Chemical, biochemical, and molecular characterization of a low vindoline *Catharanthus roseus* mutant.

Alison Edge, BSc

Biotechnology

Submitted in partial fulfillment

of the requirements for the degree of

Master of Science

Faculty of Mathematics and Science, Brock University

St. Catharines, Ontario

Abstract

The Madagascar periwinkle (*Catharanthus roseus*) is the sole source of the anticancer drug vinblastine, which is formed via the coupling of monoterpenoid indole alkaloids (MIAs) catharanthine and vindoline. A mutant line of *C. roseus* (M2-1865) with an altered MIA profile was identified in a screen of 4000 M2 lines generated by ethylmethanesulfonate (EMS) chemical mutagenesis. While this line did not accumulate vinblastine due to reduced levels of vindoline within the leaves, significant levels of 2,3-epoxide derivatives of tabersonine accumulated on the leaf surface.

Detailed nucleotide, amino acid, and enzyme activity analyses of tabersonine 3-reductase in the M2-1865 line showed that a single amino acid substitution (H189Y) diminished the biochemical activity of T3R by 95%. Genetic crosses showed the phenotype to be recessive, exhibiting standard Mendelian single-gene inheritance. The usefulness of EMS mutagenesis in elucidating MIA biosynthesis is highlighted by the results of this study.

Acknowledgements

First and foremost, I would like to thank my supervisor, Professor Vincenzo De Luca. His vast knowledge, advice, and guidance have been greatly beneficial during the past three years, and this project would not have been possible without his support and assistance. My committee members, Professor Tony Yan and Professor Charles Després, have provided indispensable resources and input for which I am very grateful.

I would also like to express my gratitude to post-doctoral fellow Dr. Vince Qu, who identified the genes that made completion of this research possible and who offered endless assistance throughout the project; Dr. Kyunghee Kim for her help, patience, and maintenance in regards to our UPLC-MS equipment; undergraduate student Michael Easson for performing the mutant screening; and our other past and present lab members for their friendship and advice during my project: Dr. Zerihun Demissie, Brent Wiens, Paulo Cázares Flores, Trevor Kidd, and Antje Thamm.

Table of Contents

List of Tables

List of Figures

Abbreviations

7DLH – 7-deoxyloganin 7-hydroxylase

7DLS – 7-deoxyloganetic acid synthase

10HGO – 10-hydroxygeraniol oxidoreductase

16OMT – 16-hydroxytabersonine 16-*O*-methyltransferase

ABA – Abscisic acid

ADH – Alcohol dehydrogenase

ADHL – ADH-like

AdoMet – *S*-adenosyl-L-methionine

CAD – Cinnamyl alcohol dehydrogenase

cDNA – Complementary DNA

CPR – Cytochrome P450 reductase

CYP – Cytochrome P450

D4H – Desacetoxyvindoline 4-hydroxylase

DAT – Deacetylvindoline 4-*O*-acetyltransferase

DL7H – 7-deoxyloganic acid 7-hydroxylase

DLGT – Deoxyloganetic acid-*O*-glucosyltransferase

DNA – Deoxyribonucleic acid

DTT – Dithiothreitol

EDTA – Ethylenediaminetetraacetic acid

EMS – Ethylmethanesulfonate

ER – Endoplasmic reticulum

G10H – Geraniol-10-hydroxylase

GES – Geraniol synthase

GFW – Gram fresh weight

GPP – Geranyl pyrophosphate

HEPES – 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

IPAP – Internal phloem-associated parenchyma

IPTG – Isopropyl β-D-1-thiogalactopyranoside

IRS – Iridoid synthase

JA – Jasmonic acid

K_d – Dissociation constant

K_M – Michaelis-Menten constant

LAMT – Loganic acid *O*-methyltransferase

LB – Luria-Bertani

LP – Leaf pair

MAT – minovincinine-19-hydroxy-*O*-acetyltransferase

MDR – Medium-chain dehydrogenase/reductase

MeJA – Methyl jasmonate

MEP – 2-C-methyl-D-erythritol-4-phosphate

MIA – Monoterpenoid indole alkaloid

mRNA – Messenger RNA

MS – Mass spectrometry

NADH/$NADP^+$ – Nicotinamide adenine dinucleotide

NADPH/$NADP^+$ – Nicotinamide adenine dinucleotide phosphate

NCBI – National Center for Biotechnology Information

NMR – Nuclear magnetic resonance

NMT – 16-methoxytabersonine-2,3-dihydro-3-hydroxytabersonine *N*-methyltransferase

NTA – Nitrilotriacetic acid

OD – Optical density

ORCA – Octadecanoid-responsive Catharanthus AP2/ERF-domain

ORF – Open reading frame

PCR – Polymerase chain reaction

PMSF – phenylmethanesulfonyl fluoride

PRX1 – Vacuolar class III peroxidase

PSPG – Plant secondary product glycosyltransferase

PVDF – Polyvinylidene fluoride

RNA – Ribonucleic acid

RT-PCR – Real-Time PCR

SAD – Sinapyl alcohol dehydrogenase

SC-Leu – Synthetic complete leucine dropout medium

SDS-PAGE – Sodium dodecyl sulfate polyacrylamide gel electrophoresis

SGD – Strictosidine β-D-glucosidase

SLS – Secologanin synthase

STR – Strictosidine synthase

T3O – Tabersonine 3-oxygenase

T3R – Tabersonine 3-reductase

T6,7E – Tabersonine 6,7-epoxidase

T6,7R – Tabersonine 6,7-reductase

T16H – Tabersonine 16-hydroxylase

T19H – Tabersonine 19-hydroxylase

TEG – Tris-HCl/EDTA/glycerol buffer

TES – Tris-HCl/EDTA/sorbitol buffer

TDC – Tryptophan decarboxylase

TLC – Thin-layer chromatography

Tris – Tris(hydroxymethyl)aminomethane

TRV – Tobacco rattle virus

UPLC – Ultra-performance liquid chromatography

UDP – Uridine diphosphate

UGT – UDP-sugar glycosyltransferase

UV – Ultraviolet

WT – Wild-type

VIGS – Virus-induced gene silencing

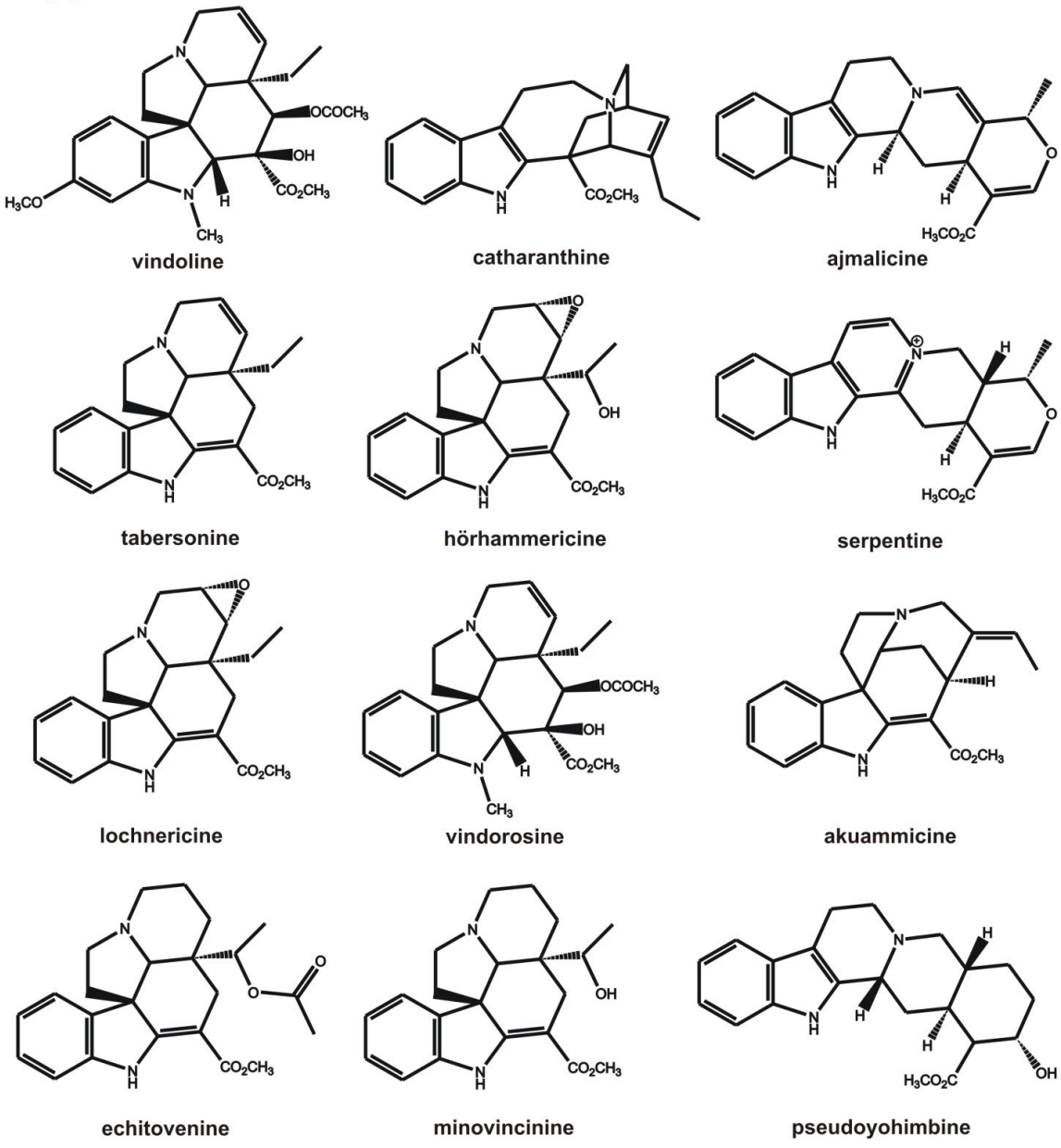
A. Introduction and Literature Review

A.1. *Catharanthus roseus* produces medicinally-active monoterpenoid indole alkaloids.

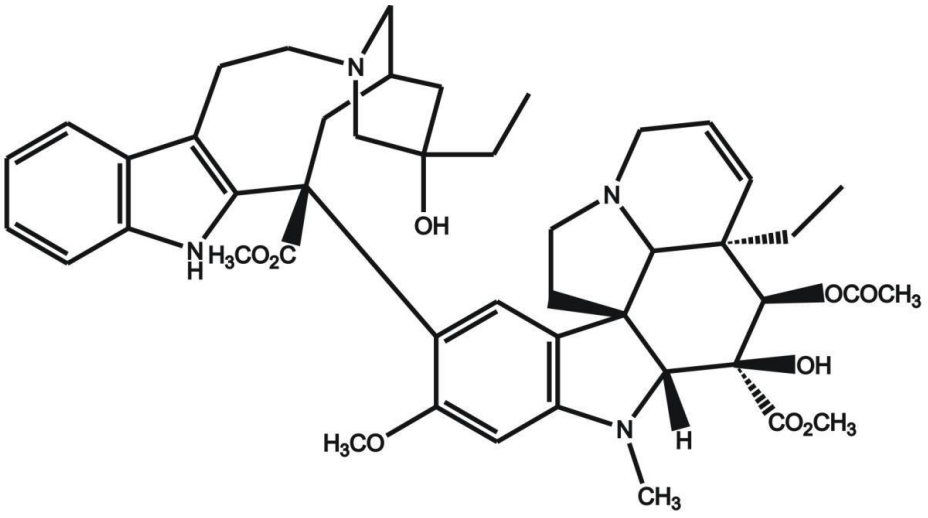
Plants synthesize a huge array of specialized metabolites, many believed to play roles in defense against pathogens, herbivory, or oxidative damage. Alkaloids are a broad class of nitrogen-containing secondary metabolites derived from amino acid precursors that display unique biological properties (De Luca et al., 2014). Monoterpenoid indole alkaloids (MIAs) are distributed across the Apocynaceae, Rubiaceae, Loganiaceae, Gentianaceae, and Nyssaceae families (Mann et al., 1994; Yamamoto et al., 2000; Asada et al., 2013).

Madagascar periwinkle (*Catharanthus roseus* (L.) G. Don), from the Apocynaceae family, contains over 100 MIAs (St-Pierre et al., 1999). Major MIA constituents (Figure 1A) of the plant include vindoline and catharanthine in the aboveground organs (Westkemper et al., 1980; Deus-Neumann et al., 1987), and ajmalicine, serpentine, tabersonine, lochnericine, and hörhammericine in roots (van der Heijden et al., 1989; Morgan and Shanks, 2000; Davioud et al., 1989; Shanks et al., 1998). Other MIAs include vindorosine, isolated as a trace compound that is found primarily in stems and also in roots (Moza and Trojánek, 1963; Cordell and Farnsworth, 1976), as well as akuammicine (Scott et al., 1971), minovincinine, echitovenine, and pseudoyohimbine (Cordell and Farnsworth, 1976). *C. roseus* is notable for the production of dimeric indole alkaloids vinblastine and vincristine (Figure 1B), which are formed via the enzymatic coupling of vindoline and catharanthine (Svoboda and Blake, 1975) and are isolated from the plant in very low yields of about 0.0005% (St-Pierre and De Luca,

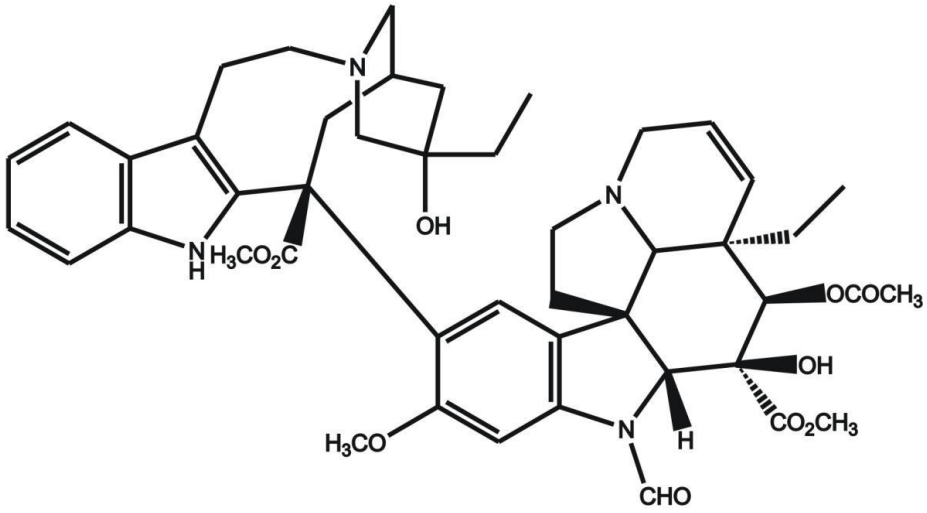
A


vindoline
catharanthine
ajmalicine
tabersonine
hörhammericine
serpentine
lochnericine
vindorosine
akuammicine
echitovenine
minovincinine
pseudoyohimbine

B



vinblastine



vincristine

Figure 1: Chemical structures of notable MIA constituents of *Catharanthus roseus*. (**A**) Major MIAs include catharanthine and vindoline in the aboveground organs, and ajmalicine, serpentine, tabersonine, hörhammericine, and lochnericine in roots; minor MIAs include vindorosine, echitovenine, minovincinine, akuammicine, and pseudoyohimbine. (**B**) *C. roseus* also accumulates very low levels of the anticancer dimers, vinblastine and vincristine, which are formed by the enzymatic coupling of catharanthine and vindoline.

1995). Many of the MIAs produced by *C. roseus* are of great interest due to their medicinal properties.

Lochnericine and hörhammericine are cytotoxic compounds (Cordell and Farnsworth, 1976) while ajmalicine and serpentine are anti-hypertensive agents (van der Heijden et al., 1989; Morgan and Shanks, 2000). Vinblastine, vincristine, and their semisynthetic derivates are potent chemotherapeutic agents employed against numerous cancers, including Hodgkin's disease, leukaemia, and lymphomas (Dewick, 2009). These dimeric alkaloids function by preventing tubulin polymerization, thereby interfering with cell structure, migration, and division (Jordan and Wilson, 2004; Gigant et al., 2005; Yang et al., 2010). These pharmaceutical properties make vinblastine and vincristine extremely valuable compounds in medicine but although total organic syntheses have been achieved, none have thus far proved economical enough to see commercial use, and these alkaloids are still extracted from plant material despite their low abundance (van der Heijden et al., 2004; Dewick, 2009).

Utilizing bioreactors as an alternative source of MIAs represents a viable method and many attempts have been made to produce these MIAs in undifferentiated plant tissue cultures with limited success. Although cell lines have been discovered that produce catharanthine (Kurz et al., 1980; Scott et al., 1980; Stöckijt and Soll, 1980), none appeared to consistently accumulate vindoline or vindoline-containing dimers (Berlin, 1984; Kurz and Constabel, 1985; Tyler et al., 1986; Eilert et al., 1987b; De Luca and Kurz, 1988; De Carolis et al., 1990). Shoots regenerated from cell cultures regained the ability to produce vindoline and vinblastine, indicating that the pathway is still present (Constabel et al., 1982). Subsequent studies have shown that the first two steps involved

in the conversion of tabersonine to vindoline have reduced activity in cell culture compared to young leaves of *C. roseus* (St-Pierre and De Luca, 1995), while the genes responsible for the last three steps in vindoline biosynthesis are not expressed at all in tissue culture (De Luca et al., 1987; Fahn et al., 1985b; St-Pierre and De Luca, 1995). Low vindoline production has been observed in some hairy root cultures established via *Agrobacterium rhizogenes* transformation (Bhadra et al., 1993; Palazón et al., 1998) and in cell suspension cultures established through *Agrobacterium tumefasciens*- and *A. rhizogenes*-mediated transformation of leaf discs (O'Keefe et al., 1997). This low vindoline production was eventually attributed to the low deacetylvindoline-4-*O*-acetylation activity of a root-specific minovincinine-19-*O*-acetyltransferase (Laflamme et al., 2001). These results suggested that achieving vinblastine production in plant cell cultures or via metabolic engineering in heterologous systems (yeast or model plants) requires further understanding of the biosynthetic pathway and the complex regulatory systems involved.

Numerous studies have been performed in the last several decades to investigate vinblastine biosynthesis. Much recent progress has been made in completing the biosynthetic pathway, with only a few steps yet to be identified.

A.2. The biosynthesis of secologanin from geraniol has been fully elucidated.

