion of the TXS, T5OH and T5AT to fluorescent proteins supported this suggestion, with similar localisation to the GA biosynthetic enzymes AtKS1, AtKO1 and AtKAO1 respectively (Helliwell *et al.*, 2001b).

TXS was found to reside in the plastids where it catalyses the cyclisation of GGPP to yield taxadiene. Due to its non-polar nature, taxadiene is likely to partition into plastidial membranes. From here, it is hydroxylated and translocated to the ER. T5OH carries out the first hydroxylation of the taxane core and was shown to be associated with both the plastidial envelope and the ER, as previously reported for the case for *A. thaliana ent*-kaurene oxidase (AtKO1) which catalyses the first oxygenation step of GA biosynthesis (Helliwell *et al.*, 1999). T5OH may therefore not only catalyse the formation of taxadiene-5α-ol, but also link the plastid and the subsequent ER-located steps. The T5AT responsible for the first acetylation step of Taxol biosynthesis, has been previously reported in *Taxus* microsomes (Wheeler *et*

al., 2001) which led the authors to suggest the enzyme to be targeted to the ER, however there has been no report confirming its sub-cellular localisation. The translational fusion of T5AT with cyan fluorescent protein has confirmed the acetyl transferase to be associated to the ER (and possibly to the cytoplasm of transgenic tobacco trichomes). These data indicated that the taxadiene synthesised in the plastids possibly requires the action of T5OH for its translocation to the ER, where the first hydroxylation step and acetylation take place in the respective presence of cytochrome P450 reductases and acetyl-CoA, both associated to the ER membranes.

The localisation of the early enzymes of Taxol biosynthesis may provide some insight into the pathway restriction previously encountered during the reconstitution of the Taxol biosynthetic pathway in yeast (Dejong *et al.*, 2005). The authors reported a pathway constraint at the T5OH step which was suggested to be related to the inadequate level of endogenous CYP-reductases for redox coupling to the cytochrome for C5 oxygenation of taxadiene. Taking into account the plastidial localisation of TXS and the absence of plastids in yeast, it could be hypothesised that there cannot be the same spatial separation of steps in this microbe as would have been the case in plant system where both plastids and ER are present in significant amount. Therefore, it can be suggested that taxadiene export from the plastid by T5OH might be an essential step, whereby taxadiene is oxidised to form more soluble intermediates by the plastid-associated T5OH. These are possibly further metabolised to the final product, taxadiene-5α-ol, once coupled to the ER-associated T5OH. The possible requirement for compartmentalisation of the early enzymes of Taxol biosynthesis might explain why individual expression of the biosynthetic genes including T5OH in yeast or bacteria resulted in active enzymes, yet

simultaneous expression of these genes in yeast did not result in the production of the expected downstream products. Furthermore, yeasts and bacteria do not have the same post-translational modification mechanisms as plants, which may be required by the yew enzymes for optimal activity. Therefore, it was reasoned that the large number of plastids and ER present in plants might make them a better platform for coupling the consecutive CYP-catalysed steps of the Taxol pathway, which may not form properly in bacterial and yeast cells.

The constitutive expression of TXS transgene in transgenic *yellow flesh* tomato mutants has been reported to produce taxadiene at levels 339 μg/g DW in ripe fruits and 160 μg/g DW in leaves of the transgenic plants (Kovacs *et al.*, 2007). However, the plants showed a slower growth rate with production of parthenocarpic fruits or fruits with less than 10 viable seeds. The pollen viability and germination test carried out on the pollen grains from these taxadiene producing tomato plants showed that 50% of the pollen were potentially viable but had a germination rate of almost nil. This observation could be explained by a possible reduction in the level of the plant hormones gibberellic acid and brassinosteroids available for pollen development, as a result of the redirection of the available GGPP by TXS, for the production of taxadiene. This sterility issue also meant that cross-pollinating these taxadiene producing *r* tomato lines to transgenic lines expressing the downstream enzymes of the Taxol biosynthetic pathway might not be successful.

As an alternative to the *r* tomato mutants, tobacco plants were used, due to their high biomass yields, robust transformation technology and strong biosafety profile. Independent transgenic tobacco lines expressing each of the enzymes involved in the

early steps of the Taxol pathway were generated during this study. These lines were alternatively crossed with the aim of assembling the required enzymes for the production of taxanes, which on purification and further modifications by chemists would give taxane-related drugs with improved solubility and activity or novel functions.

TXS expressing tobacco plants were shown to generate a minor product which was confirmed as taxa-4(5),11(12)-diene, by comparison of mass spectrum to those of the authentic standard prepared by total synthesis (Rubenstein and Williams., 1995). These taxadiene producing lines were cross-pollinated to those expressing the highest level of T5OH mRNA and the progenies expressing high levels of both transgenes were selected for GC-MS analysis. The GC-chromatogram obtained from these lines did not show the presence of the taxadiene-5α-ol which has been demonstrated to be the product of the first hydroxylation step of Taxol biosynthesis (Hefner *et al.*, 1996). However, the taxadiene drops to effectively zero in lines expressing both these transgenes but was still detected in line 21 carrying both genes but expressing only the TXS transgene. This suggested the possible near complete conversion of taxadiene in the presence of T5OH, to taxadiene-5α-ol or to other intermediates, which could have been rapidly turned over by the plants. The absence of the taxadiene peak also indicated that sufficient CYP-reductase was available for its complete hydroxylation to novel compounds in the plants.

The production of taxadiene-5α-ol has been previously reported in insignificant level in *Taxus* cell cultures (Croteau *et al.*, 2006), suggesting this step to be catalysed by the rate-limiting enzyme T5OH and also that the compound might be rapidly

degraded or of unstable nature. Furthermore, previous report of the expression of T5OH in *E.coli* strain producing ~1 g/L taxadiene have documented a level 20-fold less in the amount of taxadiene-5α-ol accumulated with the reported absence of taxadiene in the bacterial cell culture (Ajukumar *et al.*, 2010). These data might support the suggestion of a negative feedback on taxadiene production in the presence of T5OH. It can also be hypothesised that the taxadiene-5α-ol could have been rapidly turned-over in the transgenic tobacco lines or used as substrate for the synthesis of other compounds. The presence of the C5-hydroxyl group of taxadiene-5α-ol further accounts for the possible unstable nature of this compound implying that it is more susceptible to hydroxylation or acylation reactions. It can be postulated that taxadiene produced is protected whilst still in the plastids, but once moved out of these organelles to the ER, this olefin is degraded by endogenous plant enzymes, which might explain its absence in lines co-expressing the TXS and T5OH transgenes.

Similar work on tobacco was reported by Rontein *et al.*, (2007) who documented the presence of a taxane ether 5(12)-Oxa-3(11)-cyclotaxane (OCT) instead of the expected taxadiene-5α-ol, in the trichomes of transgenic *N. sylvestris* co-expressing TXS and T5OH transgenes. The three-dimensional model of OCT was reported to display a globular structure in which the oxygen atom of a C5-C12 ether linkage is partially hidden inside the structure, making this compound very stable, insensitive to air oxidation, and non-volatile in standard conditions. The GC-chromatograms obtained from the *N. tabacum* expressing TXS and T5OH were scanned for the major ions of OCT, however no peak having the mass spectrum fragmentation pattern as that of OCT was present in the transgenic tobacco crude extracts. It could

be speculated that if OCT was produced, then this compound was also degraded by plant catabolic enzymes. The C3-C11 linkage of OCT might also indicate that this compound is formed as a result of oxidation of the taxa-3(4),11(12)-diene isomer initially reported in very low level in transgenic *E. coli* (Williams *et al.*, 2000 b). Supposing that OCT is the actual oxygenated compound of taxa-3(4),11(12)-diene, it could be hypothesised that this taxadiene isomer might have been produced in the transgenic tobacco cells but was rapidly broken down by enzymes absent in the trichome cells.

Taking into account the unstable nature of taxadiene-5α-ol due to the presence of its C5-hydroxyl group, it was reasoned that the addition of an acetyl group at this position of the taxane core might make the compound more stable. A transgenic tobacco expressing the T5AT transgene was generated, which was crossed to tobacco lines co-expressing the TXS and T5OH transgenes. Progenies expressing all three transgenes were confirmed by northern analysis and their crude hexane extracts were analysed by GC-MS using the conditions under which authentic taxadiene-5α-yl acetate was previously reported (Walker *et al.*, 1999). No peak having the mass spectrum characteristic of this acetylated compound was observed in these lines. However, by-standing minor peaks were observed in the total ion current from these transgenic lines, which were absent in the Wt extract. It can be speculated that if taxadiene-5α-yl acetate was produced it would have been catabolised to other compounds which were indicated by these by-standing peaks. T5AT has been demonstrated to be localised to the ER, thus it could also be suggested that the taxadiene-5α-ol might have been degraded by other catabolic enzymes also associated to the ER, prior to being acetylated by the T5AT which shares the same

location. Addition of an acetyl group to the taxadiene-5α-ol might have also increased the volatility of this compound, hence if present, could have been lost during hexane extraction.

The expression of the early enzymes of the Taxol biosynthetic pathway was associated with slow growth of the plants, male sterility and severe dwarfism. As mentioned above, taxadiene production in *Yellow flesh* tomato plants affected the pollen growth and development, producing mostly parthenocarpic fruit in part due to incomplete pollination. The failure of the viable pollen grains to germinate might indicate a depleted level of brassinosteroids produced responsible for pollen tube growth, as a result of the redirection of GGPP by TXS for the production of taxadiene. Similarly, seed production was severely affected in transgenic *Arabidopsis* lines constitutively expressing the TXS transgene. The male sterility issues observed can therefore be associated with failure in pollen tube germination. Taxadiene itself has not been reported to be pharmacologically active, and its biological activity has not been extensively tested. The sterility problems in these transgenic lines could have also been the result of a possible cytotoxic effect of taxadiene on the pollen cells or the ovules. Previous crossing of Wt tomato lines to taxadiene producing ones gave parthenocarpic fruits with no viable seeds. This led to the speculation that taxadiene might have been toxic to the plant and affected the female reproductive system as well as pollen formation. The cytotoxic effect of taxadiene could therefore be further investigated by its exogenous application to the various stages of floral development as well as to pollen, stigma and ovaries of the Wt tobacco, tomato and *Arabidopsis* flowers. This would provide an insight to its effect on the mitotic

spindle formation, chromosomal segregation of the transgenic plant cells and to investigate the cell division of the microsporocytes and megasporocytes.

Tobacco plants expressing T5OH also demonstrated a reduced formation of viable pollen, having a very low level of pollen tube growth. This indicates that this low germination rate was the cause of male sterility in these transgenic lines. The presence of the active enzymatic form of T5OH may be competing for substrates involved in the biosynthetic pathways of the phytohormones GA or BRs, leading to depleted levels of bioactive hormones, thereby affecting pollen development and pollen tube growth. Alternatively, T5OH may have promiscuous activity and act on endogenous substrate(s) leading to the synthesis of compounds which might have had an adverse effect on the floral reproductive organs.