Most MIAs, including those of *C. roseus*, are synthesized by the assembly of a terpenoid (secologanin) and indole (tryptamine) moiety. The indole portion is derived from tryptophan while the terpenoid moiety is synthesized via the iridoid pathway from the monoterpene precursor geranyl pyrophosphate (GPP) (Meijer et al., 1993a). Tryptamine and secologanin are combined to form the central intermediate, strictosidine

(Battersby et al., 1969). Iridoid and MIA pathway intermediates have been traditionally determined via feeding of radiolabelled substrates and/or analysis of constituents in cell cultures, germinating seedlings, and developing seedlings (Battersby et al., 1966; Loew et al., 1966; Qureshi and Scott, 1968; Battersby et al., 1968; Loew and Arigoni, 1968; Battersby et al., 1969; Inouye et al., 1969; Battersby et al., 1970a; Scott et al., 1971; Uesato et al., 1984a; Uesato et al., 1986a; De Luca et al., 1986; Balsevich et al., 1986). The GPP involved in iridoid biosynthesis was initially suggested to originate from the cytosolic mevalonate pathway via feeding experiments (Goeggel and Arigoni, 1965; McCapra et al., 1965; Hall et al., 1966), though later experiments indicated that the plastid-localized 2-C-methyl-D-erythritol-4-phosphate (MEP) pathway contributes to the formation of the majority of iridoid precursors in cell and hairy root cultures (Contin et al., 1998; Hong et al., 2003) by providing geraniol via geraniol synthase (GES) (Simkin et al., 2013).

The iridoid pathway (Figure 2) begins with hydroxylation of geraniol by geraniol-10-hydroxylase (G10H), followed by further oxidation by 10-hydroxygeraniol oxidoreductase (10HGO) to form the dialdehyde, 10-oxogeranial. This substrate is converted by iridoid synthase (IRS) to iridodial and nepetalactol, the latter of which then undergoes a 3-step oxidation by 7-deoxyloganetic acid synthase (7DLS). Subsequent *O*-glucosylation is catalyzed by 7-deoxyloganetic acid glucosyltransferase (DLGT), after which the product undergoes hydroxylation by 7-deoxyloganic acid 7-hydroxylase (DL7H) and *O*-methylation by loganic acid *O*-methyltransferase (LAMT) to form loganin. Secologanin synthase (SLS) catalyzes the final ring cleavage to produce secologanin.

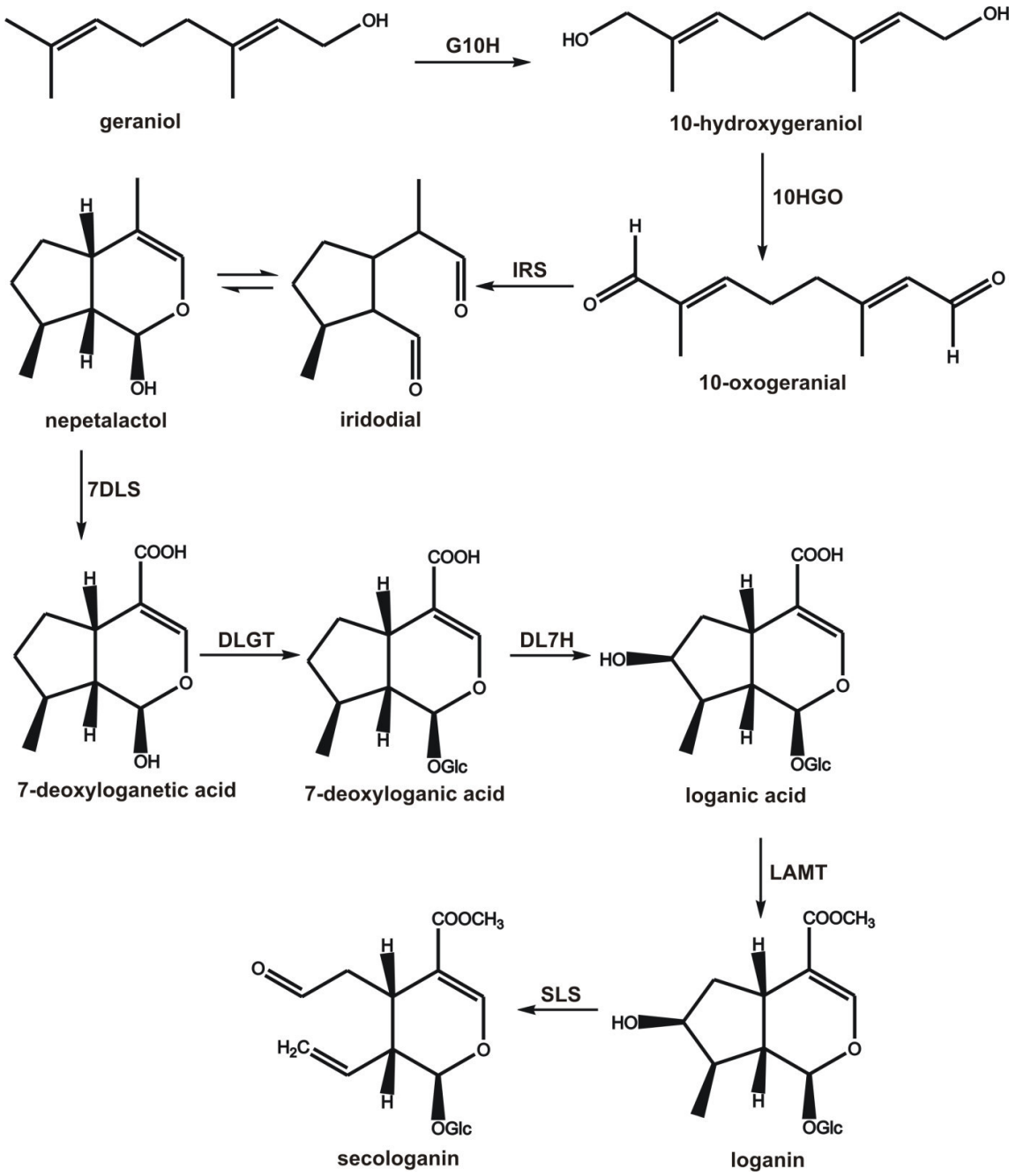


Figure 2: The conversion of geraniol to secologanin in *C. roseus* involves 8 steps that have been characterized at the biochemical and molecular levels. Enzymes: geraniol 10-hydroxylase (G10H), 10-hydroxygeraniol 10-oxidoreductase (10HGO), iridoid synthase (IRS), 7-deoxyloganetic acid synthase (7DLS), 7-deoxyloganetic acid glucosyltransferase (DLGT), 7-deoxyloganic acid 7-hydroxylase (DL7H), loganic acid O-methyltransferase (LAMT), secologanin synthase (SLS).

10-Hydroxylation was implicated as the first committed step in secologanin biosynthesis via feeding experiments (Battersby et al., 1970b; Escher et al., 1970). A cytochrome P450 (CYP)-dependent monooxygenase that exhibited 10-hydroxylase activity towards geraniol in the presence of NADPH was isolated in a microsomal preparation from *C. roseus* seedlings and partially purified (Meehan and Coscia, 1973; McFarlane et al., 1975; Madyastha et al., 1976). CYP76B6 was subsequently highly purified from *C. roseus* cell suspension cultures, characterized, and sequenced, and the corresponding gene was cloned using degenerate PCR primers based on the amino acid sequence (Meijer et al., 1993b; Collu et al., 2001; Collu et al., 2002). Expression of *CYP76B6* in yeast and transgenic *C. roseus* cell culture was used to confirm its G10H activity (Collu et al., 2001).

Further oxidation of 10-hydroxygeraniol to its dialdehyde, 10-oxogeranial, is the second step in secologanin biosynthesis. Crude protein extract from *Rauwolfia serpentina* cell cultures could perform this oxidation in the presence of $NADP^+$ (Stöckigt et al, 1981; Uesato et al., 1986b; Uesato et al., 1987). The corresponding protein was purified and characterized from *R. serpentina* cell cultures (Ikeda et al., 1991), and another $NADP^+$-dependent oxidoreductase with similar substrate specificity was also isolated from *Nepeta racemosa* (Hallahan et al., 1995). Although a *C. roseus* mRNA sequence for *10HGO* was present in the NCBI GenBank under the accession number AY352047, the report was unpublished (Mahroug et al., 2007) and the gene has only recently been cloned and characterized from *C. roseus* (Miettinen et al., 2014). Expression of the *C. roseus 10HGO* in *Escherichia coli* indicated that the protein converts 10-hydroxygeraniol to 10-oxogeranial in the presence of NAD^+ and, like the *R. serpentina* 10HGO, displays

activity towards primary alcohols (Miettinen et al., 2014).

Cyclization was initially believed to be catalyzed by a monoterpene cyclase, a class of enzymes that accept polyprenyl diphosphates to form various ring structures (Degenhardt et al., 2009; Chen et al., 2011). However, the direct precursor to iridoids was shown to be 10-oxogeranial as opposed to geranyl diphosphate (Inouye, 1978; Uesato et al., 1984a and 1984b). Crude protein extracts from *R. serpentina* cell cultures could convert 10-hydroxygeraniol into iridodial in the presence of NADH/NADPH and NAD^+/$NADP^+$ (Uesato et al., 1986b) and a corresponding cyclase that converted 10-oxogeranial to iridodial was partially purified from *R. serpentina* cell cultures (Uesato et al., 1987). The *IRS* gene, which exhibits high similarity to progesterone-5β-reductase from *Digitalis purpurea*, was recently identified from a *C. roseus* transcriptome database based on co-expression with *G10H* (Geu-Flores et al., 2012; Gavidia et al., 2007). When characterized by functional expression in *E. coli*, IRS did not metabolize progesterone but did convert 10-oxogeranial into an equilibrium mixture of nepetalactol and iridodial in an NADH/NADPH-dependent reaction (Geu-Flores et al., 2012). The *in vivo* function of IRS was supported by *in planta* virus-induced gene silencing (VIGS) experiments (Geu-Flores et al., 2012).

Oxidation of the cyclized product to form iridotrial was originally believed to be the next step (Uesato et al., 1986a), though characterization of subsequent steps in the pathway identified 7-deoxyloganetic acid as the next known intermediate (Asada et al., 2013; Salim et al., 2013; Murata et al., 2008). Three oxidations are necessary to convert iridodial/nepetalactol to 7-deoxyloganetic acid, and a bioinformatics search for *C. roseus* CYP candidate genes involved in secologanin biosynthesis recently led to the

identification of *CYP76A26* (Salim et al., 2013). VIGS experiments with *C. roseus* plants supported its role in the secologanin pathway and expression in yeast confirmed its function as 7DLS, as it converted nepetalactol via iridotrial into 7-deoxyloganetic acid in the presence of NADPH (Salim et al., 2014). Yeast expressing *7DLS* could also convert both iridodial, which isomerizes to nepetalactol, and iridotrial to 7-deoxyloganetic acid (Miettinen et al., 2014).

Conversion from 7-deoxyloganetic acid to 7-deoxyloganic acid requires glucosylation, a function that is typically performed by plant secondary product glycosyltransferases (PSPGs) that contain a characteristic 44-amino acid PSPG box motif (Vogt and Jones, 2000). The *UDP-SUGAR GLYCOSYLTRANSFERASE2* (*UGT2*) gene identified in *Gardenia jasminoides* by homology-based cloning using the conserved PSPG box sequence and functional screening encodes an iridoid glucosyltransferase that accepts 7-deoxyloganetin as substrate but not 7-deoxyloganetic acid (Nagatoshi et al., 2011). The same approach identified three UGTs in *C. roseus*, which were cloned from cell suspension cultures and young leaves (Asada et al., 2013). Functional characterization of the recombinant proteins purified from *E. coli* indicated that UGT8 displayed a high catalytic efficiency with 7-deoxyloganetic acid as the substrate, and its role as DLGT was also confirmed using VIGS (Asada et al., 2013).

Following glucosylation to 7-deoxyloganic acid, it was initially unclear whether or not 7-hydroxylation preceded *O*-methylation (Figure 3). Feeding experiments originally implicated 7-deoxyloganin as the precursor to loganin and subsequently suggested that methylation occurred before hydroxylation (Inouye et al., 1969; Battersby et al., 1970a). Cell suspension cultures of *Lonicera japonica* could convert both loganin

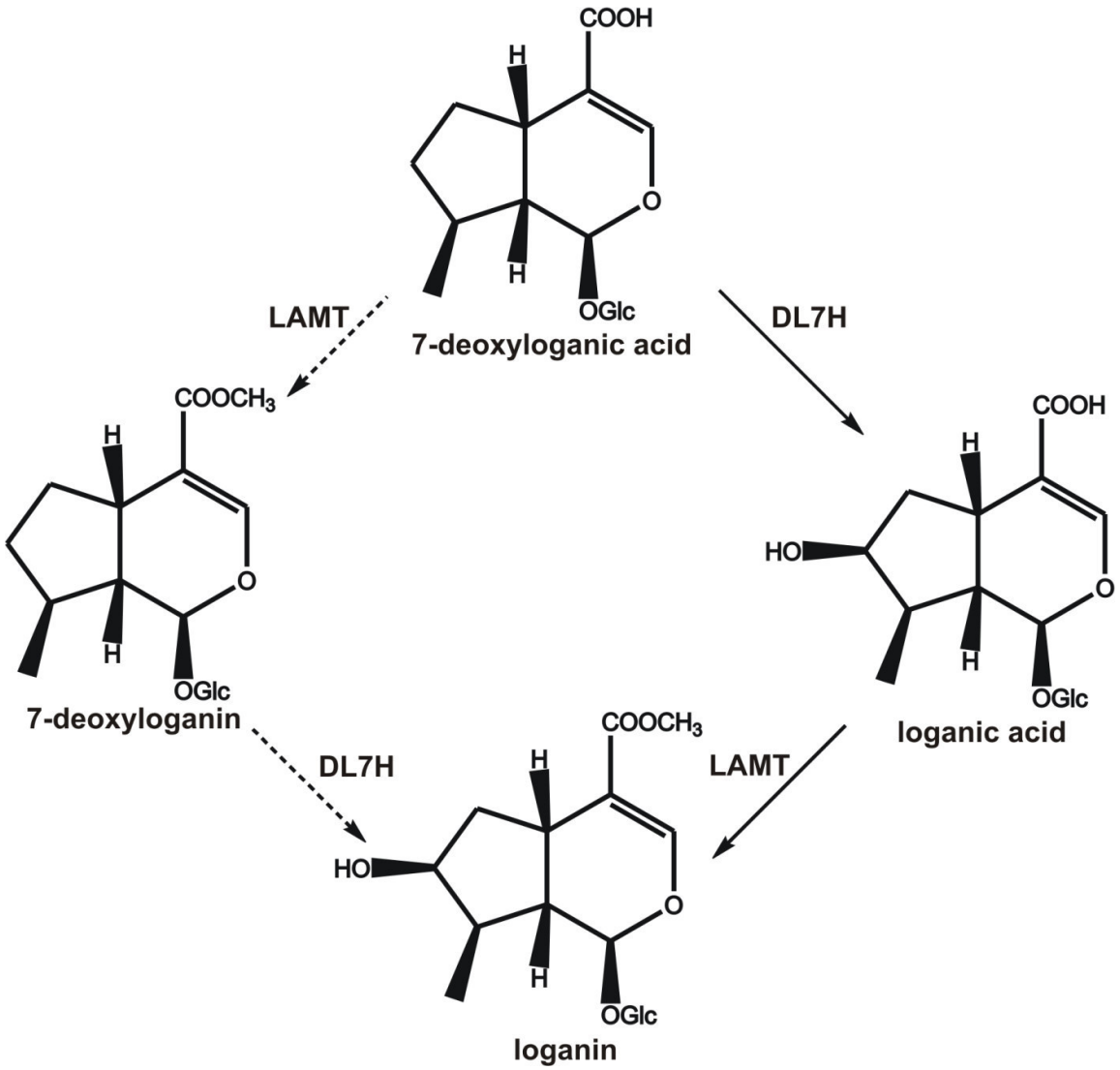


Figure 3: Proposed order of loganin biosynthesis in *C. roseus*. Solid arrows indicate the preferred pathway. Enzymes: 7-deoxyloganic acid 7-hydroxylase (DL7H), loganic acid *O*-methyltransferase (LAMT).

and 7-deoxyloganin into secologanin (Yamamoto et al., 1999) and demonstration of 7-hydroxylase activity in microsomes from *C. roseus* cell cultures resulted in the NADPH-dependent conversion of 7-deoxyloganin to loganin (Irmler et al., 2000). Similarly, the CYP enzyme DL7H was first purified from *L. japonica* cell culture and exhibited higher activity towards 7-deoxyloganin compared to 7-deoxyloganic acid (Katano et al., 2001). The *DL7H* gene was recently identified in *C. roseus* utilizing a bioinformatics approach, which produced three CYP candidates for the secologanin pathway (Salim et al., 2013). *CYP76A26* was later identified as 7DLS (Salim et al., 2014), but VIGS of *CYP72A224* resulted in a decrease in secologanin and accumulation of 7-deoxyloganic acid instead of 7-deoxyloganin in silenced *C. roseus* plants (Salim et al., 2013). Expression of the gene in yeast and functional characterization supported its identity as the *C. roseus* DL7H; the enzyme converted 7-deoxyloganic acid to loganic acid in the presence of NADPH but would not accept 7-deoxyloganin as substrate (Salim et al., 2013). This study, combined with characterization of the *O*-methyltransferase, indicated the preferred order of reactions in *C. roseus*.

A methyltransferase that converted loganic acid into loganin in the presence of *S*-adenosyl-L-methionine (AdoMet) and dithiothreitol was identified in cell-free extracts of *C. roseus* cell suspension (Madyastha et al., 1971). Likewise, the partially purified LAMT methylated loganic acid but not 7-deoxyloganic acid (Madyastha et al., 1973), and the gene was later cloned from *C. roseus* leaf epidermis cDNA libraries (Murata et al., 2008). Characterization of recombinant LAMT from both *C. roseus* and *L. japonica* confirmed its stringent substrate specificity, as neither enzyme would accept 7-deoxyloganic acid as a substrate, indicating that hydroxylation by DL7H precedes

methylation by LAMT (Murata et al., 2008; Edge, 2013). LAMT is present in the cytosol as a homodimer (Murata et al., 2008; Guirimand et al., 2011a).