Transgenic *Arabidopsis* plants constitutively expressing TXS showed severe dwarfism, in agreement with previous report (Besumbes *et al.*, 2004) suggesting a possible competitive and toxic effect of TXS or taxadiene in these plants. The same phenotype was observed in transgenic tobacco lines co-expressing TXS and T5AT, suggesting that the redirection of GGPP by TXS for the production of taxadiene compromised the synthesis of GA and BRs involved in stem elongation and plant development. Similarly, the growth of tobacco line 30 transgenic for TXS, T5OH and T5AT but expressing only the T5OH transgene was affected resulting in shorter internodes when compared to the Wt control. However, unlike tobacco, *Arabidopsis* expressing the T5OH showed no concomitant dwarf phenotypes or reduction in seed production. This observation might suggest that the active enzymatic form of T5OH

may have substrates, upon which it can act in tobacco but which are absent in *Arabidopsis.*

Individual transformation followed by crossing to "stack" transgenes can be very difficult for large numbers of transgenes because of time taken to stack all transgenes in one line and the risk of segregation of unlinked genes in later generation. Taking these hurdles into account, it was reasoned that stacking a combination of the downstream genes of the Taxol biosynthetic pathway in one plant binary vector and engineering it in plants to create diverse transgenic plants could provide a simpler way of reconstituting the Taxol pathway, preferentially in *r* mutant tomatoes. The GC-chromatogram of transgenic *r* mutant fruits expressing TXS demonstrated a major peak of taxadiene and minor by-standing peaks, which are possibly representative of other isoprenoids produced. Tobacco on the other hand has a large array of terpenes and other compounds in the leaves which have made the detection of novel taxanes by GC-MS analysis very difficult. Furthermore, the *r* mutant has a very low accumulation level of carotenoid and thus it was reasoned that using this genetic background should allow the facile extraction and down-stream processing of novel taxanes from the fruits.

To address the sterility concerns observed in lines with the individual expression of the Taxol biosynthetic genes as well as to devising a simpler process of stacking the genes of the Taxol pathway in *r* tomato, a construct was prepared carrying the genes TXS and T5OH. A PCR-directed integration method was used to introduce the TXS cassette including the CaMV35S promoter and terminator, upstream of the T5OH

cassette in the plant binary vector pGWB8. Sequencing of the amplified TXS clone demonstrated a 100% identity to the published TXS *T. baccata* homologue with the reconstruction of two *Sbf*I restriction sites upstream and towards the end of the CaMV35S terminator respectively. This CaMV35S terminator was also amplified to include an *Asc*I restriction site towards the end of its sequence where other gene cassettes can be cloned into and the final vector transferred into *yellow flesh* tomato mutant.

Although, heterologous plant expression might provide a simpler method of production when compared to semi-synthesis of yew-tree derived precursors, issues associated with deleterious phenotypes in the host plants will have to be addressed.

Future Work

As an alternative way of extracting the immediate Taxol precursors baccatin III and 10-deacetylbaccatin III from the needles of the yew trees for its conversion to Taxol, the biosynthetic pathway *en route* to these precursors could be transferred to *r* mutant tomato. The other downstream genes of the Taxol pathway leading to the production of these late intermediates will be engineered in the vector TXS+T5OH::pGWB8 which will be transformed into *yellow flesh* tomato plants. This method will combine genetic engineering and semi-synthesis, whereby the late intermediates *en route* to Taxol will be extracted from the transgenic *yellow flesh* tomato plants and converted to paclitaxel by synthetic coupling of the C13 side chain or other advanced taxanes by chemists. Alternatively, this construct carrying only TXS and T5OH can be

transformed into *r* tomato and the transgenic lines expressing both transgenes if not sterile, will be crossed with transgenic tomato lines expressing the downstream Taxol biosynthetic enzymes. Using this method, different genes of the Taxol biosynthetic pathway will be assembled in various orders in individual lines to produce the desired molecules which are to be extracted and further modified by chemists to produce novel taxanes for example with better solubility than Taxol or those which have the ability of modulating MDR in cells resistant to Taxol.

Engineering *r* tomato plants with these heterologous genes could be accompanied with deleterious phenotypic issues. To circumvent the issue of male sterility in tomato producing taxadiene, an inducible promoter can be placed upstream of the TXS cDNA in the TXS+T5OH::pGWB8 vector. For example, the tetracycline inducible promoter would give finer control over the transgene expression which could be induced at the early stages of fruit development by the exogenous application of the antibiotic tetracycline. The TXS has been previously expressed under the control of the fruit-specific PG promoter, however fruit set was greatly affected in these transgenic lines. Other fruit-specific promoters could be used such as E8, 2A11 and P119, which have been previously reported to drive the expression of exogenous genes in tomato fruits following ethylene production during fruit ripening. E8 promoter demonstrates low expression level in the absence of ethylene synthesis in unripe fruit and high expression level during ripening, 2A11 is expressed at high levels in ripening fruits but also transiently in small green fruits, while P119 promoter is expressed from the earliest stages in green fruit though to red ripe stages, with increased level during ripening. It should also be considered that ethylene is

produced during floral development, hence the expression of the transgenes could affect fruit development.

As an alternative to using a plant heterologous system, the Taxol biosynthetic pathway could be engineered in the rice pathogen *Gibberella fujikuroi,* which has been reported to produce high levels of gibberellic acid. The large GGPP pool available for GA biosynthesis could be redirected by TXS for the production of taxadiene. Furthermore the short generation time and high growth rate of the fungi might make it a more reliable, cheap and renewable way of optimising Taxol and related taxanes.

BIBLIOGRAPHY

Ajikumar, P.K. Xiao, W-H., Tyo, K.E.J., Wang, Y., Simeon, F., Leonard, E., Mucha, O., Phon, T.H., Pfeifer, B. and Stephanopoulos, G. (2010) Isoprenoid Pathway Optimisation for Taxol Precursor Overproduction in *Escherichia coli*. *Science* **330** (70) 70-74

Altstadt, T.J., Fairchild, C.R., Golik, J., Johnson, K.A., Kadow, J.F., Lee, F.Y., Long, B.H., Rose, W.C., Vyas, D.M, Wong, H., Wu, M-J and Wittman, M.D. (2001) Synthesis and Antitumour Activity of Novel C-7 Paclitaxel Ethers: Discovery of BMS-184476. *Journal of Medical Chemistry* **44**: 4577-4583

Anterola, A., Shanle, E., Perroud, P-F. and Quatrano, R. (2009) Production of taxa-4(5),11(12)-diene by transgenic *Physcomitrella patens*. *Transgenic Research* **18**: 655-660

Bai, J., Ito, N., Sakai, J., Kitabatake, M., Fujisawa, H., Bai, L., Dai, J., Zhang, S., Hirose, K., Tomida, A., Tsuruo, T. and Ando M. (2005) Taxoids and abietanes from callus cultures of *Taxus cuspidata*. *Journal of Natural Products* **68**: 497-501.

Baloglu, E. and Kingston, D.G.I. (1999) The Taxane Diterpenoids. *Journal of Natural Products*. **62**: 1448-1472

Benveniste, I., Salaun, J-P. and Durst, F. (1977) Wounding-Induced Cinnamic Acid Hydroxylase in Jerusalem Artichoke Tuber. *Phytochemistry* **16**: 69-73

Besumbes, O., Sauret-Gueto, S., Phillips, M.A., Imperial, S., Rodriguez-Concepcion, M., and Boronat, A.(2004) Metabolic Engineering of Isoprenoid Biosynthesis in Arabidopsis for the Production of Taxadiene, the First Committed Precursor of Taxol. *Biotechnology and Bioengineering* **88** (2): 168-175

Bissery, M-C., Guénard, D., Guéritte-Voegelein, F. and Lavelle, F. (1991) Experimental Antitumour Activity of Taxotere (RP 56976, NSC 628503), a Taxol Analogue. *Cancer Research* **51**: 4845-4852

Bringi, V., Kadkade, P.G., Prince, C.L. and Roach, B.L. (2011) Enhanced production of taxol and taxanes by cell cultures of *Taxus* species. Pub No. US 2011/0086397 A1

Brooks, T., Minderman, H., O'Loughlin, K.L., Pera, P., Ojima, I., Baer, M.R. and Bernacki, R.J. (2003) Taxane-based reversal agents modulate drug resistance mediated by P-glycoprotein, multidrug resistance protein, and breast cancer resistance protein. *Molecular Cancer Therapeutics* **2**: 1195-1205

Chau, M. and Croteau, R. (2004) Molecular cloning and characterization of a cytochrome P450 taxoid 2α-hydroxylase involved in Taxol biosynthesis. *Archives of Biochemistry and Biophysics*. 427:48–57

Chau, M.D., Jennewein, S., Walker, K., Croteau, R. (2004) Taxol Biosynthesis: Molecular cloning and characterization of a cytochrome P450 Taxoid 7β-Hydroxylase *Chemistry and Biology* **11**: 663-672

Chauvière, G., Guénard, D., Picot, F., Senilh, V. and Potier, P. (1981). Structural analysis and biochemical study of isolated products of the yew: *Taxus baccata* L. (Taxaceae). *Comptes Rendus de l'Académie des Sciences* **293**: 501-503.

Christen, A.A., Gibson, D.M. and Bland, J. (1991) Production of Taxol or Taxol-like compounds in cell culture. US. Patent No. 5,019,504.

Clough, S.J and Bent, A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant Journal* **16** (6): 735-753

Clouse, S.D. and Sasse, J.M (1998) Brassinosteroids: Essential Regulators of Plant Growth and Development. *Annual Review of Plant Physiology and Plant Molecular Biology* **49**: 427-51

Croteau, R., Ketchum, R.E.B., Long, R.M., Kaspera, R., and Wildung, M. R (2006) Taxol biosynthesis and molecular genetics *Phytochemistry Reviews* **5**: 75-97

Cusidó, R.M., Palazón, J., Bonfill, M., Navia-Osorio, A., Morales, C. and Pinol. M.T. (2002) Improved Palitaxel and Baccatin III Production in Suspension Cultures of Taxus media. *Biotechnology Progress* **18**: 418-423

D'Ambrosio, C., Giorio, G., Marino, I., Merendino, A., Petrozzam A., Salfi, L., Stigliani, A.L. and Cellini, F. (2004) Virtually complete conversion of lycopene into β-carotene in fruits of tomato plants transformed with the tomato *lycopene β-cyclase (tlyc-b)* cDNA. *Plant Science* **166**: 207-214

Daniell, H., Kumar, S., Dufourmantel, N. (2005) Breakthrough in chloroplast genetic engineering of agronomically important crops. *TRENDS in Biotechnology* **23** (5): 238-245

Daniewski, W.M., Gumulka, M., Anczewski, W., Masnyk, M., Bloszyk, E. and Gupta, K.K. (1998) Why the yew tree (*Taxus* baccata) is not attacked by insects. *Phytochemistry* **38**. No. 4: 1279-1282

Davuluri, G.R., Tuinen, A.V., Fraser, P.D., Manfredonia, A., Newman, R., Burgess, D., Brummell, D.A., King, S.R., Palys, J., Uhlig, J., Bramley., P.M., Pennings, H.M.J. and Bowler, C. (2005) Fruit-specific RNAi-mediated suppression of *DET1* enhances carotenoid and flavonoid content in tomatoes. *Nature Biotechnology* **23** (7) 890-895

DeJong, J.H.M., Liu, Y., Bollon, A. P., Long, R. M., Jennewein, S., Williams, D., Croteau, R. B (2005) Genetic Engineering of Taxol Biosynthetic Genes in *Saccharomyces cerevisiae Biotechnology and Bioengineering* **93** (2) 212-224

Edwards, K., Johnstone, C. and Thompson, C. (1991) A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Research* **19** (16): 1349

Eisenreich, W., Menhard, B., Hylands, P.J., Zenk, M.H., Bacher, A. (1996) Studies on the biosynthesis of Taxol: the taxane carbon skeleton is not of mevalonoid origin. *Proceedings of the National Academy of Sciences USA* **93**: 6431-6436

Elmer, W.H,m Mattina, M.J.I. and MacEachern, G.J. (1994) Sensitivity of Plant Pathogenic Fungi to Taxane Extracts from Ornamental Yews. *Biochemistry and Cell Biology* **84:** 1179-1185

Engels, B, Dahm, P. and Jennewein, S. (2008) Metabolic engineering of taxadiene biosynthesis in yeast as a first step towards Taxol (Paclitaxel) production. *Metabolic Engineering* **10**: 201-206

Expósito, O., Bonfill, M., Moyano, E., Onrubia, M., Mirjalili, M.H., Cusido, R.M., and Palazón, J. (2009) Biotechnological Production of Taxol and Related Taxoids: Current State and Prospects. *Anti-Cancer Agents in Medicinal Chemistry* **9**: 109–21.