The final step in the iridoid pathway involves oxidative cleavage of the cyclopentane ring of loganin to form secologanin. The ability to convert loganin to secologanin was observed in cell suspension cultures of *Lonicera* (Tanahashi et al., 1984), and SLS was identified as a CYP-dependent monooxygenase that would only accept loganin as a substrate in *L. japonica* cell culture (Yamamoto et al., 2000). *CYP72A1* was identified in cDNA libraries of *C. roseus* cell suspension cultures as a sequence with similar expression profiles to *G10H*, but the encoded protein did not have the capability to convert geraniol to 10-hydroxygeraniol (Vetter et al., 1992). While expression of this gene in transformed tobacco and *Arabidopsis thaliana* also failed to identify its function (Mangold et al., 1994), CYP72A1 was later identified as SLS when recombinant protein produced from *E. coli* converted loganin into secologanin in the presence of NADPH (Irmler et al., 2000). Whereas secologanin contributes the terpenoid moiety of the MIAs, the indole moiety is derived from tryptamine (Cordell, 1974; De Luca and Kurz, 1988).

A.3. Secologanin and tryptamine are combined into the central MIA intermediate, strictosidine.

Tryptamine and its derivatives are widespread in plants and act as precursors to a variety of plant hormones and secondary metabolites (Baxter and Slaytor, 1972; Facchini et al., 2000). Tryptamine was expected to be produced from the decarboxylation of tryptophan, and a protein exhibiting pyridoxal phosphate-dependent tryptophan decarboxylase (TDC) activity was detected in cell-free extracts of walnuts (*Juglans regia*

L.) (Grosse and Klapheck, 1979) and partially purified from *Phalaris tuberosa* seedlings (Baxter and Slaytor, 1972) and tomato shoots (Gibson et al., 1972). TDC was later purified and characterized from cell suspension cultures of *C. roseus* (Noé et al., 1984) and the gene was subsequently identified in a *C. roseus* seedling cDNA library and functionally expressed in *E. coli* (De Luca et al., 1989).

Condensation of tryptamine with secologanin (Figure 4) produces strictosidine (Battersby et al., 1969), a central precursor to thousands of MIAs found in various plants (Guirimand et al., 2010). Strictosidine can be converted to other MIAs by cell-free *C. roseus* extracts in the presence of NADH/NADPH (Stöckigt and Zenk, 1977a, 1977b). Strictosidine synthase (STR) was identified, purified, and characterized from soluble protein extracts of *C. roseus* cell suspension cultures (Treimer and Zenk, 1979; Mizukami et al., 1979). The *STR* gene was first cloned from *R. serpentina* (Kutchan et al., 1988) and its partial sequence was subsequently identified in *C. roseus* (McKnight et al., 1990), followed by the full-length mRNA sequence (Pasquali et al., 1992). Although strictosidine is eventually converted to catharanthine and tabersonine via potential intermediates such as preakuammicine, stemmadenine, and dehydrosecodine (Dewick, 2009; El-Sayed et al., 2004; De Luca et al., 2012), only the deglucosylation step in this segment of the pathway has been identified and characterized (Figure 5).

Two β-glucosidases specific for strictosidine were identified, partially purified, and characterized in *C. roseus* cell cultures (Hemscheidt and Zenk, 1980). Localization studies with the β-glucosidase in *C. roseus* cell cultures also suggested the presence of two forms of the protein (Stevens at al., 1993). Strictosidine β-D-glucosidase (SGD) was later purified to homogeneity from *C. roseus* cell cultures; characterization indicated that

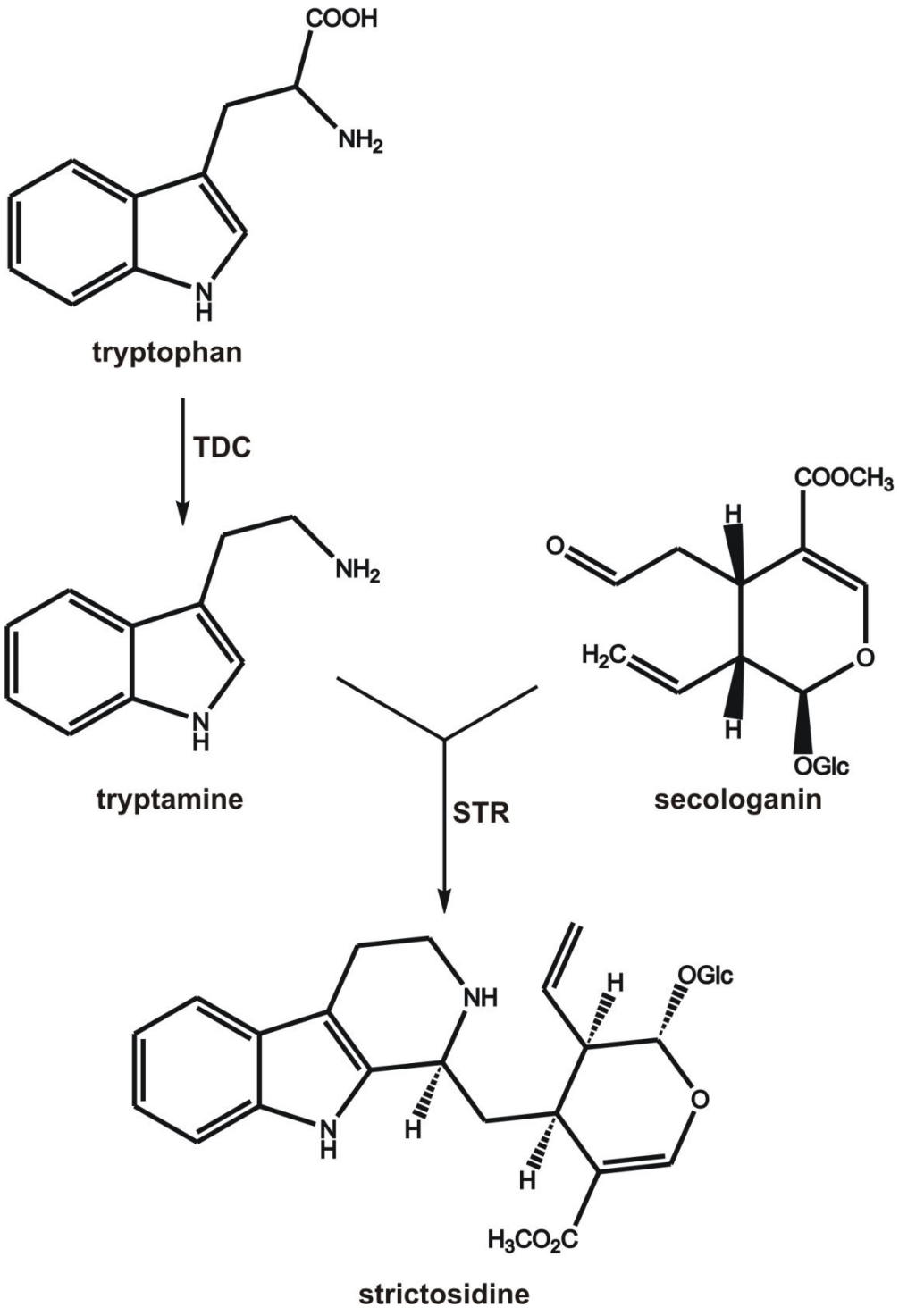


Figure 4: Biosynthesis of strictosidine, a central MIA intermediate in *C. roseus*, from tryptamine and secologanin. Enzymes: tryptophan decarboxylase (TDC), strictosidine synthase (STR).

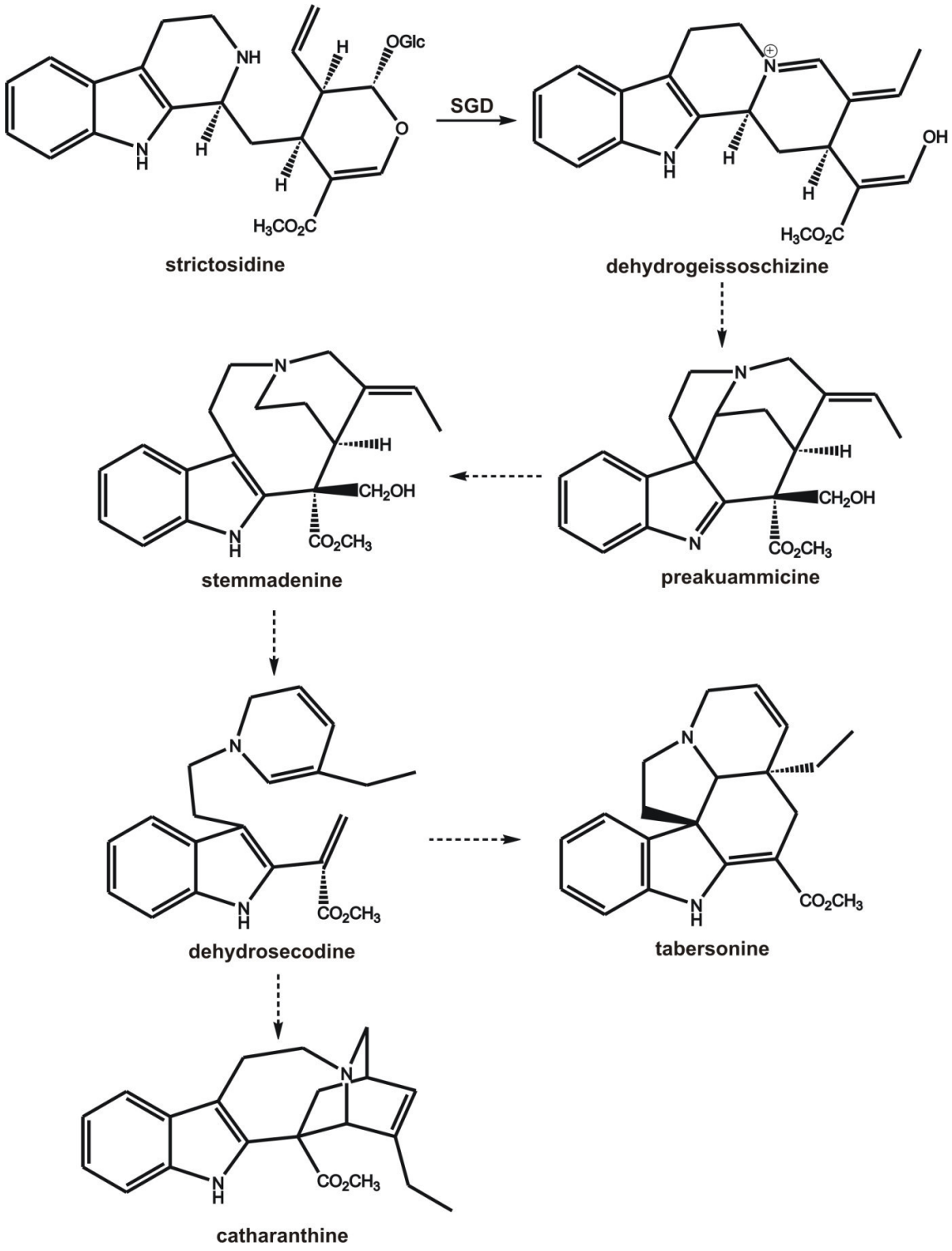


Figure 5: Proposed biosynthesis of tabersonine and catharanthine from strictosidine in *C. roseus*. Dashed arrows represent uncharacterized steps. Enzymes: strictosidine β-D-glucosidase (SGD).

the native protein is present in aggregates of multiple 63-kDa monomers (Luijendijk et al., 1998), which was supported by an additional study that observed the aggregation of 4, 8, or 12 monomers (Geerlings et al., 2000). Although amino acid sequencing of SGD was unsuccessful, degenerate PCR primers designed based on conserved regions of plant *β*-glucosidases resulted in the identification of a cDNA clone (Geerlings et al., 2000). Functional expression of the gene in yeast confirmed its function as SGD, as it exhibited high substrate specificity for strictosidine. SGD forms dehydrogeissoschizine, a reactive dialdehyde that is converted via molecular rearrangements into various corynanthe, iboga, and aspidosperma type alkaloids (De Luca and Laflamme, 2001). Feeding studies have suggested that stemmadenine is a precursor to both catharanthine and tabersonine (El-Sayed et al., 2004), but these pathways remain uncharacterized.

A.4. The pathway from tabersonine to vindoline has been fully characterized.

Although the enzymatic steps between dehydrogeissoschizine and tabersonine have not yet been identified, tabersonine is converted into vindoline in seven enzymatic steps, which have all been characterized (Figure 6). Tabersonine 16-hydroxylase (T16H) and 16-hydroxytabersonine 16-*O*-methyltransferase (16OMT) introduce the 16-methoxy group, after which tabersonine 3-oxygenase (T3O) and tabersonine 3-reductase (T3R) introduce the 3-hydroxyl and reduce the 2,3-double bond. The product is then *N*-methylated by 16-methoxy-2,3-dihydro-3-hydroxytabersonine *N*-methyltransferase (NMT) to desacetoxyvindoline, which is converted by desacetoxyvindoline-4-hydroxylase (D4H) to deacetylvindoline. Deacetylvindoline 4-*O*-acetyltransferase (DAT) catalyzes the final step to produce vindoline. A minor pathway also exists in *C. roseus*, in which tabersonine does not undergo the T16H- and 16OMT-mediated 16-methoxylation

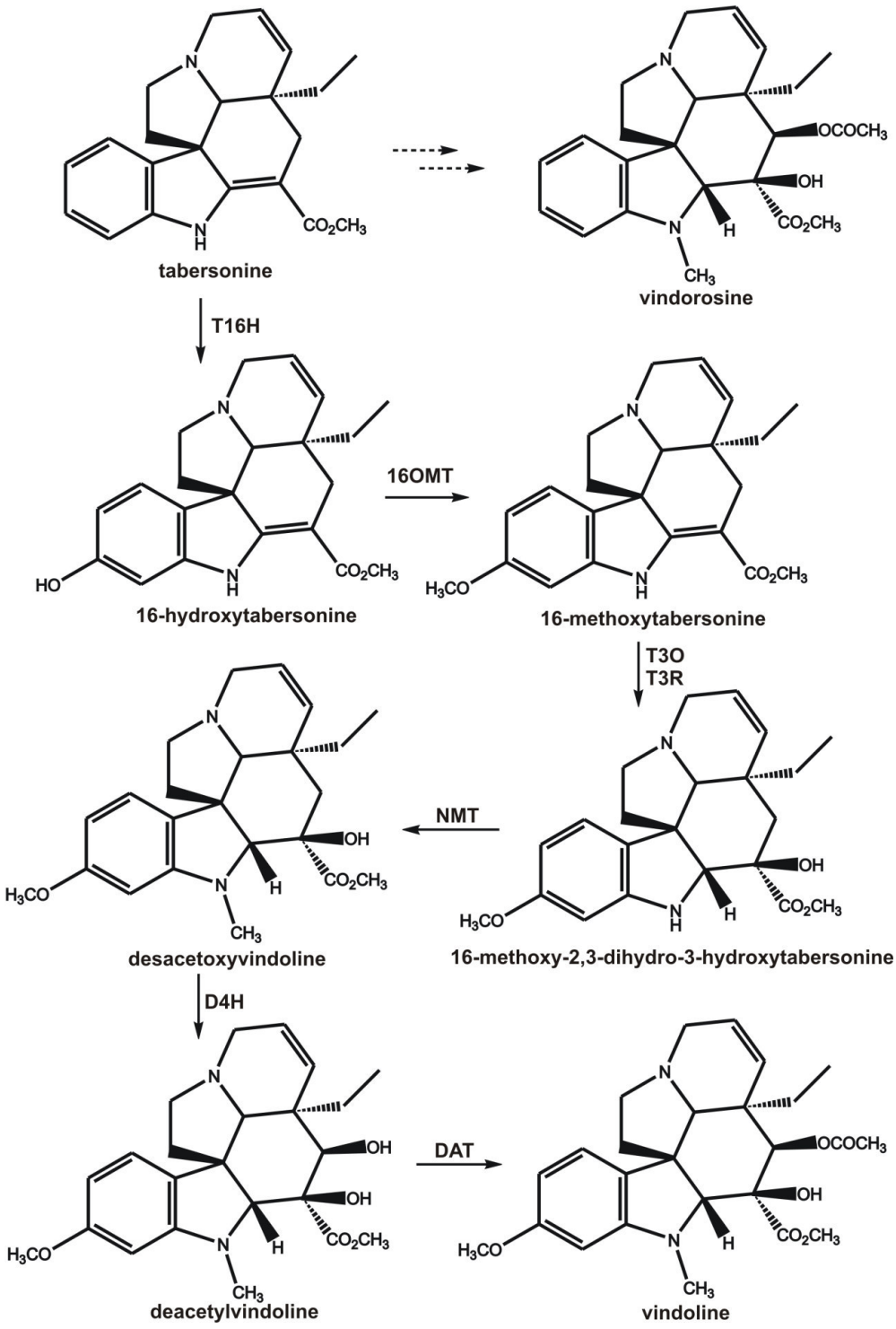


Figure 6: Conversion of tabersonine to vindoline and vindorosine in *C. roseus*. Enzymes: tabersonine 16-hydroxylase (T16H), 16-methoxytabersonine 16-*O*-methyltransferase (16OMT), tabersonine 3-oxygenase (T3O), tabersonine 3-reductase (T3R), 16-methoxytabersonine-2,3-dihydro-3hydroxytabersonine *N*-methyltransferase (NMT), desacetoxyvindoline 4-hydroxylase (D4H), deacetylvindoline 4-*O*-acetyltransferase (DAT).

but presumably undergoes the T3O-, T3R-, NMT-, D4H-, and DAT-catalyzed reactions to form vindorosine.

16-Hydroxylase activity was detected in protein extracts from young leaves of *C. roseus* in the presence of NADPH (St-Pierre and De Luca, 1995). This activity is induced by light in *C. roseus* cell cultures, which was used to identify *CYP71D12* from cDNA libraries generated from light-induced cell cultures (Schröder et al., 1999). Functional expression in *E. coli* and production of 16-methoxytabersonine in a coupled reaction with 16OMT confirmed its identity as T16H. However, later studies revealed the presence of two distinct *T16H* genes in *C. roseus*: the previously-characterized *CYP71D12* (*T16H1*) is restricted to flowers and undifferentiated cells, while the newly-identified *CYP71D351* (*T16H2*) is specific to leaves (Besseau et al., 2013). T16H1 functions as a monomer (Guirimand et al., 2011b).

AdoMet-dependent 16-*O*-methylation activity was identified in protein extracts from *C. roseus* leaves (Fahn et al., 1985a). Inclusion of the protein with T16H in a coupled assay resulted in the production of 16-methoxytabersonine from tabersonine (Schröder et al., 1999). Subsequent efforts to clone *16OMT* from cell cultures failed (Cacace et al., 2003; Schröder et al., 2002; Schröder et al., 2004) until carborundum abrasion of *C. roseus* leaves was used to isolate epidermis-enriched proteins and RNA, from which the protein was purified and the gene cloned and characterized (Levac et al., 2008). 16OMT functions as a homodimer (Levac et al., 2008; Guirimand et al., 2011b).

Conversion of 16-methoxytabersonine to the 2,3-dihydro-3-hydroxy derivative was very recently revealed to be catalyzed by a pair of enzymes. Bioinformatics approaches identified two highly-expressed epidermal genes, *CYP71D1* and an alcohol

dehydrogenase-like (*ADHL1*) gene (Kellner et al., 2015; Qu et al., 2015). Silencing of *CYP71D1* in *C. roseus* plants resulted in decreases in both vindoline and vindorosine levels and an increase in the level of 16-methoxytabersonine (Kellner et al., 2015). Identified as *T3O*, this gene was subsequently cloned into yeast. Biotransformation assays with 16-methoxytabersonine formed an unexpected product structurally similar to vincamine, leading Kellner et al. (2015) to propose that T3O generates a 2,3-epoxide that spontaneously opens to an imine alcohol which, in the absence of the next enzyme in the pathway, undergoes a structural rearrangement (Figure 7A). However, this proposal was not supported by the concurrent characterization of *T3O* and *ADHL1* performed by Qu et al. (2015). Qu et al. observed similar results in VIGS experiments, in which silencing of *T3O* reduced vindoline and increased 16-methoxytabersonine levels while silencing of *ADHL1* reduced levels of desacetoxyvindoline and resulted in accumulation of isomeric MIAs on the leaf surface (m/z = 383). Rather than the vincamine-type structure observed by Kellner et al., analysis of these MIAs via NMR and sodium borohydride reduction identified them as 2,3-epoxide derivatives of 16-methoxytabersonine. Additionally, while yeast expressing *T3O* converted 16-methoxytabersonine to the same epoxide derivatives, these compounds were accepted as substrates by recombinant ADHL1 with only low levels of conversion to 2,3-dihydro-3-hydroxylated products (Qu et al., 2015). Instead, coupled assays using both recombinant microsomal T3O and purified recombinant ADHL1 converted both tabersonine and 16-methoxytabersonine into the 2,3-dihydro-3-hydroxylated product in the presence of NADPH with reduced accumulation of the 2,3-epoxide derivatives, suggesting the involvement of an unstable intermediate that spontaneously forms the epoxide (Figure 7B) in the absence of ADHL1, renamed T3R.

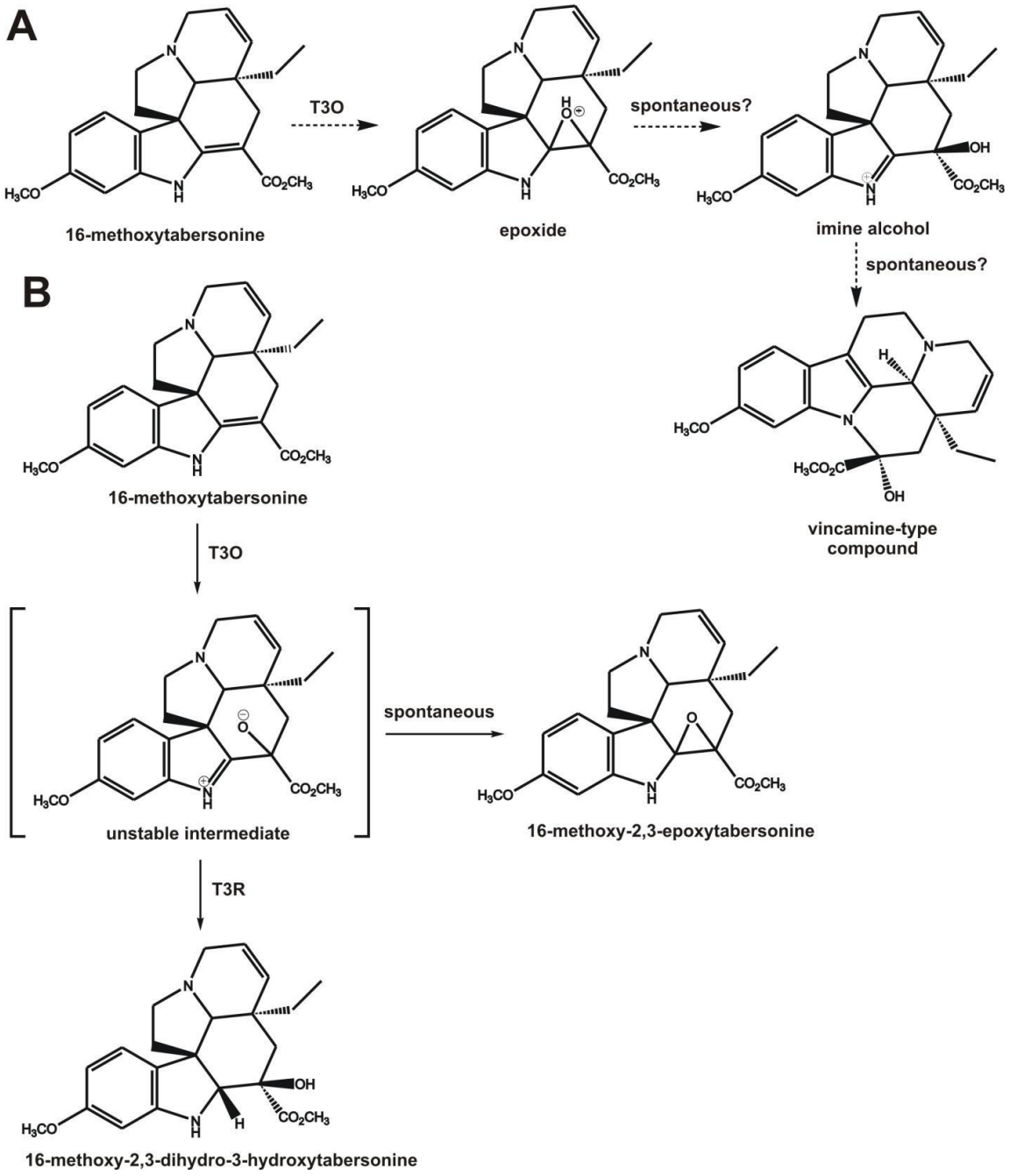


Figure 7: Vindoline biosynthesis in *C. roseus* involves two enzymes for addition of the 3-hydroxyl and reduction of the 2,3-double bond. (**A**) Kellner et al. (2015) proposed that T3O produces a 2,3-epoxide which spontaneously forms an imine alcohol and that in the absence of the second enzyme, a structural rearrangement occurs. (**B**) Characterization of both T3O and T3R by Qu et al. (2015) indicated that T3O instead produces an unstable intermediate that spontaneously forms a 2,3-epoxide in the absence of T3R. Enzymes: tabersonine 3-oxygenase (T3O), tabersonine 3-reductase (T3R).

T3O and T3R produce 16-methoxy-2,3-dihydro-3-hydroxytabersonine, the next intermediate in the pathway, which undergoes *N*-methylation in the next step.

N-Methylation of 2,3-dihydro-3-hydroxytabersonine in the presence of AdoMet was first identified in cell-free extracts of *C. roseus* leaves (De Luca et al., 1987). Since NMT activity was associated with the thylakoid membrane (De Luca et al., 1987), candidate cDNAs were identified based on sequence similarity to plastidial γ-tocopherol methyltransferases (Liscombe et al., 2010). Functional expression of candidate genes in *E. coli* identified the NMT involved in the AdoMet-dependent *N*-methylation of 2,3-dihydro-3-hydroxytabersonine or 16-methoxy-2,3-dihydro-3-hydroxytabersonine that lead to the eventual formation of vindorosine and vindoline, respectively (Liscombe et al., 2010).

A 2-oxoglutarate-dependent dioxygenase that exhibited 4-hydroxylase activity was partially purified from young leaves of *C. roseus* (De Carolis et al., 1990). The protein exhibited a strict specificity for position 4 of several alkaloid substrates; additionally, hydroxylation did not occur in the presence of the 2,3-double bond or the absence of *N*-methylation. D4H was purified to homogeneity from young leaves of *C. roseus* and characterized soon after (De Carolis and De Luca, 1993). A degenerate oligonucleotide probe based on the peptide sequence was used to screen a *C. roseus* cDNA library and identify the *D4H* gene, which was functionally expressed in *E. coli* and characterized (Vazquez-Flota et al., 1997). D4H functions as a monomer (Guirimand et al., 2011b).

The final step in vindoline biosynthesis involves 4-*O*-acetylation of deacetylvindoline. Protein extracts from *C. roseus* leaves exhibited acetyl-CoA-

dependent *O*-acetyltransferase activity and converted deacetylvindoline to vindoline (De Luca et al., 1985; Fahn et al., 1985b). DAT protein was purified from *C. roseus* leaves and characterized (Power et al., 1990) and the gene was cloned from a cDNA library using degenerate primers based on the peptide sequence (St-Pierre et al., 1998). DAT functions as a monomer (St-Pierre et al., 1998; Guirimand et al., 2011b).

In order to produce the dimeric anticancer MIAs, catharanthine and vindoline must be coupled (Figure 8) in a reaction that forms 3',4'-anhydrovinblastine (Verpoorte et al., 1997; Dewick, 2009). Crude protein extracts from *C. roseus* suspension cultures could couple catharanthine and vindoline into dimers with up to 25% conversion into 3',4'-anhydrovinblastine (Misawa et al., 1988). Partial purification of the extracts indicated that the coupling activity was closely associated with peroxidase activity (Endo et al., 1988). Similarly, partially-purified protein extracts from *C. roseus* leaves produced dimers from catharanthine and vindoline in the presence of H_2O_2, an activity that was attributed to a vacuolar basic peroxidase (Sottomayor et al., 1996). The peroxidase was subsequently purified to homogeneity from *C. roseus* leaves and characterized (Sottomayor et al., 1998). More recently, degenerate primers based on the peptide sequence of the purified protein were used to screen a cDNA library and identify the *C. roseus* vacuolar class III peroxidase (*PRX1*) gene, which coded for an identical amino acid sequence as the purified protein, although no functional recombinant expression was performed (Costa et al., 2008). 3',4'-Anhydrovinblastine is converted to vinblastine by protein extracts of both *C. roseus* leaves (Kutney et al., 1982; Baxter et al., 1979) and suspension cultures (McLauchlan et al, 1983; Endo et al., 1988), but no enzymes have been identified in these reactions.

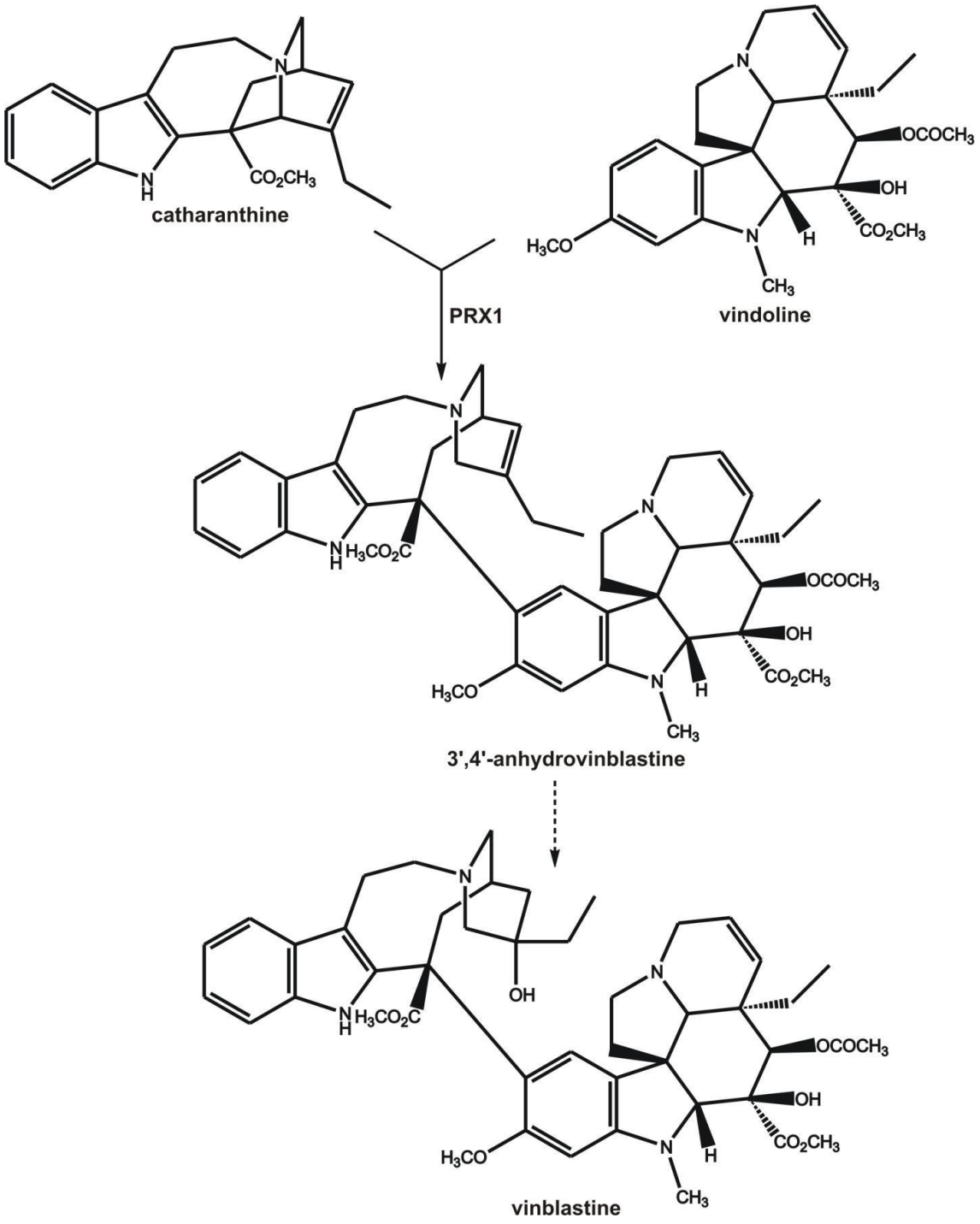


Figure 8: The coupling of catharanthine and vindoline to form the anticancer dimeric alkaloid, vinblastine, may be mediated by a peroxidase in *C. roseus*. Enzymes: vacuolar class III peroxidase (PRX1).

A.5. The root-specific pathway of *C. roseus* produces tabersonine derivatives.

Vindoline and the dimeric indole alkaloids are produced only in the aerial organs of *C. roseus*. Roots of the plant accumulate different MIAs, including 19-*O*-acetylhörhammericine, echitovenine, and lochnericine, which are derived from tabersonine via root-specific enzymatic steps.

In *C. roseus* roots, tabersonine can undergo 19-hydroxylation to 19-hydroxytabersonine (Morgan and Shanks, 1999) or 6,7-epoxidation to lochnericine (Kutney et al., 1980; Furuya et al., 1992; Shanks et al., 1998). Both 19-hydroxytabersonine and lochnericine are potential precursors in the formation of 19-*O*-acetylhörhammericine (Figure 9), while 19-hydroxytabersonine is also one of the potential precursors to echitovenine (O'Connor and Maresh, 2006; Laflamme et al., 2001; Giddings et al., 2011). The root pathway has not been well-characterized, as only tabersonine 19-hydroxylase (T19H) and minovincinine-19-hydroxy-*O*-acetyltransferase (MAT) have been characterized to date; tabersonine 6,7-epoxidase (T6,7E) activity has been demonstrated in hairy roots, while tabersonine 6,7-reductase activity has merely been proposed as a route to echitovenine (Giddings et al., 2011). Inhibitor studies with hairy root cultures suggested the involvement of two separate CYP-dependent monooxygenases in the formation of hörhammericine and lochnericine (Morgan and Shanks, 1999), which were later identified as T19H and T6,7E.

Transcriptome sequencing and coexpression analysis led to the identification of *CYP71BJ1*; yeast expressing the gene converted tabersonine into a product identified as 19-hydroxytabersonine in the presence of NADPH (Giddings et al., 2011). This T19H also accepted lochnericine as a substrate, which it converted into a product with the same

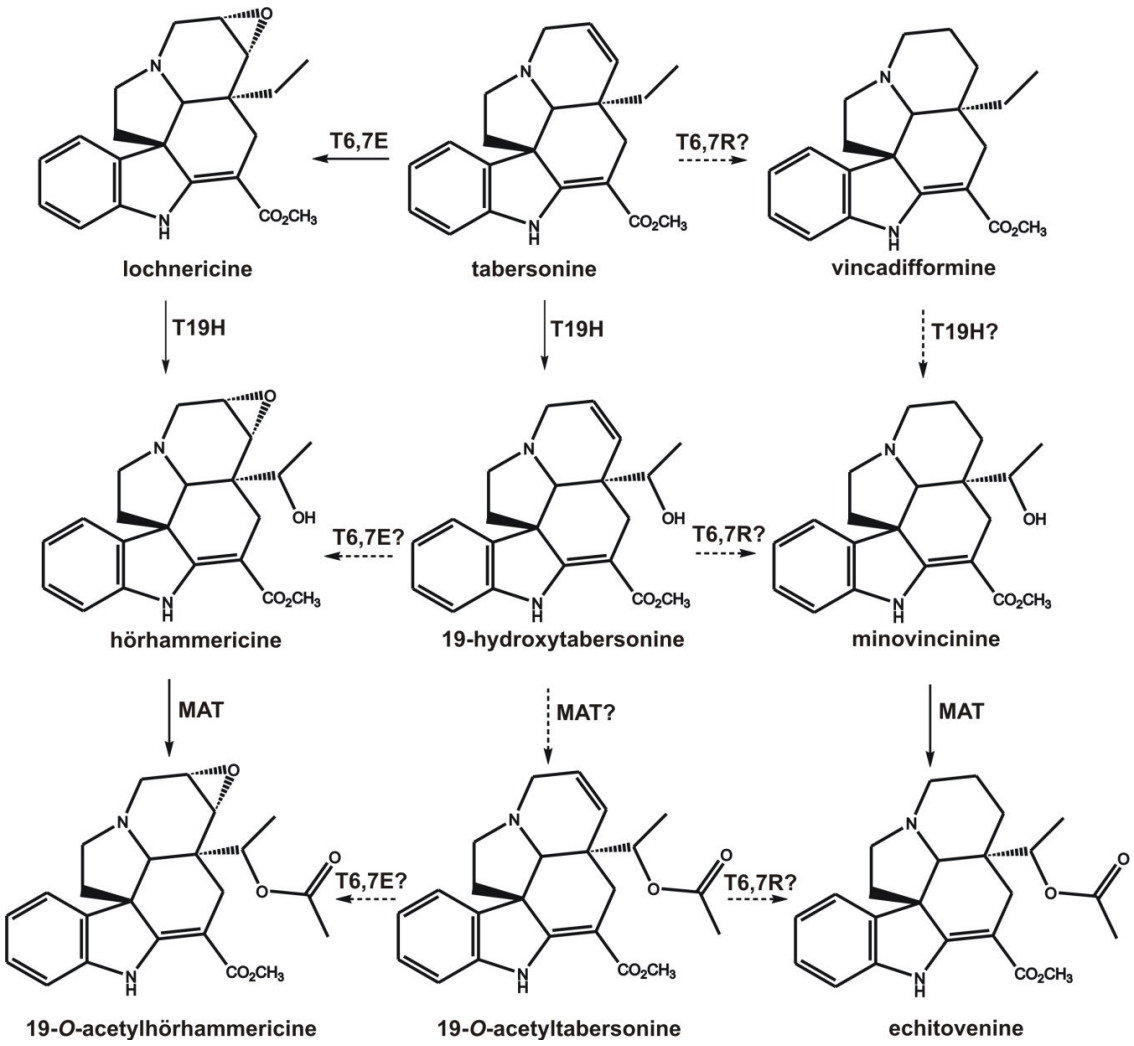


Figure 9: Proposed biosynthetic grid for the formation of MIAs found in the roots of *C. roseus*. Dashed arrows represent uncharacterized reactions. Enzymes: tabersonine 19-hydroxylase (T19H), tabersonine 6,7-epoxidase (T6,7E), tabersonine 6,7-reductase (T6,7R), and minovincinine 19-*O*-acetyltransferase (MAT). Adapted from Giddings et al. (2011).

mass and UV spectrum as hörhammericine.