Farré, G., Sanahuja, G., Naqvi, S., Bai, C., Capell, T., Zhu, C. and Christou, P. (2011) Travel advice on the road to carotnenoids in plants. *Plant Science* **179**: 28-48

Fleming, P.E., Knaggs, A.R., He, X-G, Mocek, U. and Floss, H.G. (1994) Biosynthesis of Taxoids. Mode of Attachment of the taxol Side Chain. *Journal of American Chemical Society* **116**: 4137-4138

Floss, H.G. and Mocek, U.(1995) Biosynthesis of Taxol. In: Suffness M, *Taxol - Science and Applications.* CRC Press; Boca Raton, FL, USA: 191–208.

Fraser, P.D., Kiano. J.W., Truesdale, M.R., Schuch, W., and Bramley, P.M (1999) Phytoene synthase-2 enzyme activity in tomato does not contribute to carotenoid synthesis in ripening fruit. *Plant Molecular Biology* **40:** 687-698

Fraser, P.D., Romer, S., Shipton, C.A, Mills, P.B., Kiano, J.W., Misawa, N., Drake, R.G., Schuch, W. and Bramley, P.M. (2002) Evaluation of transgenic tomato plants expressing an additional phytoene synthase in a fruit-specific manner. *Proceedings of the National Academy of Sciences USA* **99**(2) 1092-1097

Fray, R.G., Wallace, A., Fraser, P.D., Valero, D., Hedden, P., Bramley, P.M. and Grierson, D. (1995) Constitutive expression of a fruit phytoene synthase gene in transgenic tomatoes causes dwarfism by redirecting metabolites from the gibberellin pathway. *The Plant Journal* **8**(5): 693-701

Fray, R.G. and Grierson, D. (1993) Identification and genetic analysis of normal and mutant phytoene synthase genes of tomato by sequencing, complementation and co-suppression. *Plant Molecular Biology* **22**: 589-602

Galletti, E., Magnani, M., Renzulli, M.L. and Botta, M. (2007) Paclitaxel And Docetaxel Resistance: Molecular Mechanisms and Development of New Generation Taxanes. *Chemmedchem* **2**: 920-942

Giner, J-L. and Faraldos, J. A. (2003) Facile Orthoester Formation in a Model Compound of the Taxol Oxetane: Are Biologically Active Epoxy Esters, Orthoesters, and Oxetanyl Esters Latent Electrophiles? *Helvetica Chinica Acta* **86**: 3613-3622

Goldsbrough, P.B and Cullis, C.A. (1981) Characterisation of the genes for ribosomal RNA in flax. *Nucleic Acids Research* **9** (6): 1301-1309

Gómez-Galera, S., Pelacho, A.M., Gené, A., Capell, T. and Christou, P. (2007) The genetic manipulation of medicinal and aromatic plants. *Plant Cell Reports* **26**: 1689-1715

Guénard, D., Guéritte-Voegelein, F. and Potier, P. (1993) Taxol and Taxotere: Discovery, Chemistry, and Structure-Activity Relationships. *Accounts of Chemical Research* **26**: 160-167

Guénard, D., Thoret, S., Dubois, J., Adeline, M-T., Wang, Q. and Guéritte, F. (2000) Effects of the Hydrophobicity of Taxoids on their Interaction with Tubulin. *Bioorganic & Medicinal Chemistry* **8**: 145-156

Guéritte-Voegelein, F., Guénard, D., Lavelle, F., Le Goff, M-T., Mangatal, L. and Potier, P. (1991) Relationships between the Structure of Taxol Analogues and Their Antimitotic Activity. *Journal of Medical Chemistry* **34**: 992-998

Guéritte-Voegelein, F., Guénard, D. and Potier, P. (1987) Taxol and derivatives: a biogenetic hypothesis. *Journal of Natural Products* **50** (1): 9-18

Guo, B.H., Kai, G.Y., Jin, H.B. and Tang, K.X. (2006) Taxol synthesis. *African Journal of Biotechnology* **5** (1): 15-20

Hamberger, B and Bohlmann, J (2006) Cytochrome P450 mono-oxygenases in conifer genomes: discovery of members of the terpenoid oxygenase superfamily in spruce and pine. *Biochemical Society Transactions* **34** (6): 1209-1214

Han, K. H., Fleming, P. F., Walker, K., Loper, M., Scott Chilton, W., Mocek, U., Gordon, M. P. and Floss, H. G (1994) Genetic transformation of mature *Taxus* : genetically control the in vitro production drug, Taxol an approach to of the anticancer. *Plant Science* **95**: 187-196

Haseloff, J., Siemering, K.R., Prasher, D.C. and Hodge S. (1997) Removal of a cryptic intron and subcellular localization of green fluorescent protein are required to

mark transgenic *Arabidopsis* plants brightly. *Proceedings of the National Academy of Sciences* USA. **94**: 2122-2127.

Hasemann, C.A., Kurumbail, R.G., Boddupalli, S.S., Peterson, J.A. and Deisenhofer, J. (1995) Structure and function of cytochrome P450: a comparative analysis of three crystal structures. *Structure* **3**(1): 41-62

Hefner, J., Ketchum, R.E.B and Croteau, R. (1998) Cloning and Functional Expression of a cDNA Encoding Geranylgeranyl Diphosphate Synthase from *Taxus canadensis* and Assessment of the Role of this Prenyltransferase in Cells Induced for Taxol Production. *Archives in Biochemistry and Biophysics* **360**: 62-74

Hefner, J., Rubenstein, S.M., Ketchum, R.E.B, Gibson, D.M., Williams, R.M. and Croteau, R. (1996) Cytochrome P450-catalyzed hydroxylation of taxa-4(5),11(12)-diene to taxa 4(20),11(12)-diene-5α-ol: the first oxygenation step in Taxol biosynthesis. *Chemical Biology* **3**: 479-488

Heim, R., Prasher, D.C. and Tsien, R.Y. (1994) Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proceedings of the National Academy of Sciences* USA. **91**:12501-12504.

Helliwell, C.A., Chandler, P.M., Poole, A., Dennis, E.S. and Peacock, W.J. (2001a) The CYP88A cytochrome P450, ent-kaurenoic acid oxidase, catalyzes three steps of the gibberellin biosynthesis pathway. *Proceedings of the National Academy of Sciences USA* **98** (4): 2065-2070

Helliwell, C.A., Poole, A., Peacock, W.J. and Dennis, E.S. (1999) *Arabidopsis ent*-kaurene oxidase catalyses three steps of gibberellin biosynthesis. *Plant Physiology* **119**: 507-510.

Helliwell, C.A., Sullivan, J.A., Mould, R.M., Gray, J.C., Peacock, W.J and Dennis, E.S. (2001b) A plastid envelope location of Arabidopsis ent-kaurene oxidase links the plastid and endoplasmic reticulum steps of the gibberellin biosynthesis pathway *The Plant Journal* **28**: 201-208

Hezari, M., Lewis, N.G. and Croteau, R. (1995) Purification and characterization of taxa-4(5),11(12)-diene synthase from Pacific yew (*Taxus brevifolia*) that catalyses the first committed step of Taxol biosynthesis. *Archives of Biochemistry and Biophysics* **322**: 437-444

Hezari, M., Ketchum, R.E.B, Gibson, D.M. and Croteau, R (1997) Taxol production and taxadiene synthase activity in *Taxus canadensis* cell suspension cultures. *Archives of Biochemistry and Biophysics* **337**:185-190

Ho, T-I., Lee, G-H. and Peng, S-M. (1987) Structure of Taxusin. *Acta Crystallographica* **43**: 1378-1380

Holmes, D.S and Quigley, M. (1981) A rapid boiling method for the preparation of bacterial plasmids. *Analytical Biochemistry* **114** (1) 193-197

Holton, R.A., Kim, H-B, Somoza, C., Liang, F., Biediger, R.J., Boatman, P.D., Shindo, M., Smith, C.C., Kim, S., Nadizadeh, H., Suzuki, Y., Tao, C., Vu, P., Tang, S., Zhang, P., Murthi, K.K., Gentile, L.N. and Liu, J.H. (1994a) First total synthesis of Taxol. 2. Completion of the C and D rings. *Journal of American Chemical Society* **116**:1599-1600

Holton, R.A., Somoza, C., Kim, H-B, Liang, F., Biediger, R.J., Boatman, P.D., Shindo, M., Smith, C.C., Kim, S., Nadizadeh, H., Suzuki, Y., Tao, C., Vu, P., Tang, S., Zhang, P., Murthi, K.K., Gentile, L.N. and Liu, J.H. (1994b) First total synthesis of Taxol. 1. Functionalization of the B ring. *Journal of American Chemical Society* **116**: 1597-1598

Hu, G., Mei, X., Xiong, Q., Zhou, A. and Gong, W. (2002) Differential Display Reveals a cDNA Coding for an Acyltransferase-like Protein Specific to Taxol Synthesis in *Taxus chinensis. Plant Molecular Biology Reporter* **20**: 131-140

Huang, K-X., Huang, Q-L., Wildung, M.R., Croteau, R. and Scott, A.I (1998) Overproduction in *Escherichia coli*, of soluble taxadiene synthase, a key enzyme in the Taxol biosynthetic pathway. *Protein Expression and Purification* **13**: 90–96.

Huang, Q., Roessner, C.A., Croteau, R and Scott, A.I. (2001) Engineering *Escherichia coli* for the Synthesis of Taxadiene, a Key Intermediate in the Biosynthesis of Taxol. *Bioorganic & Medicinal Chemistry* **9**: 2237-2242

Itokawa, H. (2003). Taxoids occurring in the genus *Taxus*. In: Itokawa, H., Lee, K.-H. (Eds.), *Taxus* – The Genus *Taxus*. Taylor & Francis, London, UK, pp. 35–78.

Jennewein, S. and Croteau, R. (2001) Taxol: biosynthesis, molecular genetics, and biotechnological applications. *Applied Microbiology and Biotechnology* **57**: 13-19

Jennewein, S., Long, R., Williams, R.M. and Croteau, R. (2004a) Cytochrome P450 taxadiene 5α-hydroxylase, a mechanistically unusual monooxygenase catalyzing the first oxygenation step of Taxol biosynthesis. *Chemistry and Biology* **11**: 379–387

Jennewein, S., Park, H., DeJong, J. H. M., Long, R. M., Bollon, A. P. and Croteau, R. B. (2005) Coexpression in Yeast of *Taxus* Cytochrome P450 Reductase With Cytochrome P450 Oxygenases Involved in Taxol Biosynthesis. *Biotechnology and Bioengineering* **89**(5): 588-598

Jennewein, S., Rithner, C.D., Williams, R.M. and Croteau, R. (2001) Taxol biosynthesis: taxane 13α-hydroxylase is a cytochrome P450-dependent monooxygenase. *Proceedings of the National Academy of Sciences USA.* **98**: 13595–13600.