Another CYP-dependent monooxygenase in crude protein extracts from hairy roots possessed 6,7-epoxidase activity that converted tabersonine into lochnericine in the presence of NADPH, although no other substrates were tested. (Rodriguez et al., 2003). To date, the gene responsible for this reaction has not been identified or characterized. Since T19H will hydroxylate both tabersonine and lochnericine (Giddings et al., 2011), 6,7-epoxidation of 19-hydroxytabersonine may be an alternative path to hörhammericine. Hörhammericine is a known precursor to 19-*O*-acetylhörhammericine, but 19-hydroxytabersonine has not been tested as a substrate for MAT.

When *DAT* was identified and characterized, some deacetylvindoline 4-*O*-acetylation activity by a root-specific DAT homolog was also observed, though this protein primarily produced the 19-*O*-acetylated tabersonine derivates found in roots (St-Pierre et al., 1998; Cordell and Farnsworth, 1976). Cloning and functional expression of the homolog identified it as MAT, which converted hörhammericine to 19-*O*-acetylhörhammericine and also produced echitovenine from crude root extracts, with minovincinine as the likely substrate (Laflamme et al., 2001). Transformation of hairy roots with DAT inhibits *O*-acetylation by MAT, reducing levels of 19-*O*-acetylhörhammericine and increasing levels of hörhammericine (Magnotta et al., 2007).

The difference between the aerial- and root-specific pathways highlights the compartmentalization of MIA biosynthesis in *C. roseus*. Numerous studies have investigated the localization of the vinblastine pathway, which has revealed a complex system of both cellular and subcellular compartmentalization.

A.6. Vinblastine biosynthesis takes place across multiple cell types.

Enzymes involved in MIA biosynthesis localize to multiple cell types and organelles, which neccessitates both intra- and inter-cellular trafficking of intermediates. Although transport of MIAs in *C. roseus* has not been well-studied, localization of enzymes and intermediates has been comprehensively investigated.

The MEP pathway, which provides most of the geraniol for secologanin biosynthesis, is localized to the plastids (El-Sayed and Verpoorte, 2007; Simkin et al., 2013) of internal phloem-associated parenchyma (IPAP) cells (Burlat et al., 2004; Mahroug et al., 2007; Simkin et al., 2013). Similarly, the early steps of the iridoid pathway are also expressed in IPAP cells. This includes *G10H*, *IRS*, and *DLGT* (Burlat et al., 2004; Mahroug et al., 2007; Geu-Flores et al., 2012; Asada et al., 2013), although additional localization studies have also suggested that G10H may be active in the leaf epidermis (Murata and De Luca, 2005). Studies with *7DLS* and *DL7H* suggest that they are expressed within the leaf (Salim et al., 2013; Salim et al., 2014), while the later steps in secologanin biosynthesis, including *LAMT* and *SLS*, are known to be expressed in the leaf epidermis (Irmler et al., 2000; Murata and De Luca, 2005; Murata et al., 2008; Guirimand et al., 2011a; Guirimand et al., 2011b). G10H and SLS are anchored to the endoplasmic reticulum (ER) membrane with their active sites on the cytosolic side (Madyastha et al., 1977; Guirimand et al., 2009; Guirimand et al., 2011a), whereas IRS and LAMT are cytosolic proteins (Madyastha et al., 1973; Geu-Flores et al., 2012).

TDC, STR, and SGD are preferentially detected in epidermal cells (St-Pierre et al., 1999; Murata and De Luca, 2005; Guirimand et al., 2011b). TDC and SGD are cytosolic proteins, and tryptamine and secologanin must subsequently be transported to the vacuole to undergo the STR-mediated conversion to strictosidine (Deus-Neumann

and Zenk, 1984; De Luca and Cutler, 1987; McKnight et al., 1991; Guirimand et al., 2011a).

Vindoline and the dimeric indole alkaloids are found exclusively in leaves and stems, where vindoline localizes to specialized idioblast and laticifer cells in the mesophyll (Westkemper et al., 1980; Deus-Neumann et al., 1987; Balsevich and Bishop, 1989; Roepke et al., 2010). Vindoline synthesis takes place in the leaf epidermis and in the laticifer and idioblast cells within the leaf. T16H and 16OMT localize to leaf epidermal cells (Murata and De Luca, 2005; Guirimand et al., 2011b; Besseau et al., 2013), and expression studies suggest that *T3O*, *T3R*, and *NMT* may also be expressed in the epidermis (Qu et al., 2015). The final two enzymes in vindoline synthesis, D4H and DAT, localize to the laticifers and idioblasts (St-Pierre et al., 1999). T16H is anchored to the ER as a monomer (St-Pierre and De Luca, 1995; Guirimand et al., 2011b). T3O is also likely to be ER membrane-associated, whereas 16OMT and T3R are cytosolic proteins (Levac et al., 2008; Guirimand et al., 2011b; Qu et al., 2015). NMT activity has been found to be associated with the thylakoid membrane of chloroplasts (De Luca and Cutler, 1987; De Luca et al., 1987) while D4H and DAT operate in the nucleocytoplasmic compartment following passive diffusion into the nucleus (Guirimand et al., 2011b).

Of particular interest is catharanthine, which is synthesized within the leaf but subsequently secreted to the leaf surface via a catharanthine transporter, CrTPT2, which is expressed in the leaf epidermis (Roepke et al., 2010; Yu and De Luca, 2013). Catharanthine is also equally distributed to both aboveground and underground tissues (Westkemper et al., 1980; Deus-Neumann et al., 1987; Balsevich and Bishop, 1989;

Roepke et al., 2010). This spatial separation of catharanthine and vindoline is believed to be responsible for the extremely low levels of vinblastine that are produced, as the two MIA monomers must come together for dimerization in the vacuole, where PRX1 is localized (Sottomayor et al., 1996; Costa et al., 2008). Notably, the by-products of T3O in *T3R*-silenced plants accumulate on the leaf surface like catharanthine, indicating transport by CrTPT2 or another unidentified transporter (Qu et al., 2015).

The tissue-specific expression of the genes in vinblastine biosynthesis is one part of the complex regulation of the pathway, in which many genes are also under various developmental- and environment-specific controls.

A.7. The vinblastine biosynthetic pathway is under complex regulatory controls.

MIA biosynthesis in *C. roseus* is under strict regulation by factors such as developmental stage, pathogen attack, light, nutrients, and plant hormones. During seedling development, early-stage enzymes TDC and STR exhibit peak activity by 104 hours (5 d), while late-stage enzymes NMT, D4H, and DAT peak by 150 hours (6 d) (De Luca et al., 1986; De Luca et al., 1988; De Carolis et al., 1990). Interestingly, T16H activity does not peak until day 9 of seedling development (St-Pierre and De Luca, 1995). In mature plants, co-expression of MIA biosynthetic genes, including those in the iridoid pathway, is observed with highest activity in younger leaves (De Luca et al., 1985; De Carolis et al., 1994; Murata et al., 2008; Salim et al., 2014). Environmental factors, regulators, and signalling compounds can exhibit further influence on MIA biosythesis; for example, mechanical wounding or exposure to ethylene has no effect on catharanthine but increases vindoline levels in seedlings (Vázquez-Flota et al., 2004).

Plant growth regulators, including auxin, cytokinins, and abscisic acid (ABA), all

have effects on MIA accumulation that have been studied in *C. roseus* tissue culture. Auxin reduces alkaloid accumulation by down-regulating TDC and STR activity at the transcriptional level (Eilert et al., 1986; Eilert et al., 1987a; Roewer et al., 1992; Pasquali et al., 1992; Arvy et al., 1994). Cytokinins such as zeatin have been shown to enhance G10H activity, increase alkaloid accumulation, and increase conversion of secologanin to ajmalicine (Decendit et al., 1992; Decendit et al., 1993). Increased catharanthine and vindoline accumulation by a *C. roseus* cell culture in response to ABA was reported in one study (Smith et al., 1987) but this response was not consistently reproduced (El-Sayed and Verpoorte, 2002).

In addition to growth regulators, signalling compounds such as jasmonic acid (JA) and methyl jasmonate (MeJA) have also been studied for their effects on MIA metabolism in *C. roseus*. MeJA and JA induce MIA accumulation in hairy roots (Rijhwani and Shanks, 1998; Morgan and Shanks, 1999), cell culture (Gantet et al., 1999), and seedlings (Vazquez-Flota and De Luca, 1998b) by inducing expression of multiple genes (van der Fits and Memelink, 2000; Geerlings et al., 2000). MeJA increases transcription of *GES*, *G10H*, *TDC*, *STR*, and *SGD* in cell culture (Geerlings et al., 2000; Collu et al., 2001; Simkin et al., 2013) and enhances activities of TDC, STR, D4H, and DAT in seedlings (Aerts et al., 1994). MeJA-mediated induction of TDC and D4H activities in seedlings occurs at the transcriptional, post-transcriptional, and post-translational levels (Vazquez-Flota and De Luca, 1998b). A class of MeJA- and JA-responsive transcription factors, Octadecanoid-responsive Catharanthus AP2/ERF-domain (ORCA), are responsible for mediating several MIA biosynthetic genes in response to MeJA and JA and include ORCA1, ORCA2, and ORCA3 (Menke et al.,

1999; van der Fits and Memelink, 2000; Memelink et al., 2001). ORCA2 has been shown to transactivate *STR* (Menke et al., 1999), while ORCA3 binds to the promoter regions of *STR* and *TDC* (van der Fits and Memelink, 2001). Overexpression of *ORCA3* in cell culture has been shown to increase expression of *TDC*, *STR*, *SGD*, and *D4H*, with no effect on *G10H* or *DAT* expression (van der Fits and Memelink, 2000). Effects of MeJA on plants are dependent on the stage of development: MeJA dramatically enhances MIA accumulation if applied before germination, and the ability of MeJA to enhance production decreases sharply with age (Aerts et al., 1994). Exogenous factors such as pathogen infection, simulated by treatment with fungal elicitors, also affect MIA metabolism, although this action has not been well-studied.

MIA accumulation and TDC and STR activities are transcriptionally induced in cell culture by the addition of fungal elicitor preparations or yeast extract (Eilert et al., 1986; Eilert et al., 1987a; Roewer et al., 1992; Pasquali et al., 1992; Menke et al., 1999). In contrast, G10H activity is slightly reduced by elicitor treatment (Moreno et al., 1996). Enhanced accumulation of catharanthine and ajmalicine has been observed in cell culture in response to tetramethyl ammonium bromide and *Aspergillus niger* homogenate (Zhao et al., 2001). Investigation of the regulatory effects of light, however, has been found to be essential to MIA biosynthesis.

Light is an important regulator of vindoline biosynthesis. Etiolated (dark-grown) seedlings accumulate tabersonine and catharanthine as the major alkaloids due to low T16H, D4H, and DAT activities despite high TDC, STR, and NMT activities (De Luca et al., 1986; Balsevich et al., 1986; De Luca et al., 1988; De Carolis et al., 1990; St-Pierre and De Luca, 1995). Exposing etiolated seedlings to light results in the rapid conversion

of tabersonine to vindoline (Balsevich et al., 1986). Light exposure induces both *D4H* gene expression and protein accumulation, and etiolated and light-grown seedlings exhibit different major isoforms of D4H protein, suggesting that its light-mediated activation may involve a post-translational modification in addition to transcriptional regulation (De Carolis et al., 1990; Vazquez-Flota and De Luca, 1998a). Light exposure also induces DAT and T16H activity by as much as 10-fold and 6-fold, respectively (De Luca et al., 1986; De Luca et al., 1988; De Carolis et al., 1990; St-Pierre and De Luca, 1995; Schröder et al., 1999). Additionally, red-light treatment has been shown to be sufficient for inducing D4H and DAT activity, while far-red light has the opposite effect, suggesting the involvement of phytochrome in the process (Vazquez-Flota and De Luca, 1998; Aerts and De Luca, 1992).

MIA biosynthesis in *C. roseus* is clearly a complex system. Improvement of MIA accumulation in cell culture or metabolic engineering of the pathway require a greater understanding of the enzymatic steps, genes, and regulatory processes involved. To this end, many new methods have been generated as tools to investigate the vinblastine pathway in *C. roseus*.

A.8. Techniques for identification and screening of candidate MIA pathway genes.

Traditional approaches that were used in earlier investigations to identify steps in vinblastine biosynthesis involved the use of radiolabelled substrates for feeding to cell cultures to identify intemediates, isolation and purification of target enzymes from plants or cell cultures, protein sequencing for determining peptide sequences, and design of degenerate primers based on protein sequences or homologous genes to probe cDNA libraries for candidate genes. Although many genes have been identified in this manner,

the methods are time-consuming and not always successful (De Luca et al., 2012). More efficient approaches are necessary for the completion of the pathway, and much recent progress has been made. Efforts have focused on large-scale transcriptome sequencing projects for *C. roseus* and other medicinal plants to generate searchable databases (Facchini et al., 2012; Xiao et al., 2013 [http://phytometasyn.ca]; Góngora-Castillo et al., 2012a; Góngoroa-Castillo et al., 2012b [http://www.medicinalplantgenomics.msu.edu]; the Medicinal Plant Transcriptome Project [http://www.uic.edu/pharmacy/MedPlTranscriptome/index.html]). The availability of these transcriptomic databases has led to the recent identification of several iridoid and MIA biosynthetic genes using bioinformatics approaches to search for homologous genes across MIA- and/or iridoid-producing species. Additionally, the tissue-specific expression of the MIA pathway led to the development of the carborundum abrasion technique (Murata and De Luca, 2005; Murata et al., 2008; Levac et al., 2008) for isolation of epidermal RNA and proteins. Epidermis-enriched cDNA libraries produced via random sequencing of the RNA allows for comparison of gene expression between the leaf epidermis and whole leaf and identification of candidate genes involved in epidermis-specific pathways. Once candidate genes have been identified, an additional useful method has been developed for screening and confirming gene functions: virus-induced gene silencing (VIGS).

VIGS experiments in *C. roseus* utilize a tobacco rattle virus (TRV) vector system (Burch-Smith et al., 2004), into which a 300- to 500-bp fragment of the target gene is cloned (De Luca et al., 2012; Liscombe and O'Connor, 2011). *C. roseus* seedlings or mature plants are subsequently infiltrated with transformed *A. tumfasciens* carrying the

VIGS constructs for silencing of the target gene within the plant. Monitoring the change in MIA profile thus allows for confirmation of gene function or screening of candidate genes, and several MIA gene functions have been identified or supported utilizing this technique (Geu-Flores et al., 2012; Salim et al., 2014; Asada et al., 2013; Salim et al., 2013; Kellner et al., 2015; Qu et al., 2015).

An additional approach to identifying candidate genes lies in the screening and profiling of different cultivars or mutant plant lines that exhibit altered MIA profiles. For example, screening of 50 different cultivars led to the identification of an altered profile in *C. roseus* cv. Vinca Mediterranean Dp Orchid, which accumulates 10 times less vindoline than the reference standard *C. roseus* cv. Little Delicata due to reduced T16H activity (Magnotta et al., 2006). Similarly, mutagenesis techniques can result in changes in secondary metabolism, due to alterations in regulatory or functional genes that can be identified via screening in large-scale experiments. Mutagenesis approaches include treatments with chemical mutagens, irradation, and random genetic insertions to generate a population of mutants (Kim et al., 2006; Greene et al., 2003). Chemical mutagenesis often involves ethylmethanesulfonate (EMS), an alkylating agent that forms adducts with nucleotides, primarily (>99%) guanine residues (Greene et al., 2003). Guanine alkylation by EMS produces O^6-ethylguanine, which pairs with thymine instead of cytosine and results in replacement of G/C pairs with A/T pairs through DNA repair mechanisms (Greene et al., 2003; Haughn and Somerville, 1987). These replacements can produce silent, missense, or nonsense mutations, which correspond to codons which result in the same amino acid, a different amino acid, or a premature stop codon, respectively. EMS-induced mutations are typically made up of ~30% silent, ~65% missense, and ~5%

nonsense mutations (McCallum et al., 2000). EMS is an attractive mutagen as it produces high levels of random point mutations, at ~1 mutation per 200-500 kbp, with low rates of chromosome breaks that would otherwise lead to inversions or deletions (Greene et al., 2003; Kurowska et al., 2011). Resulting mutants may therefore display gains or losses of function.

In EMS mutagenesis experiments, seeds are first soaked in the mutagen and germinated to generate the M1 generation of plants, which undergo self-pollination to yield the M2 generation seeds (Bush and Krysan, 2010). Screening can be subsequently performed on the M2 generation using a variety of methods depending on the plant and the targets of interest. PCR-based methods can be used to screen for mutations in specific, known gene targets (Kwok, 2011) while identification of novel genes involved in primary or secondary metabolic pathways may involve screening by visual phenotype or metabolic profiling. For example, genes involved in saponin biosynthesis in oats have been identified using chemically-mutagenized plants that were screened based on fluorescence under UV light (Papadopoulou et al., 1999; Qin et al., 2010; Geisler et al., 2013); similarly, EMS-induced mutants of *C. roseus* with increased salt tolerance, antibacterial activity, or alkaloid content have been selected based on altered morphology such as dwarf phenotypes, though these mutants exhibit many pleiotropic effects that complicate identification of mutations and functional genes (Verma and Singh, 2010; Kulkarni et al., 1999; Rai et al., 2003; Kumari et al., 2013).

In the absence of morphological phenotypes, such as with mutations that affect a secondary metabolic functional gene rather than a central regulatory element, alkaloid profiling is easily accomplished in *C. roseus* by dipping leaves in chloroform to separate

the surface alkaloids, followed by acid-base extraction of surface and whole leaf alkaloids and analysis via thin-layer chromatography (TLC) (Roepke et al., 2010). Once a mutant with an altered profile of interest is obtained, identification of the mutation can be accomplished through a variety of approaches, including gene expression and enzymatic profiling (Magnotta et al., 2006) and transcriptome sequencing (Qin et al., 2010; Geisler et al., 2013; Thamm, 2014). This approach has been successfully used in identifying and characterizing a high-ajmalicine-accumulating mutant line of *C. roseus*, from which three new transcription factors involved in MIA biosynthesis and a clade of perakine reductase-like enzymes were identified (Czerwinski, 2011; Thamm, 2014), demonstrating the potential of EMS mutagenesis in novel MIA gene discovery.

A.9. Objectives and experimental design.

Following the same approach that led to the discovery of the high-ajmalicine mutant (Czerwinski, 2011), another collection of 4000 EMS-mutagenized M2 lines of *C. roseus* was screened for altered MIA profiles with the aim of identifying novel functional or regulatory genes involved in MIA biosynthesis. The screening led to the identification of line M2-1865, which produced low levels of vindoline within the leaves and accumulated 2,3-epoxide derivatives of tabersonine on the leaf surface.

Since the 2,3-epoxides are produced by T3O in the absence of T3R (Qu et al., 2015), it was hypothesized that the M2-1865 phenotype was the result of: 1) a change in *T3R* gene expression due to either a mutation in a regulatory gene or in *T3R*, and/or 2) a change in enzymatic activity of T3R due to a mutation affecting protein function. Investigation of the hypothesis was carried out utilizing detailed analyses of MIA gene expression, T3R nucleotide and amino acid sequences, and enzymatic activity of

recombinant T3R. While no changes in MIA gene expression were observed between the parental line (WT) and M2-1865 line, the mutant T3R possessed a single amino acid substitution (H189Y) which reduced recombinant enzymatic activity by 95%.

B. Materials and Methods

B.1. General materials.

All chemicals were purchased from Sigma-Aldrich unless otherwise specified. Tabersonine was purchased from Omnichem while other alkaloids were purified by previous lab members from extracted plant material. Reagents for RNA extraction and cDNA synthesis were purchased from Invitrogen. Restriction enzymes were purchased from New England Biolabs. Rabbit anti-c-Myc and rabbit His-probe primary antibodies were purchased from Santa Cruz Biotechnology, while IRDye800 Goat anti-rabbit secondary antibody was purchased from Rockland-Inc.

B.2. Plant materials and mutant screening.

C. roseus plants were cultivated in the Brock University greenhouse under long day conditions (16 hours light, 8 hours dark) at 28ºC. Plants were watered 2-3 times per week and fertilized twice per month.

EMS-mutagenized M2 seeds of *C. roseus* cv. Pacifica Apricot were obtained from Ball Seed Co. (Ball Horticultural Company, West Chicago IL). MIAs were extracted from seedlings as described below and 4000 lines were initially screened by undergraduate student Michael Easson via thin-layer chromatography. Candidate mutants with altered MIA profiles were subsequently verified using UPLC-MS.

B.3. Alkaloid extraction.

Leaf pairs were freshly harvested and weighed. Surface MIAs were extracted according to the protocol developed by Roepke et al. (2010) via dipping the leaves in 2 mL chloroform for one hour, followed by acid/base extraction: chloroform was extracted with an equal volume of acidified water (pH 3.5-4.5), which was subsequently basified

(pH 9-10) and extracted with an equal volume of ethyl acetate; the ethyl acetate was then evaporated to dryness and the extracted MIAs were dissolved in 200 µL HPLC-grade methanol. After chloroform dip, leaves were then extracted in 1 mL HPLC-grade methanol for 1 hour with shaking. MIAs were analyzed and quantified via UPLC-MS.

B.4. Crossings and phenotype segregation.

M2-1865 was first confirmed to be homozygous for the mutant phenotype via analysis of MIA profiles in M3 offspring obtained by self-fertilization. Reciprocal crossings were subsequently made between parental line (WT) and M2-1865 plant lines. For the F1 generation, 10 offspring from WT male x M2 female and 4 offspring from WT female x M2 male crossings were obtained. 10-20 seeds of each self-fertilized plant in the F1 generation were collected to generate the F2 generation, of which 60 6-week-old offspring were randomly selected for analysis. Phenotypic analysis was carried out primarily by determining levels of 2,3-epoxides on the leaf surface, which were only observed in homozygous mutants. Vindoline levels were also quantified. Leaf pair 2 (LP2) was used for phenotypic analysis of M3, F1, and F2 generation plants.

B.5. RNA extraction and cDNA synthesis.

RNA was extracted from freshly-harvested plant tissues of three individual plants of both WT and M2-1865 lines using TRIzol reagent according to the manufacturer's instruction and was purified using overnight LiCl precipitation at 4°C. Reverse transcription was then performed on a T100 Thermal Cycler from Bio-Rad using SuperScript II Reverse Transcriptase, Oligo d(T) 12-18 Primer, and 1-5 µg total RNA according to the manufacturer's instructions.

B.6. Real-Time PCR and gene expression.

cDNA extracted from WT and M2-1865 plant lines as described (B.5.) was used as the template in RT-PCR reactions using iTaq SYBR Green Supermix (Bio-Rad) and 5 µM primers in a total reaction volume of 10 µL. The reactions were performed with a C1000 Thermal Cycler with CFX96 Real-Time System (Bio-Rad) under the following conditions: 95 °C for 3 min, followed by 40 cycles of 95 °C for 10 s and 58 °C for 30 s. Gene expression was determined relative to 60S ribosomal subunit as the reference gene. Primer sequences for RT-PCR are listed in Supplementary Table 1.

B.7. Plant protein extraction

Young leaves (LP 1-3) were freshly harvested and pulverized with sea sand in 3 mL buffer (50 mM HEPES, pH 7.5) per gram of tissue. The resulting slurry was filtered through three layers of cheesecloth and centrifuged at 5000 rpm, 4ºC, for 10 min. The supernatant was partially purified through a 100-mL column packed with Sephadex G-25 Medium (Amersham Biosciences) at a flow rate of 1 mL/min. The protein eluted in the first 30 mL of the void fraction and the 30-70% ammonium sulfate fraction was desalted through a PD-10 Sephadex G-25 M Column (GE Healthcare). The protein extracts of WT and M2-1865 lines were assayed for T3R activity.

B.8. Isolation of microsomal protein from yeast expressing *T3O*.

Recombinant yeast transformed with the pESC-Leu2d construct containing the *C. roseus* cytochrome P450 reductase (*CPR*) (Salim et al., 2014) and *T3O* genes (Qu et al., 2015) were obtained from post-doctoral fellow Vince Qu. 2 mL of overnight yeast culture was inoculated into 200 mL of SC-Leu medium containing 2% glucose and grown at 30°C with shaking at 200 rpm for 24 hours, after which the medium was switched to SC-Leu containing 2% galactose to induce protein expression for another 24 hours. Cells

were harvested by centrifugation at 6000 rpm for 10 min and resuspended in 2-3 mL TES buffer (10 mM Tris-HCl, 1 mM EDTA, and 0.6 M sorbitol; pH 7.5). Cells were broken in a TissueLyser II (Qiagen) by shaking at 30 Hz for 5 min at 4°C with acid-washed glass beads. The lysate was centrifuged at 10 000 g for 10 min at 4°C and the supernatant was transferred to ultracentrifuge tubes. Microsomes were pelleted at 100 000 g for 1 hour at 4°C and resuspended in TEG buffer (10 mM Tris-HCl, 1 mM EDTA, and 10% glycerol; pH 7.5). Microsomal protein concentration was determined via Bradford assay. Microsomes were then flash-frozen in liquid nitrogen and stored at -80°C.

B.9. *T3R* cloning and sequencing.

Cloning of the WT and M2-1865 *T3R* genes was carried out according to the protocol established by Qu et al. (2015). cDNA was synthesized from RNA extracted from young leaves of *C. roseus* WT and M2-1865 lines as described (B.5., B.6.). The *T3R* open reading frame (ORF) was amplified using High-Fidelity Phusion DNA Polymerase (New England BioLabs) with cloning primers designed by post-doctoral fellow Vince Qu: the forward primer 5'-ATAAGGATCCAATGGCTGCAAAGTCAGTGAAGG-3', which contained a *BamH*I site (underlined), and the reverse primer, 5'-AGAAGTCGACGGGTGATTTGAAAGTG TTTCCAATG-3', which contained a *Sal*I site (underlined). The PCR fragment was inserted into pGEMT-Easy vector (Promega), which was transformed into competent *Escherichia coli* strain TOP10 via heat shock. Positive transformants were identified using blue/white screening and PCR. Recombinant plasmid constructs were purified and sent to Eurofins Genomics for sequencing. The *T3R* ORF fragment was subsequently digested from the pGEMT-Easy vector using *BamH*I-HF and *Sal*I-HF and inserted into

the pET30-b expression vector (Novagen). Ligation was confirmed via double digests with *BamH*I-HF/*Sal*I-HF, *BamH*I-HF/*Xho*I, *Nco*I-HF/*Sal*I-HF, and *Nco*I-HF/*Xho*I. Competent *E. coli* strain TOP10 was transformed with the constructs and recombinant plasmids were purified from positive transformants as identified via PCR. Recombinant plasmids were subsequently transformed into *E. coli* strain BL21-DE3 for protein expression.

B.10. Recombinant protein extraction from *E. coli* expressing *T3R*.

Purification of recombinant protein was performed according to Qu et al. (2015). Recombinant BL21-DE3 *E. coli* expressing pET30-b-*T3R* were grown at 37°C and 250 rpm to OD_{600} = 0.400 in LB medium containing 50 µg/mL of kanamycin. Protein expression was subsequently induced with 0.1 mM IPTG overnight at 10°C with shaking at 200 rpm. Cells were pelleted by centrifugation at 6000 rpm for 10 min at 4°C and resuspended in 4 mL lysis buffer for purification (10 mM Tris-HCl, 100 mM NaCl, 10 mM imidazole, 1 mM DTT, 1 mM PMSF, and 10% glycerol; pH 7.5) or desalting (10 mM Tris-HCl, 100 mM NaCl, 1 mM PMSF, and 10% glycerol; pH 7.5). Cells were then lysed using a Fisher Scientific Sonic Dismembrator Model 100 with six rounds of 0.5-second pulses at setting 10 for 10 seconds, with 10-second cooling periods in-between. The lysate was subsequently pelleted at 8000 rpm for 10 min at 4°C.

For purification via Ni-NTA affinity chromatography, the supernatant was harvested and loaded onto an equilibrated column packed with 2 mL Ni-NTA resin (Qiagen). The resin was washed with 50 mL wash buffer (10 mM Tris-HCl, 100 mM NaCl, 30 mM imidazole, and 10% glycerol; pH 7.5) and the recombinant protein was eluted with elution buffer (10 mM Tris-HCl, 100 mM NaCl, 300 mM imidazole, and 10%

glycerol; pH 7.5) in 500-µL fractions. Protein concentration was measured via Bradford assay and fractions were flash-frozen in liquid nitrogen before storing at -80°C.

For desalting of total crude protein extracts, 2.5 mL of the supernatant was eluted through an equilibrated PD-10 Sephadex G-25 M Column (GE Healthcare) with 3.5 mL of lysis buffer. Protein concentration was measured via Bradford assay and the total desalted protein was flash-frozen in liquid nitrogen and stored at -80°C in 500-µL aliquots.

B.11. SDS-PAGE and immunoblotting.

Purified T3R (5 µg), total soluble *E. coli* protein (10 µg), total insoluble *E. coli* protein (5 µg), or yeast microsomal protein (50 µg) were separated via SDS-PAGE at 200 V for 50-60 min using 12% resolving gel. Gels were transferred to a 0.45 µm Amersham Hybond-ECL PVDF protein blotting membrane (GE Healthcare) using a Trans-Blot SD Semi-Dry Transfer Cell (Bio-Rad). The membrane was blotted with 0.5 µg/L rabbit anti-c-Myc primary antibody for detection of recombinant T3O or with 0.4 µg/mL rabbit His-probe primary antibody for detection of His-tagged recombinant T3R. Both membranes were blotted with 0.1 µg/L IRDye800 Goat anti-rabbit secondary antibody. Membranes were visualized with a Li-COR Odyssey 9120 scanner.

B.12. T3R *in vitro* assays.

In order to test for T3R activity in the WT and M2-1865 plant protein extracts, 200-µL assays were prepared using 100 µg recombinant microsomal T3O, 568 µg partially-purified plant protein extract, 25 µM tabersonine, 1 mM NADPH, and 50 mM HEPES buffer, pH 7.5.

Assays using purified recombinant T3R were prepared with 50 µg recombinant

microsomal T3O, recombinant *E. coli* protein (4 µg purified recombinant T3R, 50 µg total desalted WT extract, or 500 µg total desalted M2 extract), 25 µM tabersonine or 16-methoxytabersonine, 1 mM NADPH, and 50 mM HEPES buffer, pH 7.5, to a total volume of 100 µL.

Both assays were incubated at 30°C for 1 hour and were subsequently basified and extracted twice with ethyl acetate. The ethyl acetate was evaporated to dryness and alkaloids were dissolved in 120 µL HPLC-grade methanol for analysis via UPLC-MS.

B.13. UPLC-MS.

Extracted alkaloids from plant material and *in vitro* assays were analyzed on an Acquity UPLC system (Waters) equipped with a BEH C18 column (2.1 x 50 mm; particle size 1.7 µm) maintained at 35°C, a photodiode array detector, and a mass spectrometer. The solvent systems used were solvent A (6:14:80 methanol:acetonitrile:5 mM ammonium acetate) and solvent B (24:64:10 methanol:acetonitrile:5 mM ammonium acetate). The following linear elution gradient was used: 0-0.5 min 99% A, 1% B, 0.3 mL/min; 0.5-0.6 min 99% A, 1% B, 0.4 mL/min; 0.6-8.0 min 1% A, 99% B, 0.4 mL/min; 8.0-8.3 min 99% A, 1% B, 0.4 mL/min; 8.3-10 min 99% A, 1% B, 0.3 mL/min. The mass spectrometer was operated in positive ionization mode at a capillary voltage of 3.1 kV, a cone voltage of 48 V, a desolvation gas flow of 600 L/hr, a desolvation temperature of 350°C, and a source temperature of 150°C.

MIAs were quantified utilizing standard curves generated from known concentrations. As no standards were available for 16-methoxy-2,3-epoxytabersonine or 16-methoxy-2,3-dihydro-3-hydroxytabersonine, they were quantified using standards of 2,3-epoxytabersonine and 2,3-dihydro-3-hydroxytabersonine, respectively.

B.14. Sequence alignments and phylogenetic analysis.

Nucleotide and amino acid sequence alignments were performed using the M-Coffee tool at http://tcoffee.crg.cat (Notredame et al., 2000; Wallace et al., 2006; Moretti et al., 2007; Di Tommaso et al., 2011). Phylogenetic analysis was performed with the Phylogeny.fr tool at http://www.phylogeny.fr using ProbCons for sequence alignment, GBlocks for curation, PhyML for the phylogeny, and TreeDyn for rendering (Dereeper et al., 2010; Dereeper et al., 2008; Do et al., 2005; Castresana, 2000; Guindon and Gascuel, 2003; Anisimova and Gascuel, 2006; Chevenet et al., 2006). Species and accessions for the sequences used in the phylogenetic analysis are listed in Supplementary Table 2.

C. Results

C.1. M2-1865 has reduced levels of vindoline and accumulates tabersonine 2,3-epoxide derivatives on the leaf surface.