Jennewein, S., Rithner, C. D., Williams, R. M, and Croteau, R. (2003) Taxoid metabolism: Taxoid 14β-hydroxylase is a cytochrome P450-dependent monooxygenase *Archives of Biochemistry and Biophysics* **413**(2): 262-270

Jennewein, S., Wildung, M.R., Chau, M., Walker, K., Croteau, R. (2004b) Random sequencing of an induced *Taxus* cell cDNA library for identification of clones involved in Taxol biosynthesis. *Proceedings of the National Academy of Sciences USA* **101**: 9149–9154

Kai, G., Zhao, L., Zhang, L., Li, Z., Guo, B. and Zhao, D. (2005) Characterisation and Expression Profile Analysis of a New cDNA Encoding Taxadiene Synthase from *Taxus media*. *Journal of Biochemistry and Molecular Biology* **38** (6): 668-675

Keegstra, K. and Cline, K. (1999) Protein import and routing systems of chloroplasts. *Plant cell.* **11**: 557-570

Ketchum, R.E.B., Gibson, D.M., Croteau, R.B. and Schuler, M.L. (1999a) The kinetics of taxoid accumulation in cell suspension cultures of *Taxus* following elicitation with methyl jasmonate. *Biotechnology and Bioengineering* **62**:97-105

Ketchum, R.E.B., Horiguchi, T., Qiu, D., Williams, R.M. and Croteau, R.B (2007a) Administering cultured *Taxus* cells with early precursors reveals bifurcations in the taxoid biosynthetic pathway. *Phytochemistry* **68** (3): 335-341

Ketchum, R.E.B., Rithner, C.D., Qiu, D., Kim, Y., Williams, R.M., Croteau, R.B. (2003) *Taxus* Metabolomics: Methyl Jamonate Preferentially Induces Production of Taxoids Oxygenated at C-13 in *Taxus* x *media* Cell Cultures. *Phytochemistry* **62**: 901-909.

Ketchum, R.E.B. Tandon, M., Gibson, D.M., Begley, T. and Shuler, M.L. (1999b) Isolation of Labelled 9-Dihydrobaccatin III and Related Taxoids from Cell Cultures of *Taxus canadensis* Elicited with Methyl Jasmonate. *Journal of Natural Products* **62**: 1395-1398

Ketchum, R.E.B., Wherland, L. and Croteau, R.B. (2007b) Stable transformation and long-term maintenance of transgenic *Taxus* cell suspension cultures. *Plant Cell Reports* **26**: 1025-1033

Kim, G-T., Fujioka, S., Kozuka, T., Tax, F.E., Takatsuto, S., Yoshida, S. and Tsukaya, H. (2005) CYP90C1 and CYP90D1 are involved in different steps in the brassinosteroid biosynthesis pathway in *Arabidopsis thaliana*. *The Plant Journal* **41**: 710-721

Kim, G-T. and Tsukaya, H. (2002) Regulation of the biosynthetis of plant hormones by cytochrome P450s. *Journal of Plant Research* **115**: 169-177

Kobayashi, J., Hosoyama, H., Wang, X-X., Shigemori, H., Sudo, Y. and Tsuruo, T. (1998) Modulation of multidrug resistance by *Taxus* pine C and other taxoids from japanese yew. *Bioorganic & Medicinal Chemistry Letters* **8**: 1555-1558

Kobayashi, J. and Shigemori, H. (2002) Bioactive Taxoids from the Japanese Yew *Taxus cuspidata*. *Medicinal Research Reviews* **22** (3): 305-328

Koepp, A.E., Hezari, M., Zajicek, J., Stofer-Vogel, B., LaFever, R.E., Lewis, N.G. and Croteau, R. (1995) Cyclization of geranylgeranyl diphosphate to taxa-4(5),11(12)-diene is the committed step of Taxol biosynthesis in Pacific yew. *Journal of Biology and Chemistry* **270**: 8686-8690

Kovacs, K., Zhang, L. Linforth, R. S. T., Whittaker, B., Hayes, C. J. and Fray, R. G (2007) Redirection of carotenoid metabolism for the efficient production of taxadiene [taxa-4(5),11(12)-diene] in transgenic tomato fruit. *Transgenic Research* **16**: 121-126

Kumaran, R.S., Jung, H. and Kim, H.J. (2011) *In vitro* screening of taxol, an anticancer drug produced by the fungus, *Colletotrichum capsici. Engineering for Life Sciences* **11**(1): 1-8

Larkin. P.J., Miller, J.A.C., Allen, R.S., Chitty, J.A., Gerlach, W.L., Frick, S., Kutchan, T.M. and Fist, A.J. (2007) Increasing morphinan alkaloid production by over-expressing codeinone reductase in transgenic *Papaver somniferum. Plant Biotechnologoy Journal* **5**: 26-37

Li, J-Y, Strobel, G., Sidhu, R., Hess, W.M. and Ford, E.J. (1996) Endophytic Taxol-producing fungi from bald cypress, *Taxodium distichum. Microbiology* **142**:2223-2226

Lin, X., Hezari, M., Koepp, A. E., Floss, H. G. and Croteau, R. (1996) Mechanism of Taxadiene Synthase, a diterpene cyclase that catalyzes the first step of Taxol biosynthesis in Pacific Yew *Biochemistry* **35**: 2968-2977

Long, B.H. and Fairchild, C.R. (1994) Paclitaxel Inhibits Progression of Mitotic Cells to G1 Phase by Interference with Spindle Formation without Affecting Other Microtubule Functions during Anaphase and Telephase. *Cancer Research* **54**: 4355-4361

Malik, S., Cusidó, R.M., Mirjalili, M.H., Moyano, E., Palazón, J. and Bonfill. M. (2011) Production of the anticancer drug taxol in *Taxus baccata* suspension cultures: A review. *Process Biochemistry* **46**: 23-34

Menhard, B., Eisenreich, W., Hylands, P.J., Bacher, A. and Zenk, M.H., (1998) Taxoids from cell cultures of *Taxus chinenesis. Phytochemistry* **49**: 113–125.

Miele, E., Spinelli, G.P., Tomao, F. and Tomao, S. (2009) Albumin-bound formulation of paclitaxel (Abraxane® ABI-007) in the treatment of breast cancer. *International Journal of Nanomedicine* **4**: 99-105

Minderman, H., Brooks, T.A., O'Loughlin, K.L., Ojima, I., Bernacki, R.J. and Baer, M.R. (2004) Broad-spectrum modulation of ATP-binding cassette transport proteins by the taxane derivatives ortataxel (IDN-5109, BAY 59-8862) and tRA96023. *Cancer Chemotheraphy and Pharmacology* **53**: 363-369

Morant, M., Bak, S., Møller, B. L. and Werck-Reichhart, D. (2003) Plant cytochromes P450: tools for pharmacology, plant protection and phytoremediation. *Current Opinion in Biotechnology* **14**: 151-162

Nelson, D.R. (1999) Cytochrome P450 and the Individuality of Species. *Archives of Biochemistry and Biophysics* **369** (1): 1-10

Nicolaou, K.C., Guy, R.K., Pitsinos, E.N. and Wrasidlo, W. (1994a) A Water-Soluble Prodrug of Taxol with Self-Assembling Properties. *Angewandte Chemie International Edition* **33** (15/16): 1583 - 1587

Nicolaou, K.C., Ueno, H., Liu, J-J., Nantermet, P.G., Yang, Z., Renaud, J., Paulvannan, K. and Chadha, R. (1995) Total Synthesis of Taxol. 4. The Final Stages and Completion of the Synthesis. *Journal of American Chemical Society* **117**: 653-659

Nicolaou, K.C., Yang, Z., Liu, J.J., Ueno, H., Nantermet, P.G., Guy, R.K., Claiborne, C.F., Renaund, J., Couladouros, E.A., Paulvannan, K., Sorensen, E.J. (1994b) Total synthesis of Taxol. *Nature* **367**: 630-634

Nicoletti, M.I., Colombo, T., Rossi, C., Monardo, C., Stura, S., Zucchetti, M., Riva, A., Morazzoni, P., Donati, M.B., Bombardelli, E., D'Incalci, M. and Giavazzi, R. (2000) IDN5109, a Taxane with Oral Bioavailability and Potent Antitumour Activity. *Cancer Research* **60**: 842-846

Nims, E., Dubois, C.P., Roberts, S.C. and Walker, E.L. (2006) Expression profiling of genes involved in paclitaxel biosynthesis for targeted metabolic engineering. *Metabolic Engineering* **8**: 384-394

Odgen L.(1988) *Taxus* (yews) - a highly toxic plant. *Vet. Hum. Toxicol.* **30**:563–564.

Ojima, I., Bounaud, P-Y., Takeuchi, C., Pera, P. and Bernacki, R. J. (1998) New taxanes as highly efficient reversal agents for multi-drug resistance in cancer cells. *Bioorganic & Medicinal Chemistry Letters* **8**: 189-194

Ojima. I., Slater, J.C., Kuduk, S.D., Takeuchi, C.S., Gimi,R.H Sun,C-M., Park,Y.H., Pera, P., Veith, J.M and Bernacki, R.J. (1997) Syntheses and Structure-Activity Relationships of Taxoids Derived from 14β-Hydroxy-10-deacetylbaccatin III. *Journal of Medicinal Chemistry* **40**: 267-278

Paine, J.A., Shipton, C.A., Chaggar, S., Howells, R.M., Kennedy, M.J., Vernon, G., Wright, S.Y., Hinchliffe, E., Adams, J.L., Silverstone, A.L. and Drake, R. (2005) Improving the nutritional value of Golden Rice throught increased pro-vitamin A content. *Nature Biotechnology* **23** (4) 482-487

Pandi, M., Kumaran, R.S., Choi, Y-K., Kim, H.J. and Muthumary, J. (2011) Isolation and detection of taxol, an anticancer drug produced from *Lasiodiplodia theobromae*,

an endophytic fungus of the medicinal plant *Morinda citrifolia*. *African Journal of Biotechnology* **10**(8): 1428-1435

Pompon, D., Louerat, B., Bronine, A. and Urban, P. (1996) Yeast expression of animal and plant P450s in optimized redox environments. *Methods in Enzymology* **272**:51–64

Ro, D-K., Arimura, G-I., Lau, S.Y.W., Piers, E. and Bohlmann, J. (2005) Loblolly pine abietadienol/abietadienal oxidase *PtAO* (CYP720B1) is a multifunctional, multusubstrate cytochrome P450 monooxygenase. *Proceedings of National Academic of Sciences USA* **102**: 8060-8065

Roberts, S.C., Naill, M., Gibson, D.M. and Shuler, M.L. (2003) A simple method for enhancing paclitaxel release from *Taxus Canadensis* cell suspension cultures utilizing cell wall digesting enzymes. *Plant Cell Reports* **21**: 1217-1220

Rohr, J. (1997) Biosynthesis of Taxol. *Angewandte Chemie International Edition* **36**: 2190-2195

Rontein, R., Onillon, R., Herbette, G., Lesot, A., Werck-Reichhart, D., Sallaud, C., Tissier, A. (2007) CYP725A4 from Yew Catalyzes Complex Structural Rearrangement of Taxa-4(5),11(12)-diene into the Cyclic Ether 5(12)-Oxa-3(11)-cyclotaxane.*The Journal of Biological Chemistry* **283** (10): 6067-6075

Rubenstein, S.M. and Williams, R.M (1995) Studies on the Biosynthesis of Taxol: Total Synthesis of Taxa-4(20),11(2)-diene and Taxa-4(5),11(2)-diene. The first committed biosynthetic intermediate. *Journal of Organic Chemistry* **60**: 7215-7223

Sambrooks, J., Fritsch, E.F., Maniatis, T. (1989) *Molecular cloning: a laboratory Manual.* 2nd Edition. Cold Spring Harbour Laboratory Press, New York

Schiff, P.B., Fant, J. and Horwitz, S.B. (1979) Promotion of microtubule assembly *in vitro* by taxol. *Nature* **277**: 665–667.

Schoendorf, A., Rithner, C.D., Williams, R.M., and Croteau, R.B. (2001) Molecular cloning of a cytochrome P450 taxane 10β-hydroxylase cDNA from *Taxus* and functional expression in yeast *Proceedings of the National Academy of Sciences USA* **96**; 1501-1506

Shi, Q-W. and Kiyota, H. (2005) New Natural Taxane Diterpenoids from *Taxus* species since 1999. *Chemistry & Biodiversity* **2**: 1597-1623

Shimomura, O., Johnson, F.H. and Saiga, Y. (1962) Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, *Aequorea*. *Journal of Cellular Comparative Physiology* **59**: 223-239.

Snyder, J.P., Nettles, J.H., Cornett, B., Downing, K.H. and Nogales, E. (2001) The binding conformation of Taxol in β-tubulin: A model based on electron

crystallographic density. *Proceedings of National Academic of Sciences USA* **98** (9): 5312-5316

Stierle, A., Strobel, G. and Stierle, D. (1993) Taxol and taxane production by *Taxomyces andreannae*, an endophytic fungus of Pacific yew. *Science* **260**: 214-226

St-Pierre, B., Laflamme, P., Alarco, A.M., and De Luca, V. (1998) The terminal O-acetyltransferase involved in vindoline biosynthesis defines a new class of proteins responsible for coenzyme A-dependent acyl transfer. *Plant Journal* **14**: 703-713.

Strobel, G.A., Stierle, A. and Hess, W.M. (1993). Taxol formation in yew – *Taxus* . *Plant Science* **92**: 1-12

Strobel, G., Yang, X.S., Sears, J., Kramer, R., Sidhu, R.S. and Hess, W.M. (1996) Taxol from *Pestalotiopsis microspora*, an endophytic fungus *of Taxus wallachiana*. *Microbiology* **142**: 435-440

Suffness, M. and Wall, M.E. (1995) Discovery and development of Taxol. *Taxol Science and applications*. CRC Press: 97-121

Tissier, A., Sallaud, C., and Rontein, D. (2006) Essai au champ pluriannuel de tabac dont la production en cembranes a été fortement diminuée par ARN interference. *Librophyt SAS* French Patent AN FR2006:000188

Torres, K. and Horwitz, S.B. (1998) Mechanisms of Taxol-induced Cell Death are Concentration Dependent. *Cancer Reseach* **58**: 3620-3626

Ueda, Y., Matiskella, J.D., Mikkilineni, A.B., Farina, V., Knipe, J.O., Rose, W.C., Casazza, A.M. and Vyas, D.M. (1995) Novel Water-Soluble Phosphate Derivatives Of 2'-Ethoxy Carbonylpaclitaxel As Potential Prodrugs Of Paclitaxel: Synthesis And Antitumour Evaluation. *Bioorganic & Medicinal Chemistry Letters* **5**(3): 247-252

Walker, K.D. and Croteau, R. (2000a) Molecular cloning of a 10-deacetylbaccatin III-10-*o*-acetyl transferase cDNA from *Taxus* and functional expression in *Escherichia coli. Proceedings of the National Academy of Sciences USA* **97**: 583-587

Walker, K.D. and Croteau, R. (2000b) Taxol biosynthesis: molecular cloning of a benzoyl-CoA:taxane 2α-*O*-benzoyltransferase cDNA from *Taxus* and functional expression in *Escherichia coli. Proceedings of the National Academy of Sciences USA* **97**:13591-13596

Walker, K.D., Ketchum, R.E.B, Hezari, M., Gatfield, D., Goleniowski, M., Barthol, A. and Croteau, R. (1999) Partial purification and characterization of acetyl coenzyme A: taxa-4(20),11(12)-dien-5α-ol-*o*-acetyl-transferase that catalyses the first acetylation step of Taxol biosynthesis. *Archives of Biochemistry and Biophysics* **464**: 273-279

Walker, K.D., Klettke, K., Akiyama, T. and Croteau, R. (2004) Cloning, heterologous expression, and characterization of a phenylalanine aminomutase involved in taxol biosynthesis. *Journal of Biological Chemistry* **279**: 53947– 53954.

Walker, K., Schoendorf, A. and Croteau, R. (2000) Molecular cloning of a taxa-4(20),11(12)-dien-5α-ol-*O*-acetyl transferase cDNA from *Taxus* and functional expression in *Escherichia coli*. *Archives of Biochemistry and Biophysics* **374**: 371–380

Walker, K., Fujisaki, S., Long, R. and Croteau, R. (2002a) Molecular cloning and heterologous expression of the C-13 phenylpropanoid side chain-CoA acyltransferase that functions in Taxol biosynthesis. *Proceedings of the National Academy of Sciences USA* **99**: 12715–12720.

Walker, K., Long, R. and Croteau, R. (2002b) The final acylation step in Taxol biosynthesis: Cloning of the taxoid C13-side-chain N-benzoyltransferase from *Taxus*. *Proceedings of the National Academy of Sciences USA*: **99** (14): 9166-9171

Wani, M.C., Taylor, H.L., Wall, M.E., Coggon, P. and McPhail, A.T. (1971) Plant antitumor agents VI: The isolation and structure of Taxol, a novel antileukemic and anti- tumor agent from *Taxus brevifolia*. *Journal of American Chemical Society* **93**: 2325–2327.

Wang, J., Li, G., Lu, H., Zheng, Z., Huang, Y. and Su, W. (2000) Taxol from *Tubercularia* sp. strain TF5, an endophytic fungus of *Taxus mairei*. *FEMS Microbiology Letters* **193**: 249-253

Wang, W., Shi, Q., Outang, T., Zhu, P. and Cheng, K (2002) Cloning, expression and characterization of taxadiene synthase, a diterpene cyclase from *Taxus chinensis*. *Acta Botanica Sinica* **44** (2) : 181 - 187

Werck-Reichhart, D. and Feyereisen, R. (2000) Cytochromes P450: a success story. *Genome Biology* **1**(6): 1-9

Wheeler, A.L., Long, R.M., Ketchum, R.E.B., Rithner, C.D., Williams, R.M. and Croteau, R. (2001) Taxol Biosynthesis: Differential Transformations of Taxadien-5a-ol and Its Acetate Ester by Cytochrome P450 Hydroxylases from *Taxus* Suspension Cells. *Archives of Biochemistry and Biophysics* **390**(2): 265–278

Wildung, M.R. and Croteau, R. (1996) A cDNA clone for taxadiene synthase, the diterpene cyclase that catalyzes the committed step of Taxol biosynthesis. *Journal of Biology and Chemistry* **271**: 9201-9204

Williams, D.C., Carroll, B.J., Jin, Q., Rithner, C.D., Lenger, S.R., Floss, H.G., Coates, R.M. Williams, R.M. and Croteau,R (2000a) Intramolecular proton transfer in the cyclization of geranylgeranyl diphosphate to the taxadiene precursor of Taxol catalyzed by recombinant taxadiene synthase. *Chemical Biology* **7**: 969–977.

Williams, D.C., Wildung, M.R., Jin, A.Q., Dalal, D., Oliver, J.S., Coates, R.M. and Croteau, R. (2000b) Heterologous expression and characterization of a "pseudomature" form of taxadiene synthase involved in paclitaxel (Taxol) biosynthesis and evaluation of a potential intermediate and inhibitors of the multistep diterpene cyclization reaction. *Archives in Biochemistry and Biophyics.* **379**:137–146.

Wilson, C.R., Sauer, J-M. and Hooser, S.B. (2001) Taxines: a review of the mechanism and toxicity of yew (*Taxus* spp.) alkaloids. *Toxicon* **39**: 175-185

Wu, S., Schalk, M., Clark, A., Miles, R.B., Coates, R. and Chappell, J. (2006) Redirection of cytosolic or plastidic isoprenoid precursors elevates terpene production in plants. *Nature biotechnology* **24** (11): 1441-1447

Wurbs, D., Ruf, S. and Bock, R. (2007) Contained metabolic engineering in tomatoes by expression of carotenoid biosynthesis genes from the plastid genome. *The Plant Journal* **49**: 276-288

Ye, X., Al-Babili, S., Klöti, A., Zhang, J., Lucca, P., Beyer, P. and Potrykus, I. (2000) Engineering the Provitamin A (β-Carotene) Biosynthetic Pathway into (Carotenoid-Free) Rice Endosperm. *Science* **287**: 303-305

Young, D.H., Michelotti, E.L., Swindell, C.S and Krauss, N.E (1992) Antifungal properties of Taxol and various analogues. *Experientia* **48**: 882-885

Yukimune, Y., Tabata, H., Higashi, Y. and Hara, Y. (1996) Methyl jasmonate-induced overproduction of paclitaxel and baccatin III in *Taxus* suspension cells. *Phytochemistry* **54**: 13–17

Zamir, L. and Caron, G. (2003) Semi-synthesis of a protected baccatin iii compound. Patent No: US 6,576,777 B2

Zabarovsky, E.R. and Winberg, G. (1990) High efficiency electroporation of ligated DNA into bacteria. *Nucleic Acids Research* **18**: 5912

Zhang, C. and Fevereiro, P.S. (2007) The Effect of Heat Shock on Paclitaxel Production in *Taxus yunnanensis* Cell Suspension Cultures: Role of Abscisic Acid Pretreatment. *Biotechnology and Bioengineering* **96** (3): 506-514

Zhang, C.H. and Wu, J-Y. (2003) Ethylene inhibitors enhance elicitor-induced paclitaxel production in suspension cultures of *Taxus* spp. Cells. *Enzyme and Microbial Technology* **32**: 71–77

Zhang, C.H., Mei, X.G., Liu, L. and Yu, L.J. (2000) Enhanced paclitaxel production induced by the combination of elicitors in cell suspension cultures of *Taxus chinensis. Biotechnology Letters* **22**: 1561–1564

Zhong, S., Lin, Z., Fray, R. and Grierson, D. (2008) Improved plant transformation vectors for fluorescent protein tagging. *Transgenic Research* **17**: 985-989

Zubrod, C.G., Schepartz, S.A. and Leiter, J. (1966) The chemotherapy program of the NCI: history, analysis, and plans. *Cancer Chemotherapy Reports* **50**: 349–355.

APPENDICES

Appendix I

Details of the primers used for the cloning of plasmids, RT-PCR analysis and for probe preparation for northern analysis. *x* – *AttB* adaptor sites

Gene/Vector name	Primer Name	Sequences 5′ ⟶ 3′
T5OH	5tax 5alp	CAC CAA AAT GGA CGC CCT GTA TAA GAG C
	3tax 5alp	TTC TCC TTC AAT TGA CTA TGG TCT CGG
	35S-Promoter-F	CTATCCTTCGCAAGACCCTTC
	T5OH-F	ATG GAC GCC CTG TAT AAG AGC
	T5OH-Rns	TGG TCT CGG AAA CAG TTT AAT G
	T5OH-PstRev	GCCCGTCCCTGGAGTGCTCTATG
	5a-OH-MidF	ATGATAAGCAGGAACAGGATC
TXS	TaxSynRNS	TAC TTG AAT TGG ATC AAT ATA AAC TTT TC
	TaxSynF	ATG GCT CAG CTC TCA TTT AAT G
	TaxMid2	CCATCTTTGAGTAATTCAAAATGTTTAG
	TaxMid1	CCCTGTATGTATTCCATCGATCTGCTGG
	Taxint35	ATAAACCTCCGCCACTCGGTGTC
	TaxSyn-R	TCATACTTGAATTGGATCAATATAAACTT
	Taxint53	CCTGAGCTCCCCTGCCTCCACTG
T10BOH	5Tax10hyd	GTCGAAATGAATACCTTCAGTTTCCTC
	3Tax10hyd	TGATTACGATCTGGGAAATACTTTCAG
T5AT	Acetyl-Rns	TAC TTT AGC CAC ATA TTT TTT CAT CA
	5Tax5acyl	TTAGA ATGG AGA AGA CAG ATT TAC AC
	3Tax5acyl	TCGTTGCACTTTCATACTTTAGCCAC
T13OH	13αOHASP	*aga aag ctg ggt*A TTA AGA TCT GGA ATA GAG
	13αOH SP	*aaa aag cag gct* ATG GAT GCC CTT AAG CAA
	13αOH-MidF	GTTTCCCCTTCCATTCTTTTCGTTACC
	13AOH-MidR	TTCAATAAACTGCTGTGGAGTGCTCG
	13αOH-RNS	AGATCTGGAATAGAGTTTAATGGG
	13αOH-F	ATGGATGCCCTTAAGCAATTG
T7OH	7βOH-R	TCAGGATCTGGCGATAAGTTTTATTGG
	7βOH-F	ATGGATGCCCTTTCTCTTGTAAAC
TXS-T5OH Vector	TXSRev-Sbf1/Pst1	CATTCTGCAGGCGCGCCAGGAAACAGCTAT
	TXSF-Pst1	CCACCTGCAGCCATGGAGTCAAAGATTC
	Txs-Terminator	CGTGTTGTTGACCGGGCCTTG
pCR8 primers	*att*B1 adapter	GG GGA CAA GTT TGT ACA AAA AAG CAG GCT
	*att*B2 adapter	TGG GTC GAA AGA ACA TGT TTC ACC AGG GG
	GW1	GTTGCAACAAATTGATGAGCAATGC
	GW2	GTTGCAACAAATTGATGAGCAATTA

Appendix II

Phenol solutions (In a fume hood)

Phenol Chloroform

200 g Phenol
400 mL 10 mM Tric Cl, pH 7.5
0.2 g 8-hydroxyquinolene
200 mL Chloroform
8 mL Isoamylalcohol
Stir on magnetic stirrer for 15 minutes, allow the phases to separate overnight.