In order to determine the MIA profile of line M2-1865, MIAs were extracted from the leaf surface and leaf body by acid/base extraction (as described in B.3.) and analyzed by UPLC-MS. Initial UPLC analyses suggested that the leaf body of M2-1865 (Figure 10B) contained significantly reduced levels of vindoline (RT = 4.61 min), vindorosine (RT = 4.77 min), vinblastine (RT = 5.00 min), and 3',4'-anhydrovinblastine (RT = 5.85 min) compared to those of the parental line (WT) (Figure 10A). In contrast, the M2-1865 leaf surface extracts (Figure 10D, 10G-J) accumulated new MIAs [16-methoxy-2,3-epoxytabersonine (RT = 3.40 and 3.72 min, m/z = 383) and 2,3-epoxytabersonine (RT = 3.54 and 3.79 min, m/z = 353)] not found on the leaf surface of WT (Figure 10C). Incubation of recombinant T3O with NADPH and either 16-methoxytabersonine or tabersonine generated the same reaction products found on the leaf surface of M2-1865 (Figure 10E-F, 10K-L) [16-methoxy-2,3-epoxytabersonine (RT = 3.38 and 3.70 min, m/z = 383) and 2,3-epoxytabersonine (RT = 3.52 and 3.78 min, m/z = 353)] (Qu et al., 2015). Quantification of MIAs from leaf pair (LP) 1 to LP5 (Figure 11) indicated that vindoline levels were reduced in M2-1865 by as much as 75%, vindorosine by ~50%, and 3',4'-anhydrovinblastine by ~50% compared to WT. Vinblastine levels in the mutant were too low to be quantified. Furthermore, M2-1865 accumulated similar levels of catharanthine, tabersonine, and 16-methoxytabersonine compared to levels found in WT. The M2-1865 plants did not display any morphological changes compared to WT plants. The results suggest that accumulation of 2,3-epoxides in M2-1865 may occur as a result of a change

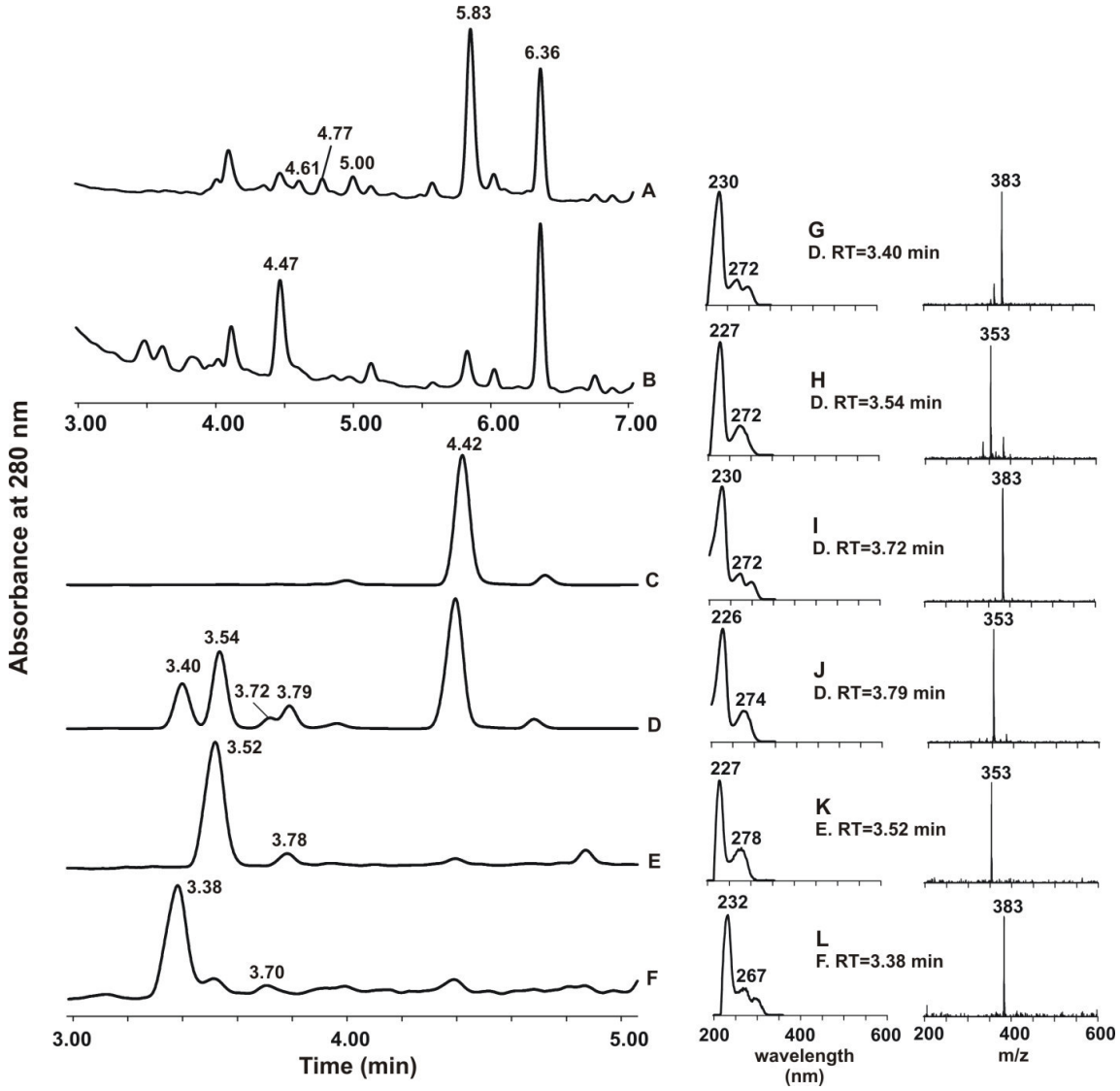


Figure 10: LC profiles of MIA extracts from WT and M2-1865 plants suggest reduced vindoline and accumulation of tabersonine 2,3-epoxide derivatives in the mutant. Extracts of leaf bodies after chloroform dip from WT (**A**) and M2-1865 (**B**) suggest reduced levels of vindoline, vindorosine, and vindoline-containing dimers [vindoline (4.61 min), vindorosine (4.77 min), vinblastine (5.00 min), 3',4'-anhydrovinblastine (5.83 min), comigrating tabersonine and 16-methoxytabersonine (6.36 min)]. Leaf surface extracts from WT (**C**) and M2-1865 (**D**) plants show accumulation of four compounds (3.40-3.79 min) in the mutant that do not accumulate in WT and that have similar retention times, UV absorbance, and mass spectra (**G-J**) to tabersonine 2,3-epoxide derivatives (**K-L**) produced in assays with recombinant T3O using tabersonine (**E**) or 16-methoxytabersonine (**F**) as substrate [catharanthine (4.42 min)].

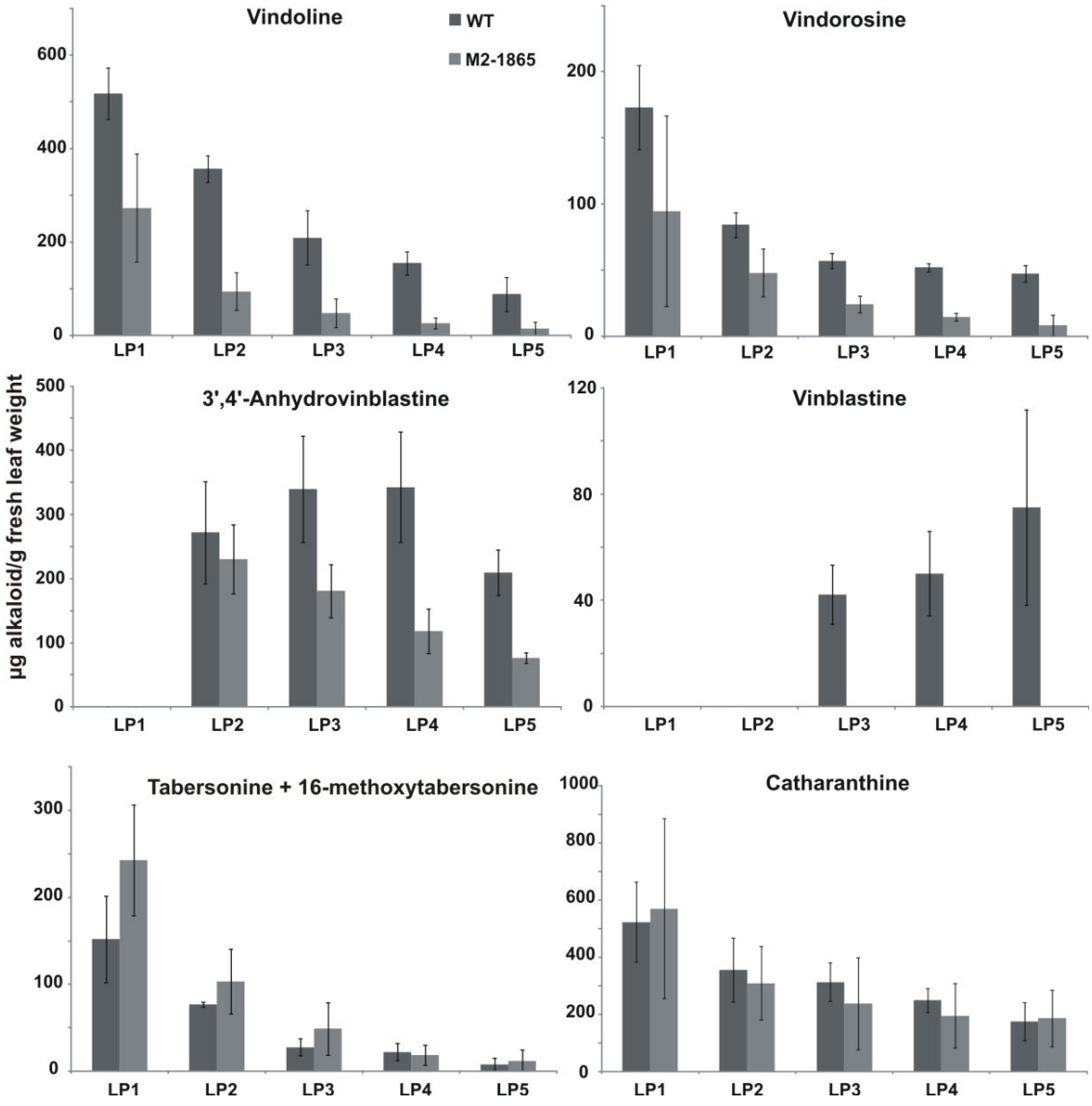


Figure 11: Quantification of MIAs in WT and M2-1865 plant lines by UPLC-MS shows that levels of vindoline, vindorosine, and vindoline-containing dimers are reduced in M2-1865 plants compared to WT. Vinblastine levels in M2-1865 were too low to be quantified. Tabersonine and 16-methoxytabersonine co-migrated and were thus quantified together. Results are the average levels in three individual plants of each line. LP1 = youngest leaf pair.

in the gene expression and/or enzyme activity of T3R.

C.2. Protein extracts from M2-1865 plants do not display T3R activity.

In order to test for T3R activity, protein was extracted from WT and M2-1865 plant lines (as described in B.7.). The crude protein extracts were desalted and concentrated by ammonium sulfate precipitation. The 30-70% desalted ammonium sulfate fraction was used for *in vitro* assays (as described in B.12.) in combination with recombinant yeast microsomes expressing *T3O*, with tabersonine as the substrate and NADPH as the cofactor. UPLC-MS analysis of the reactions indicated that tabersonine (RT = 6.15-6.26 min, m/z = 337) was converted by T3O into 2,3-epoxytabersonine (RT = 3.61 and 3.89 min, m/z = 353) in the absence of plant extracts (Figure 12A). WT plant extract did not display any activity towards the substrate (Figure 12B). In the presence of both T3O and WT plant extract, epoxides did not accumulate as WT protein converted the intermediate produced by T3O into 2,3-dihydro-3-hydroxytabersonine (RT = 4.93 min, m/z = 355) (Figure 12C), but this activity was not observed with M2-1865 protein extract (Figure 12D). This contributed to the likelihood that either change in T3R gene expression or protein functionality in the mutant plant was responsible for the mutant phenotype.

C.3. WT and M2-1865 plant lines show similar expression levels of *T3R* and other MIA biosynthetic genes.

Expression levels of *T3R* in various plant tissues and of other MIA biosynthetic genes in LP1 were compared using real-time PCR as described in B.6. (Figure 13). *T3R* was highly expressed in leaves of both WT and M2-1865 plants, with LP1 and LP2 showing similar expression levels between the two plant lines, and was poorly expressed

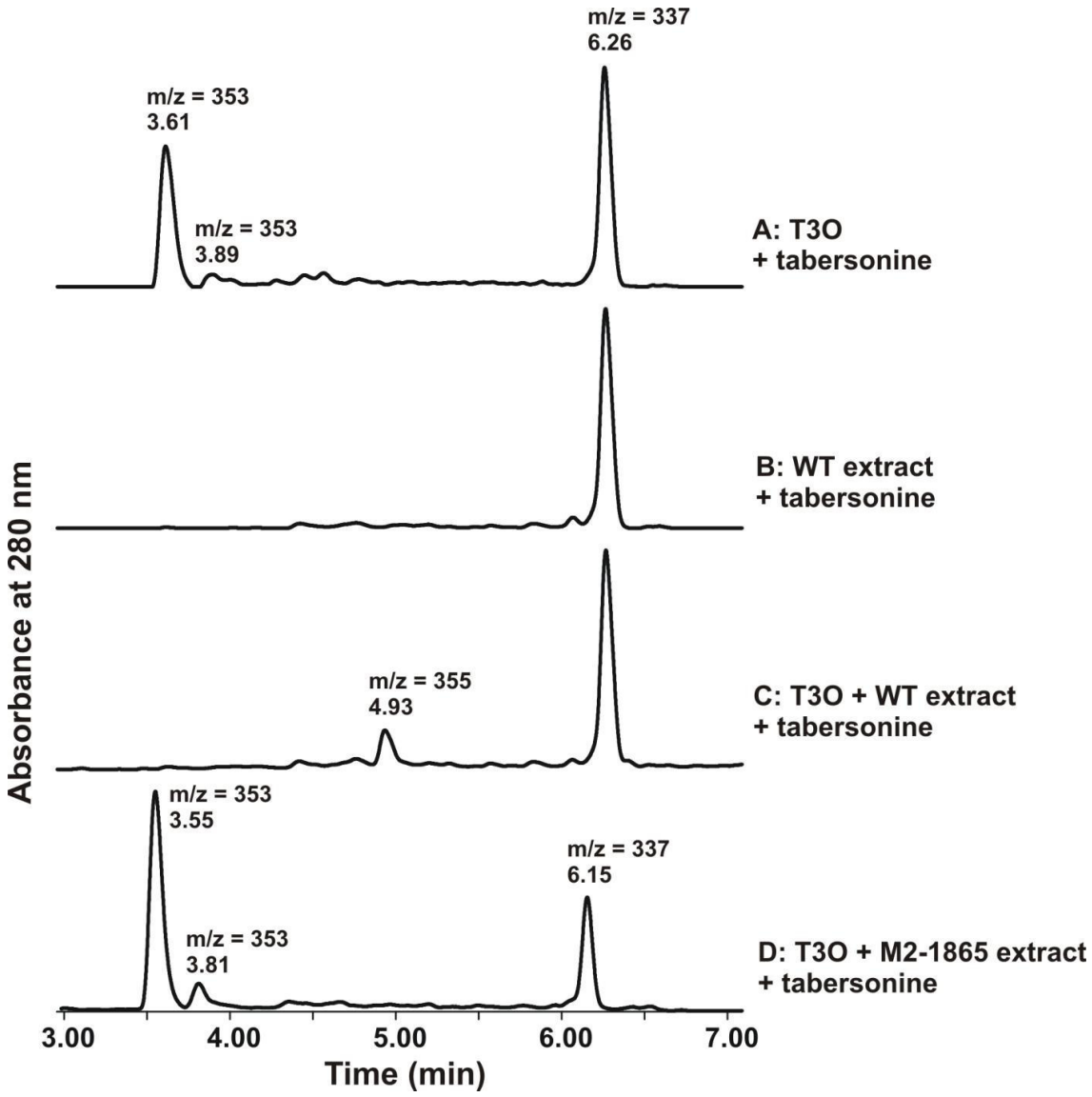


Figure 12: The conversion of tabersonine to 2,3-dihydro-3-hydroxytabersonine only takes place in assays containing both recombinant T3O and WT plant protein extracts but not in those containing M2-1865 plant protein extracts. Assays contained 1 mM NADPH, 25 µM tabersonine, +/- 100 µg recombinant yeast microsomal T3O, and +/- 568 µg WT or M2-1865 plant protein extract [tabersonine (RT 6.15 min, m/z 337), 2,3-epoxytabersonine (RT 3.55 and 3.81 min, m/z 353), 2,3-dihydro-3-hydroxytabersonine (RT 4.93 min, m/z 355)].

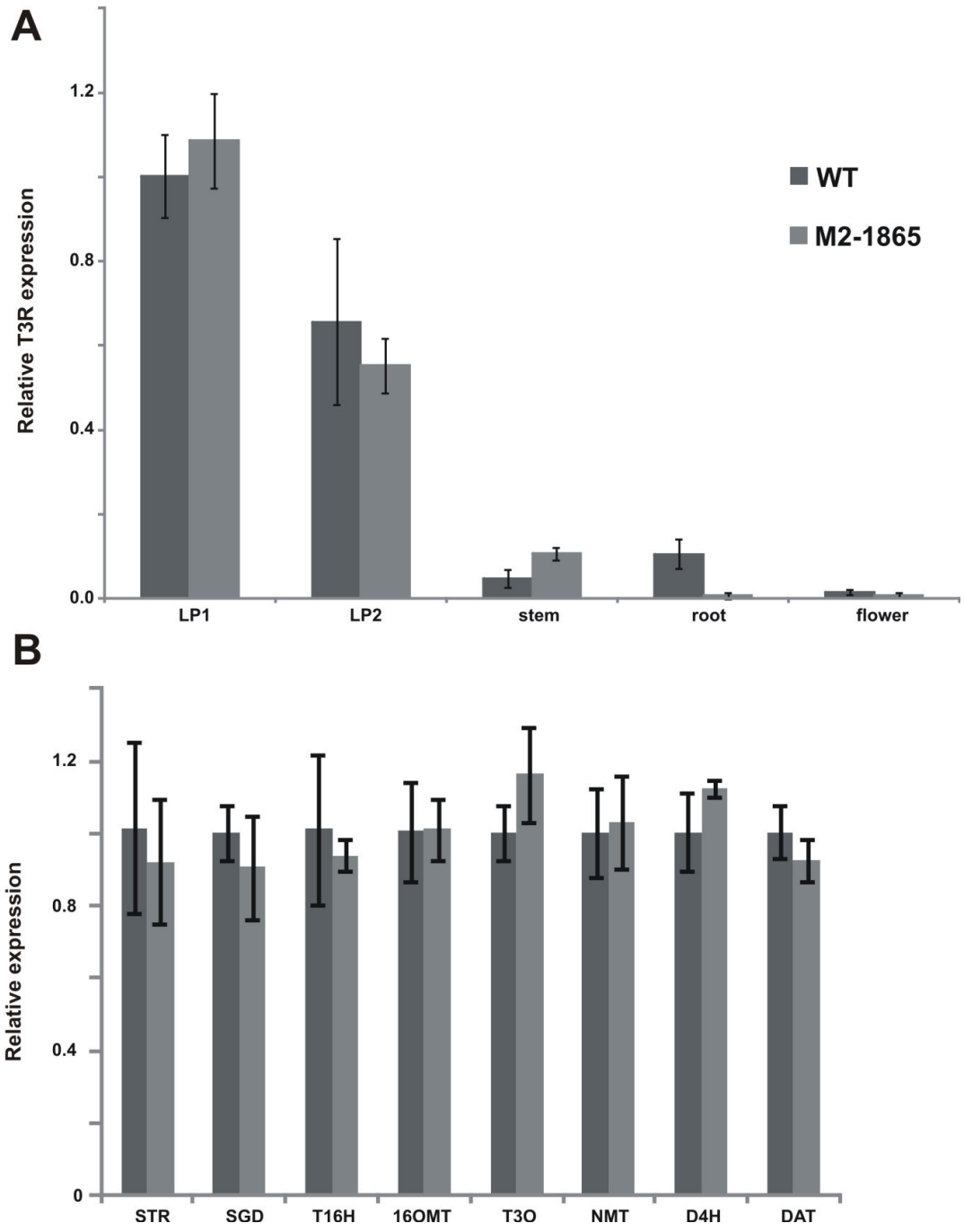


Figure 13: WT and M2-1865 plant lines show similar expression levels of *T3R* as well as other MIA pathway genes. (**A**) *T3R* expression in various organs. (**B**) Expression of other MIA pathway genes in LP1. Expression was determined relative to 60S ribosomal subunit as reference gene using three individual plants from each line. Genes: strictosidine synthase (STR), strictosidine *β*-D-glucosidase (SGD), tabersonine 16-hydroxylase (T16H), 16-methoxytabersonine 16-*O*-methyltransferase (16OMT), tabersonine 3-oxygenase (T3O), tabersonine 3-reductase (T3R), 16-methoxytabersonine-2,3-dihydro-3hydroxytabersonine *N*-methyltransferase (NMT), desacetoxyvindoline 4-hydroxylase (D4H), deacetylvindoline 4-*O*-acetyltransferase (DAT).

in other organs. Similarly, LP1 of both WT and M2-1865 plants exhibited comparable expression of the other known MIA biosynthetic genes from *STR* to *DAT*. These similar gene expression profiles between WT and M2-1865 plants, together with the apparent lack of T3R activity in mutant plant extracts (Figure 12) triggered the molecular cloning of mutant T3R to investigate its molecular and biochemical properties compared with T3R from the WT plants.