Phenol mixture

100 g Phenol
14 ml meta-cresol
0.1 g 8-hydroxyquinolene
30 ml water

Growth media

Luria-Bertani (LB) Media

10 g Tryptone
5 g Yeast extract
10 g NaCl
Adjust pH to 7.0 with 5N NaOH
SDW to 1L
Autoclave at 121°C for 30 mins (15 psi)

LB agar (500 ml)

5 g Tryptone
2.5 g Yeast extract
2.5 g NaCl
3 g Agar
SDW to 500 mL
Autoclave at 121°C for 30 mins (15 psi)

Liquid MS (1L)

4.3 g Murashige Skoog Salts
30 g Sucrose
Adjust pH to 5.8
Autoclave at 121°C for 30 mins (15 psi)

MSR3 medium (1L) (For tobaccao micropropagation)

4.3 g Murashige Skoog Salts
30 g Sucrose
6.5 g Agar
Adjust pH to 5.8
Autoclave at 121°C for 30 mins (15 psi)

M1 medium (1L)

4.3 g Murashige Skoog Salts
30 g Sucrose
6.5 g Agar
Adjust pH to 5.8
Autoclave at 121°C for 30 mins (15 psi)
When ready to pour (at about 60°C) add:
- 500 mg/L Indole Acetic Acid (IAA)
- 750 mg/L Zeatin
- 0.01 g/L Nystatin

M13 or Regeneration medium (1L)

4.3 g Murashige Skoog Salts
30 g Sucrose
6.5 g Agar
Adjust pH to 5.8
Autoclave at 121°C for 30 mins (15 psi)
When ready to pour (at about 60°C) add:
- 400 mg/L Augmentin
- 500 mg/L Indole Acetic Acid (IAA)
- 750 mg/L Zeatin
- 70 mg/L Kanamycin
- 0.01 g/L Nystatin
- 5 mg/L Tetracyclin

General Solutions

Alexander's strain solution was prepared by adding the following constituents in the order given below and stored in the dark.

10 mL 95% alcohol
1 mL Malachite green (1% solution in 95% alcohol)
50 mL Distilled water
25 mL Glycerol
5 mL Acid fuchsin (1% solution in water)
0.5 mL Orange G (1% solution in water)
4 mL Glacial acetic acid
Add distilled water (4.5 mL) to a total of 100 mL.

Basta (DL – phosphinothricin) (10 mg/ml)
Add 10 ml sterile distilled water to 100 mg Basta
Filter sterilisation through Millipore 0.22 µM
Store at -20°C

IAA (1 mg/ml) (Indole Acetic Acid)

Add 1N KOH drop wise to 100 mg IAA until dissolved.
Make up to 100 mL with SDW and store at -20°

Kanamycin Monosulphate (100 mg/L)

Add 10 ml sterile distilled water to 1 g Kanamycin
Filter sterilisation through Millipore 0.22 µM
Store at -20°C

Nystatin (20mg/ml)

Add 10 ml methanol to 200 mg Nystatin
Store at -20°C

Plant RNA Extraction buffer (500 ml)

5 g of 1,5-Naphthalendisulfonic acid Disodium
30 g of Sodium 4-aminosalicylate
25 ml Phenol mixture
The volume was adjusted to 500 ml and transferred to either a brown bottle or a bottle covered with Aluminium foil to prevent exposure to light. The extraction buffer was

stored at 4°C and shaken prior to use.

Rapid DNA extraction buffer

200 mM Tric Cl, pH 7.5
250 mM NaCl
25 mM EDTA
0.5 % SDS (Sodium dodecyl sulphate)

Rifampicin (50 mg/ml)

Add 10 ml methanol to 0.5 g Rifampicin
Store at -20°C

R3 vitamins (1L)

1 g Thiamine
0.5 g Nicotinic acid
0.5 g Pyridoxine
Filter sterile with 20 µm Millipore filters

STET solution

10 mM Tris-Cl (pH 8.0)
0.1 M NaCl
1 mM EDTA (pH 8.0)
5% (v/v) Triton X-100

1.1 M spermidine stock

1 g Spermidine powder (S0266-1G, SIGMA)
68 ml H_2O
(The solution deaminates quickly and releases NH_3. It can only be stored at -20°c for 1 month)

Tetracycline (5mg/ml)

Add 10 ml of 70% EtOH to 50 mg Tetracycline
Store at -20°C

5 X TBE Buffer

108 g Tris
55 g Boric Acid
40 mL 0.5 EDTA, pH 8.0

Zeatin (1mg/ml)

Add 1N KOH drop wise to 100 mg Zeatin until dissolved.
Make up to 100 mL with SDW and store at -20°C

Sodium Phosphate Buffer (1M, pH 6.6)

35.2 ml of 1M Disodium hydrogen phosphate
64.8 ml of 1M Sodium dihydrogen phosphate

20 x SSC (Saline-Sodium citrate) (1L)

175.3 g Sodium chloride
88.2 g Sodium citrate
Dissolved with deionised water and make up to 1L
Adjust pH with HCl to pH 7.0
Autoclave at 121°C for 30 mins to sterilise

Northern Blot Gel Running Buffer

1 g agarose
90 ml SDW
2 ml Sodium Phosphate buffer (1M, pH 6.6)
8 ml formaldehyde pH7 (add when cool)

Running buffer for Northern Gel

593.5 ml SDW
50 ml formaldehyde pH7
6.5 ml Sodium phosphate buffer (1M, pH 6.6)

Ethidium bromide buffer

1000 µl deionised formamide
500 µl SDW
330 µl formaldehyde

40 µl of 0.5 M EDTA (pH 8)
40 µl of 1M Sodium phosphate buffer, pH 6.5

Bromophenol blue loading dye

0.25 % (w/v) Bromophenol Blue
0.25 % (w/v) Xylenecynol FF
40 % (w/v) sucrose in SDW
Filter sterilise

Pre-hybridisation solution

50 % (v/v) deionised formamide
20 % (v/v) 5 M NaCl
16 % (v/v) of 50% Dextran sulphate solution
10 % (v/v) of 10% SDS (sodium dodecyl sulfate) solution (added last)
4 % (v/v) of 200 µg/ml Salmon sperm DNA (ssDNA)
ssDNA was denatured at 95 °C for 10 minutes prior to adding to the pre-hybridisation solution.

50 x TAE (per litre)

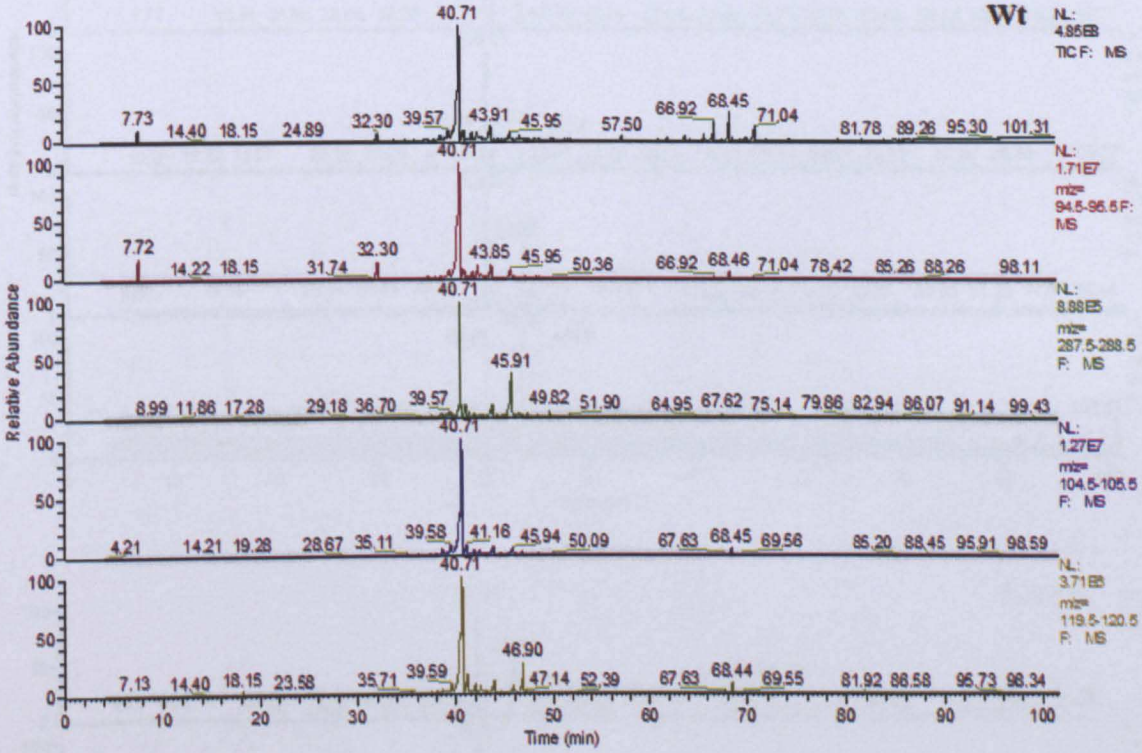
Tris Base 242 g
Glacial HOAc 57.1 ml
0.5M EDTA 8.0 100 ml
Add deionised water to 500 ml
Autoclave at 121°C for 30 mins (15 psi) to sterilise

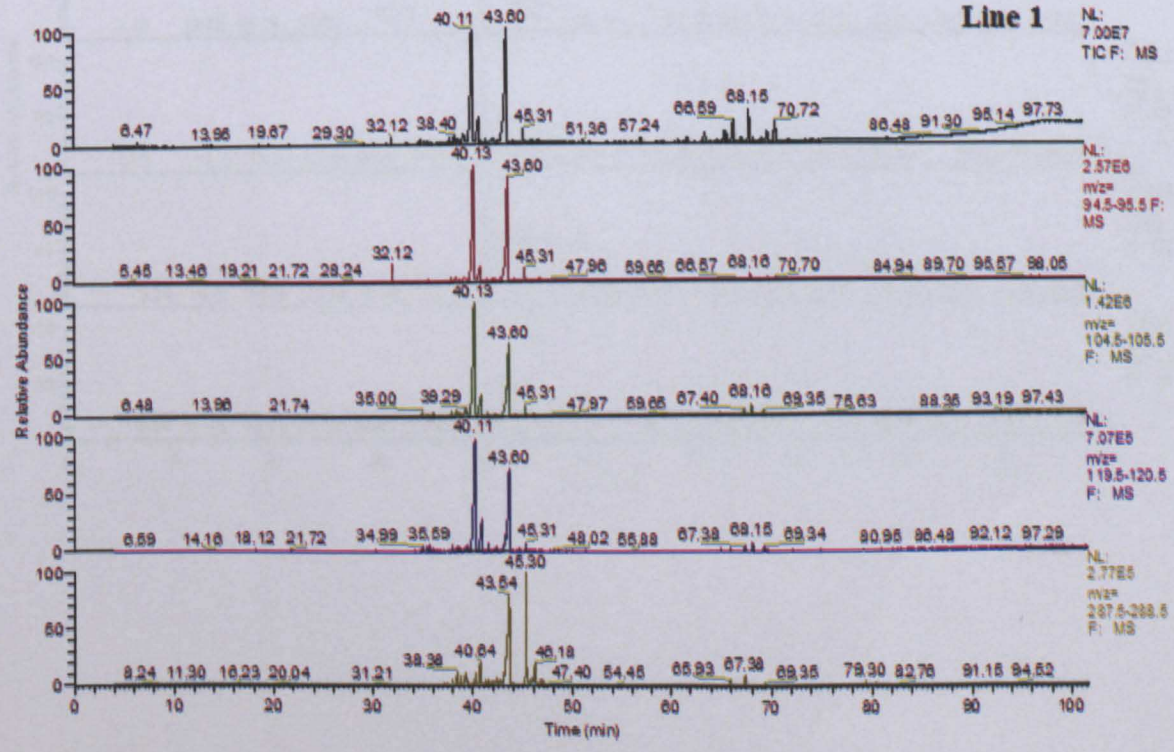
SOC Medium (1x Stock) -for higher efficiency transformation/ liter

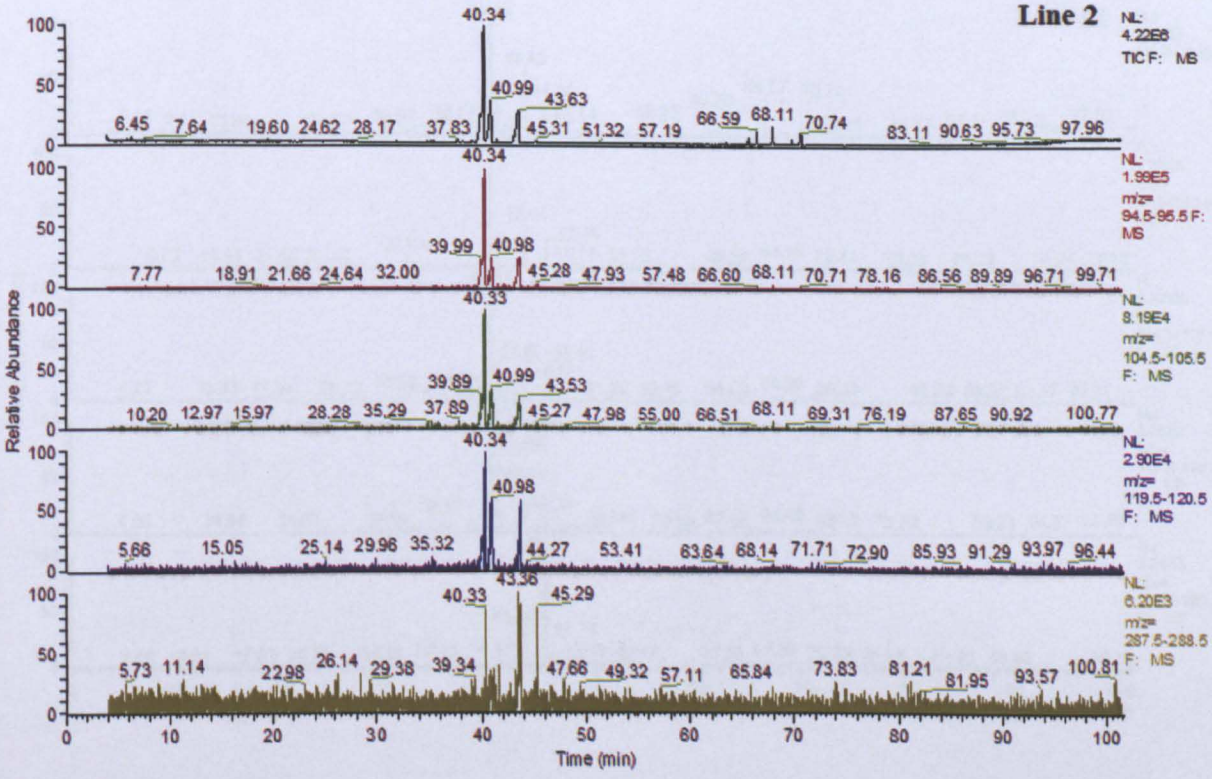
Bacto-Yeast Extract 5 g
Bacto-Tryptone 20 g
5M NaCl 2 ml
KCl 0.19 g
1M MgCl2 10 ml
1M MgSO4 10 ml
D-glucose 3.6 g

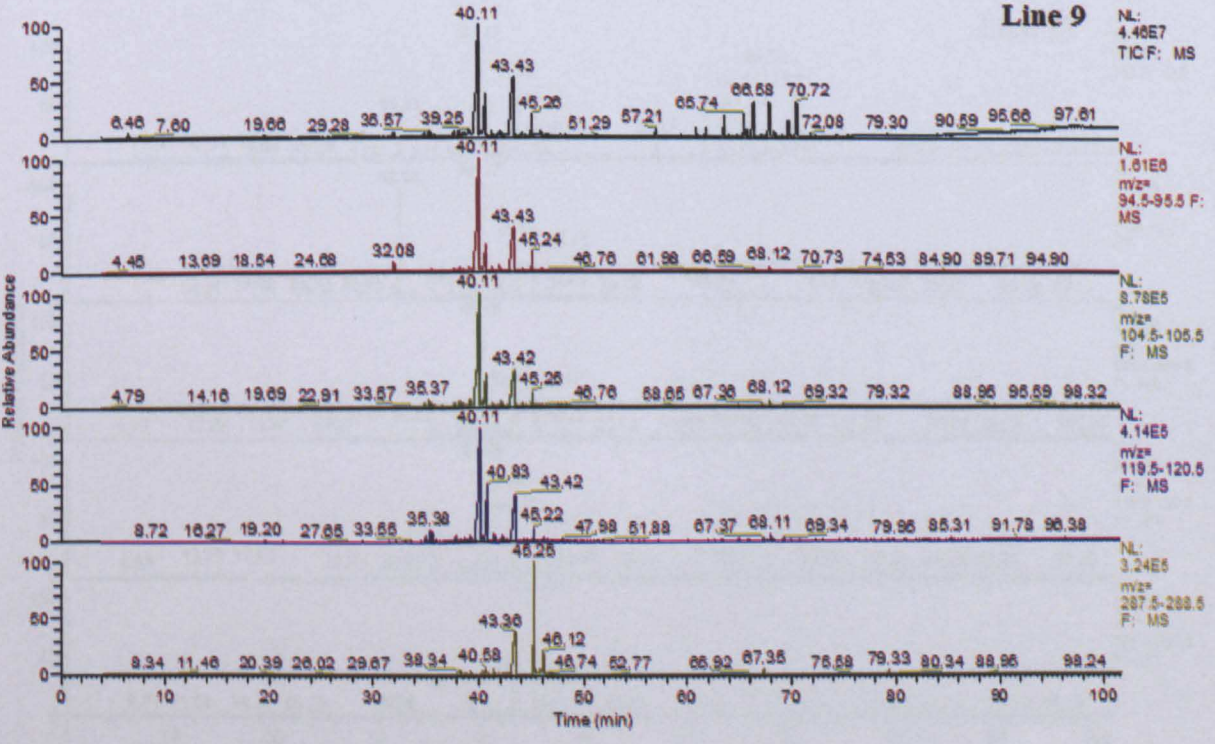
Appendix III

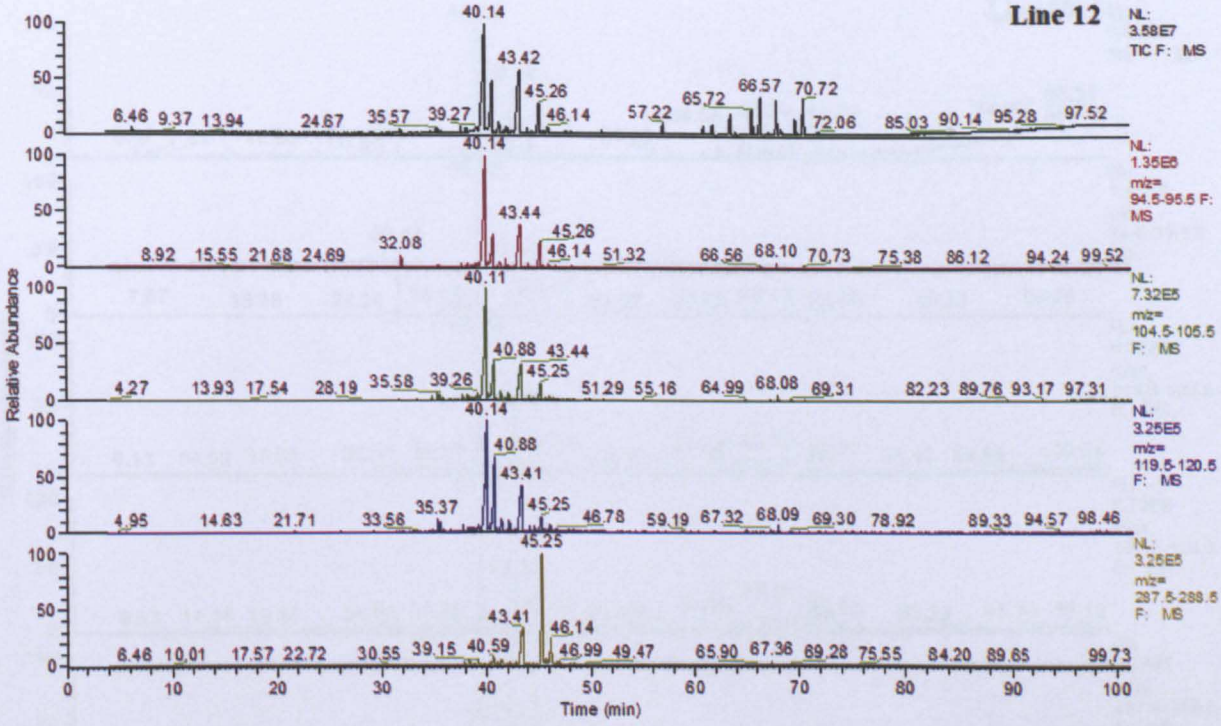
Chromatograms for each extract from tobacco lines D expressing TXS and T5OH, scanned for taxadiene-5α-ol ions.

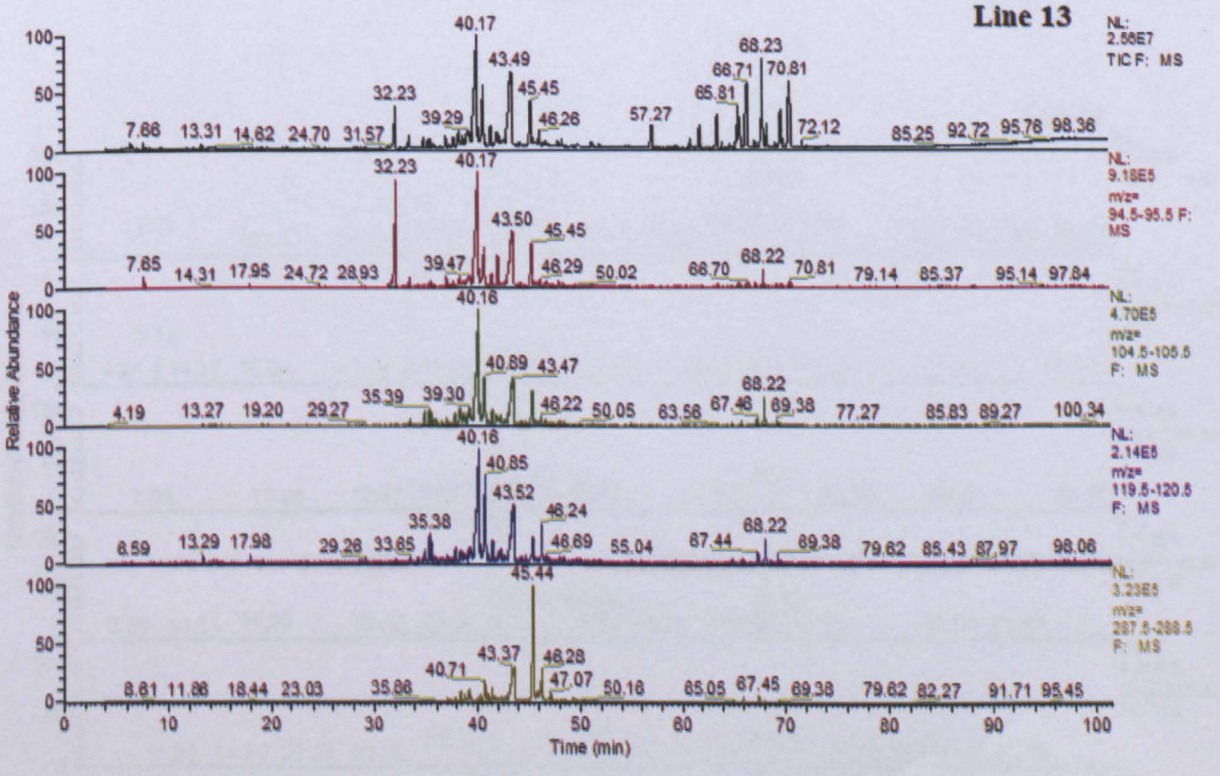


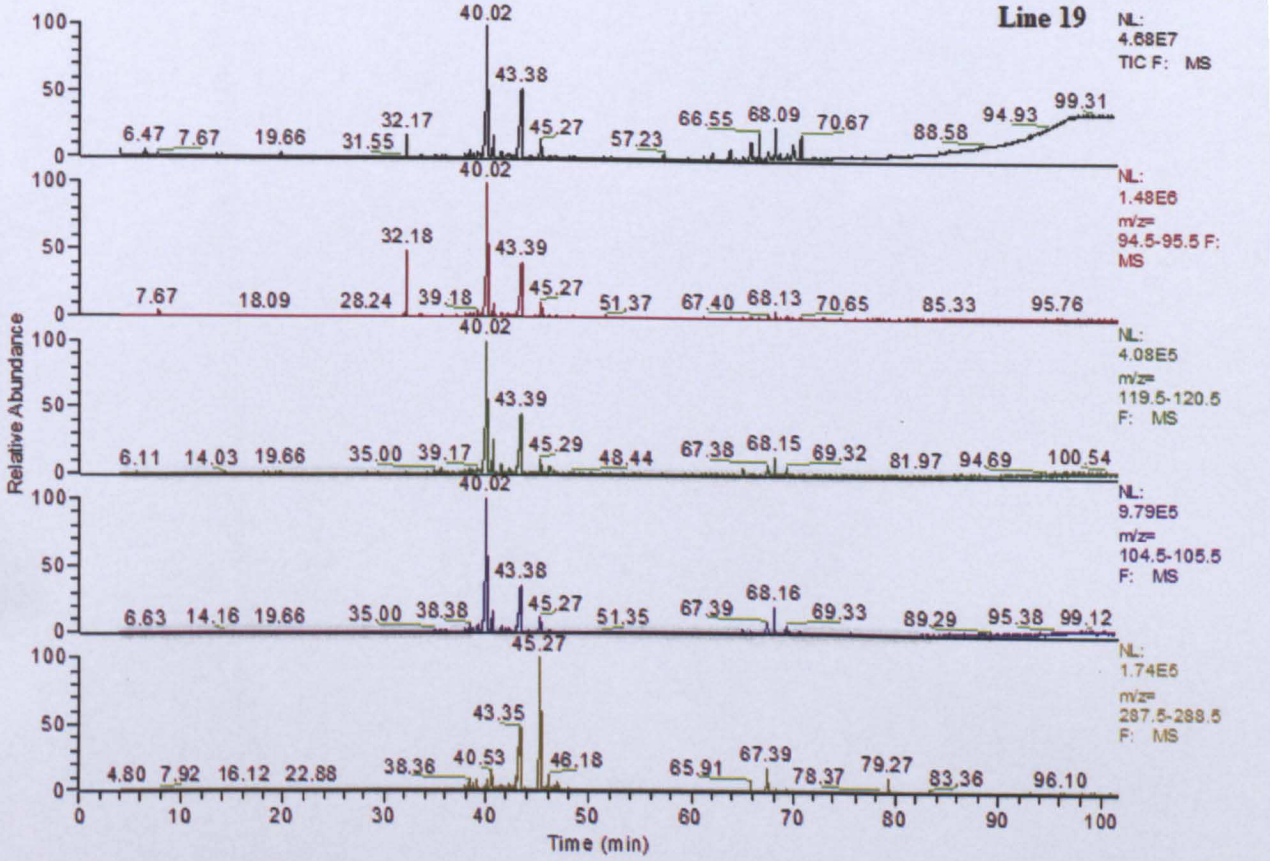


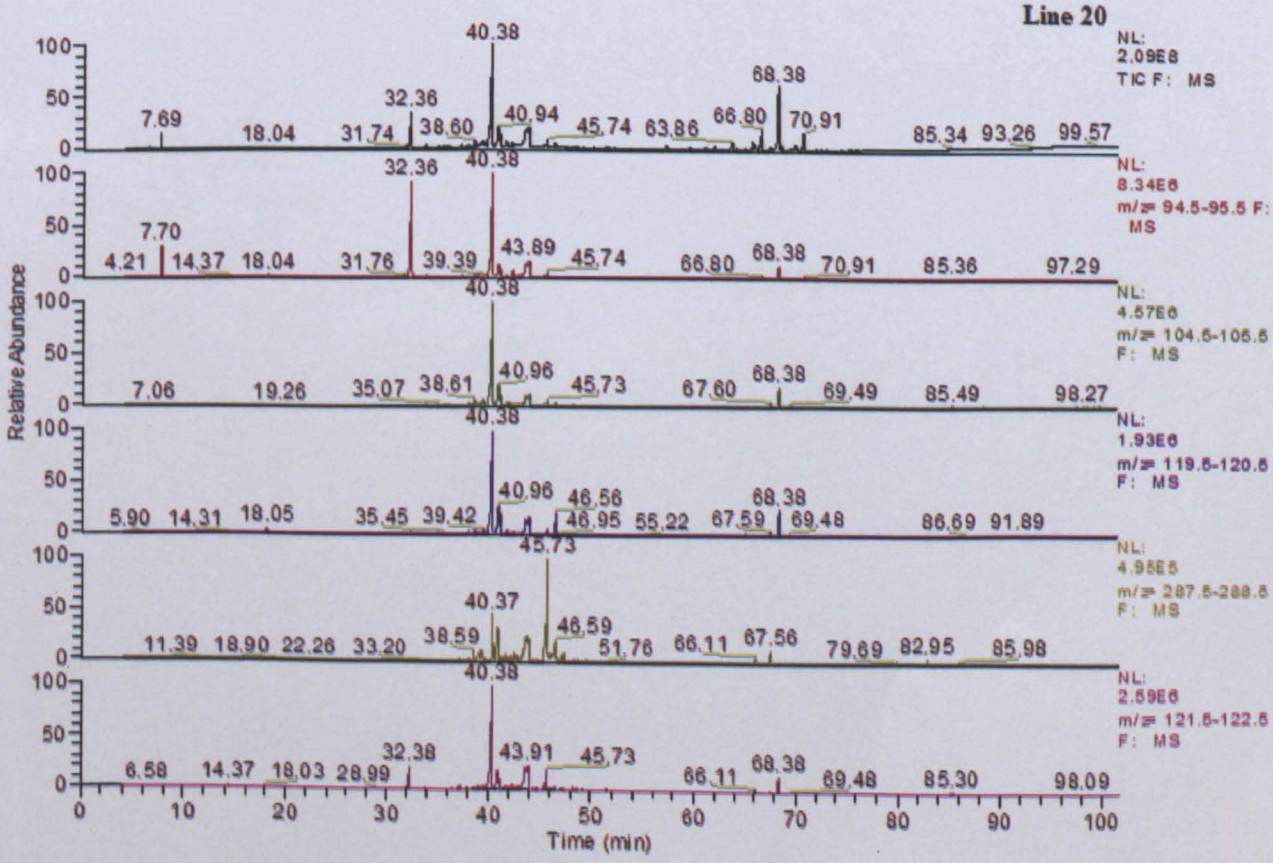

Line 2
NL: 4.22E6 TIC F: MS
NL: 1.99E5 m/z= 94.5-95.5 F: MS
NL: 8.19E4 m/z= 104.5-105.5 F: MS
NL: 2.90E4 m/z= 119.5-120.5 F: MS
NL: 6.20E3 m/z= 287.5-288.5 F: MS
Relative Abundance
Time (min)


Line 9
NL: 4.46E7 TIC F: MS
NL: 1.61E6 m/z= 94.5-95.5 F: MS
NL: 8.78E5 m/z= 104.5-105.5 F: MS
NL: 4.14E5 m/z= 119.5-120.5 F: MS
NL: 3.24E5 m/z= 287.5-288.5 F: MS
Relative Abundance
Time (min)


Line 12
NL: 3.58E7 TIC F: MS
NL: 1.35E6 m/z= 94.5-95.5 F: MS
NL: 7.32E5 m/z= 104.5-105.5 F: MS
NL: 3.25E5 m/z= 119.5-120.5 F: MS
NL: 3.26E5 m/z= 287.5-288.5 F: MS
Relative Abundance
Time (min)


Line 13
NL: 2.56E7 TIC F: MS
NL: 9.18E5 m/z= 94.5-95.5 F: MS
NL: 4.70E5 m/z= 104.5-105.5 F: MS
NL: 2.14E5 m/z= 119.5-120.5 F: MS
NL: 3.23E5 m/z= 287.5-288.5 F: MS
Relative Abundance
Time (min)


Line 19
NL: 4.68E7 TIC F: MS
NL: 1.48E6 m/z= 94.5-95.5 F: MS
NL: 4.08E5 m/z= 119.5-120.5 F: MS
NL: 9.79E5 m/z= 104.5-105.5 F: MS
NL: 1.74E5 m/z= 287.5-288.5 F: MS
Relative Abundance
Time (min)


Line 20
NL: 2.09E8 TIC F: MS
NL: 8.34E6 m/z= 94.5-95.5 F: MS
NL: 4.57E6 m/z= 104.5-105.5 F: MS
NL: 1.93E6 m/z= 119.5-120.5 F: MS
NL: 4.95E5 m/z= 287.5-288.5 F: MS
NL: 2.59E6 m/z= 121.5-122.5 F: MS
Relative Abundance
Time (min)