C.4. Mutant T3R has two nucleotide substitutions but possesses a single amino acid substitution, H189Y.

Using cDNA synthesized from WT and M2-1865 leaf RNA (see section B.5.), *T3R* was cloned into pGEMT-Easy and transformed into *E. coli* as described in section B.9. Recombinant plasmids encoding WT and M2-1865 *T3R* genes were sequenced and compared (Figure 14) to show that mutations of cytosine residues had occurred at positions 565 and 903 to thymine and adenine, respectively, in the mutant sequence. The mutation at position 903 altered the codon from ACC to ACA, which both code for threonine, and was therefore a silent mutation. However, the mutation at position 565 changed the codon from CAT to TAT, which resulted in a substitution mutation, H189Y, in the mutant protein (Figure 15).

C.5. Mutant T3R activity is reduced by 95% in recombinant protein extracts from *E. coli*.

To investigate the change in catalytic activity produced by the H189Y mutation, *WT-T3R* and *M2-T3R* were transferred into pET30-b (as described in section B.9.) for expression as a 6xHis-tagged protein in *E. coli*. Ni-NTA affinity chromatography was used for purification of the 6xHis-tagged proteins (B.10.); however, while WT-T3R was

```
WT-T3R       1 ATGGCTGCAAAGTCAGTGAAGGCTTTTGGTTTGGCTCTTAAGGATTCATCTGGGCTTTTCTCT   63
M2-T3R       1 ATGGCTGCAAAGTCAGTGAAGGCTTTTGGTTTGGCTCTTAAGGATTCATCTGGGCTTTTCTCT   63

WT-T3R      64 CCATTCAACTTCTCAAGAAGGGCTACAGGGGAACACGATGTTCAATTGAAAGTATTATATTGT  126
M2-T3R      64 CCATTCAACTTCTCAAGAAGGGCTACAGGGGAACACGATGTTCAATTGAAAGTATTATATTGT  126

WT-T3R     127 GGTGTTTGCAATTTCGATAACTTAATGAGGAGAAACAAATACGGGAGGACCAAGTTTCCCTAT  189
M2-T3R     127 GGTGTTTGCAATTTCGATAACTTAATGAGGAGAAACAAATACGGGAGGACCAAGTTTCCCTAT  189

WT-T3R     190 GTTTTCGGGCATGAAATTGTGGGTGTTGTAACTGAGGTTGGTTCCAATGTTAAGAAATTCAAA  252
M2-T3R     190 GTTTTCGGGCATGAAATTGTGGGTGTTGTAACTGAGGTTGGTTCCAATGTTAAGAAATTCAAA  252

WT-T3R     253 ATCGGTAACAAAGTTGGGGTTAGCTTTATCGTTGATACCTGTAGAGAATGTGAAAGGTGTAAA  315
M2-T3R     253 ATCGGTAACAAAGTTGGGGTTAGCTTTATCGTTGATACCTGTAGAGAATGTGAAAGGTGTAAA  315

WT-T3R     316 ATTGGACAGCAAATAGCCTGTAAAAAAGCAGTATCATCGGACGGCTTCTTCGAGACACCAGGT  378
M2-T3R     316 ATTGGACAGCAAATAGCCTGTAAAAAAGCAGTATCATCGGACGGCTTCTTCGAGACACCAGGT  378

WT-T3R     379 TACGGTGGTTGTTCAAATATATTTGTGGCTGACGAGAATTATGTCATACTTTGGCCTGAAAAC  441
M2-T3R     379 TACGGTGGTTGTTCAAATATATTTGTGGCTGACGAGAATTATGTCATACTTTGGCCTGAAAAC  441

WT-T3R     442 CTTCCTATGGATTCTGGGGCACCCCTGCTGTGTATAGGAATTACATGTTATAATCCCTTAAGA  504
M2-T3R     442 CTTCCTATGGATTCTGGGGCACCCCTGCTGTGTATAGGAATTACATGTTATAATCCCTTAAGA  504

WT-T3R     505 CGTTTTGGACTTGATAAACCTGGAGTTAGAGTTGGTATAGTTGGTCTAGGAGCAGTTGGACAT  567
M2-T3R     505 CGTTTTGGACTTGATAAACCTGGAGTTAGAGTTGGTATAGTTGGTCTAGGAGCAGTTGGATAT  567

WT-T3R     568 TTAGCTATTAAATTTGCAAAAGCTTTTGGTGCTAGGGTTACTTTGATCAGTTCATCCCCTGGA  630
M2-T3R     568 TTAGCTATTAAATTTGCAAAAGCTTTTGGTGCTAGGGTTACTTTGATCAGTTCATCCCCTGGA  630

WT-T3R     631 AAAAAGGATGAGGCTTTTCAGAAATTTGGTGTAGATTCTTTCTTGGTCAGCAGTAATGCAGAG  693
M2-T3R     631 AAAAAGGATGAGGCTTTTCAGAAATTTGGTGTAGATTCTTTCTTGGTCAGCAGTAATGCAGAG  693

WT-T3R     694 GAAATGCAGGCTGCAGCTGAAACTCTGGATGGTATCCTAGACACTGTACCAGTGGTTCACCCC  756
M2-T3R     694 GAAATGCAGGCTGCAGCTGAAACTCTGGATGGTATCCTAGACACTGTACCAGTGGTTCACCCC  756

WT-T3R     757 CTTGAGCCACTCTTTGCTTTACTGAAACCTCTTGGGAAACTTATCATTATAGGTGAACCGCAT  819
M2-T3R     757 CTTGAGCCACTCTTTGCTTTACTGAAACCTCTTGGGAAACTTATCATTATAGGTGAACCGCAT  819

WT-T3R     820 AAGCCTTTTGAGGTATCCGCAATGTCCCTCATGGAGGGTGGAAAAATAATTAGCGCGAGTACG  882
M2-T3R     820 AAGCCTTTTGAGGTATCCGCAATGTCCCTCATGGAGGGTGGAAAAATAATTAGCGCGAGTACG  882

WT-T3R     883 GGTGGAAGTATAAAGGACACCCAAGAGATAGTCGATTTTGCAGCAGAACATAACGTAGTTGCA  945
M2-T3R     883 GGTGGAAGTATAAAGGACACACAAGAGATAGTCGATTTTGCAGCAGAACATAACGTAGTTGCA  945

WT-T3R     946 GATGTTGAGGTTATCCCCGTGGACTATGTGAACACTGCCATGGAGCGTCTTGATAAAGCTGAT 1008
M2-T3R     946 GATGTTGAGGTTATCCCCGTGGACTATGTGAACACTGCCATGGAGCGTCTTGATAAAGCTGAT 1008

WT-T3R    1009 GTGAAGTATCGTTTCGTGATTGACATTGGAAACACTTTCAAATCACCCGTCGACTTCTAA 1068
M2-T3R    1009 GTGAAGTATCGTTTCGTGATTGACATTGGAAACACTTTCAAATCACCCGTCGACTTCTAA 1068
```

Figure 14: *T3R* cloned from WT and M2-1865 plant lines shows mutations of cytosines at positions 565 and 903 (highlighted) to thymine and adenine, respectively, in the mutant sequence.

```
WT-T3R       1 MAAKSVKAFGLALKDSSGLFSPFNFSRRATGEHDVQLKVLYCGVCNFDNLMRRN   54
M2-T3R       1 MAAKSVKAFGLALKDSSGLFSPFNFSRRATGEHDVQLKVLYCGVCNFDNLMRRN   54

WT-T3R      55 KYGRTKFPYVFGHEIVGVVTEVGSNVKKFKIGNKVGVSFIVDTCRECERCKIGQ  108
M2-T3R      55 KYGRTKFPYVFGHEIVGVVTEVGSNVKKFKIGNKVGVSFIVDTCRECERCKIGQ  108

WT-T3R     109 QIACKKAVSSDGFFETPGYGGCSNIFVADENYVILWPENLPMDSGAPLLCIGIT  162
M2-T3R     109 QIACKKAVSSDGFFETPGYGGCSNIFVADENYVILWPENLPMDSGAPLLCIGIT  162

WT-T3R     163 CYNPLRRFGLDKPGVRVGIVGLGAVGHLAIKFAKAFGARVTLISSSPGKKDEAF  216
M2-T3R     163 CYNPLRRFGLDKPGVRVGIVGLGAVGYLAIKFAKAFGARVTLISSSPGKKDEAF  216

WT-T3R     217 QKFGVDSFLVSSNAEEMQAAAETLDGILDTVPVVHPLEPLFALLKPLGKLIIIG  270
M2-T3R     217 QKFGVDSFLVSSNAEEMQAAAETLDGILDTVPVVHPLEPLFALLKPLGKLIIIG  270

WT-T3R     271 EPHKPFEVSAMSLMEGGKIISASTGGSIKDTQEIVDFAAEHNVVADVEVIPVDY  324
M2-T3R     271 EPHKPFEVSAMSLMEGGKIISASTGGSIKDTQEIVDFAAEHNVVADVEVIPVDY  324

WT-T3R     325 VNTAMERLDKADVKYRFVIDIGNTFKSPVDF  355
M2-T3R     325 VNTAMERLDKADVKYRFVIDIGNTFKSPVDF  355
```

Figure 15: T3R cloned from WT and M2-1865 plant lines shows a single amino acid substitution, H189Y (highlighted), in the mutant protein.

successfully purified, the method failed to isolate purified M2-T3R (Figure 16A). While the reasons for this failure are not clear, the application of protein extracts from *E. coli* expressing *M2-T3R* to the Ni-NTA resin decolourized the resin under various native and denaturing conditions that were tried and as a result, recombinant mutant T3R was not retained on the resin. Since both soluble recombinant 6xHis-tagged proteins were still detected at similar levels in the same amount of total desalted crude protein extracts in a Western blot (Figure 16B), the desalted crude extracts were used to compare WT-T3R and M2-T3R activities. The coupled enzyme assay also required microsomal T3O protein, which was obtained from recombinant yeast expressing *T3O* (Figure 16C).

Recombinant *E. coli* extracts with WT-T3R or M2-T3R proteins were assayed (B.12.) with tabersonine or 16-methoxytabersonine as substrates in the presence of NADPH and +/- recombinant yeast microsomes expressing *T3O*. Reaction products of basified enzyme assays were extracted with ethyl acetate, dried, and redissolved in methanol for analysis by UPLC-MS (Figure 17). Enzyme assays for T3O activity converted tabersonine (RT = 6.18-6.24 min, m/z = 337) and 16-methoxytabersonine (RT = 6.14-6.20 min, m/z = 367) to the corresponding 2,3-epoxide derivatives [(Figure 17A; RT = 3.52 and 3.78 min, m/z = 353) and (Figure 17B; RT = 3.38 min, m/z = 383)]. Neither recombinant bacterial protein extract exhibited any activity towards either substrate (Figure 17C-D, 17G-H), but when both T3O and recombinant WT-T3R or M2-T3R protein extracts were present, both MIAs were converted into their respective 2,3-dihydro-3-hydroxy derivatives [(Figure 17E, 17I; RT = 4.89 min, m/z = 355) and (Figure 17F, 17J; RT = 4.66 min, m/z = 385)]. Protein extracts from *E. coli* expressing *M2-T3R* exhibited lower levels of conversion to the 2,3-dihydro-3-hydroxylated products, with

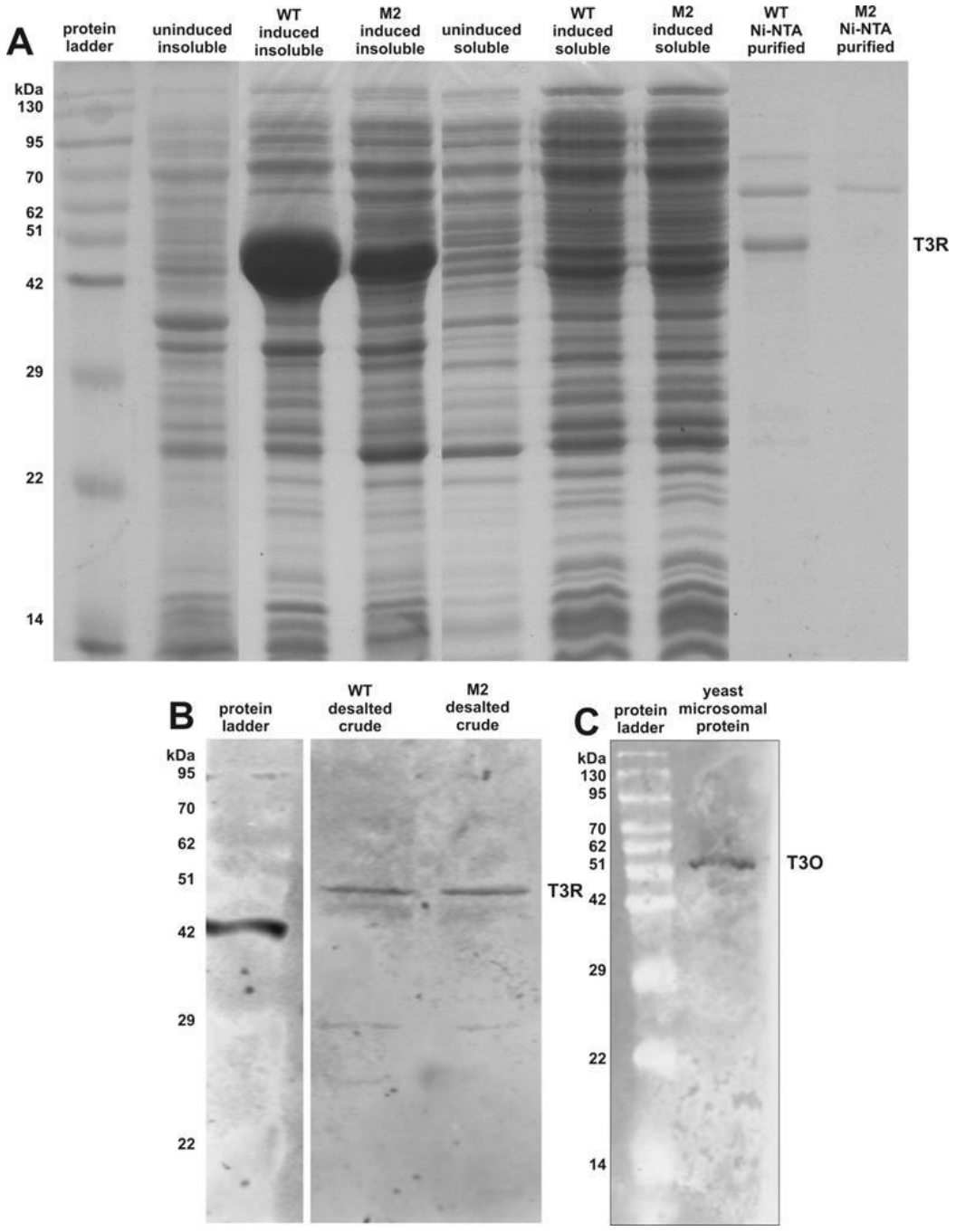


Figure 16: Production and affinity purification of soluble recombinant WT- and M2-T3R protein in *E. coli*. (**A**) SDS-PAGE of 5 µg insoluble, 10 µg soluble, and 5 µg Ni-NTA affinity purified protein extracts from uninduced and induced recombinant *E. coli* expressing 6xHis-tagged *WT-T3R* and *M2-T3R*. A band corresponding to T3R was purified from *E. coli* expressing *WT-T3R* but not from *E. coli* expressing *M2-T3R*. (**B**) Western blot of SDS-PAGE showing similar levels of 6xHis-tagged T3R in 10 µg of desalted crude protein extracts from recombinant *E. coli* expressing *WT-T3R* or *M2-T3R*. (**C**) Western blot of SDS-PAGE showing T3O in 50 µg recombinant yeast microsomal protein.

higher accumulation of the 2,3-epoxides (Figure 17I-J), compared to protein extracts from *E. coli* expressing *WT-T3R* (Figure 17E-F). More detailed comparisons of 2,3-dihydro-3-hydroxy product formation by both extracts (Figure 17K) indicated that protein from *E. coli* expressing *M2-T3R* exhibits ~5% T3R activity ($7.06x10^{-5} \pm 0.75x10^{-5}$ and $1.21x10^{-4} \pm 0.15x10^{-4}$ pkat/mg total protein) compared to protein from *E. coli* expressing *WT-T3R* ($1.25x10^{-3} \pm 0.23x10^{-3}$ and $2.29x10^{-3} \pm 0.51x10^{-3}$ pkat/mg total protein) towards tabersonine and 16-methoxytabersonine, respectively.

C.6. Line M2-1865 displays a recessive MIA phenotype that exhibits standard Mendelian inheritance.

Line M2-1865 was confirmed to be homozygous by analysis of MIA profiles of six offspring in the M3 generation, which were obtained by self-fertilization of the M2 generation and which accumulated similar levels of 2,3-epoxides and vindoline as the M2 generation (Supplementary Figure 1). Reciprocal crossings were then made between WT and M2-1865 plants in order to assess the pattern of inheritance (as described in B.4.). MIA analysis of the F1 generation of these crosses indicated that vindoline and vindorosine levels returned to those found in the WT (Figure 18A), while the 2,3-epoxides found on the leaf surface of M2-1865 plants disappeared from all fifteen offspring in the F1 generation (Supplementary Figure 2). 10-20 seeds were collected and germinated from each self-fertilized F1 plant to generate the F2 generation, from which 60 6-week-old seedlings were randomly selected for MIA analysis and quantification. 16 homozygous mutants were identified by the presence of significant levels of 2,3-epoxides on the leaf surface, while the remaining 44 offspring contained no significant levels of epoxides, giving an approximately 3:1 ratio of WT:M2-1865 MIA phenotypes (Figure

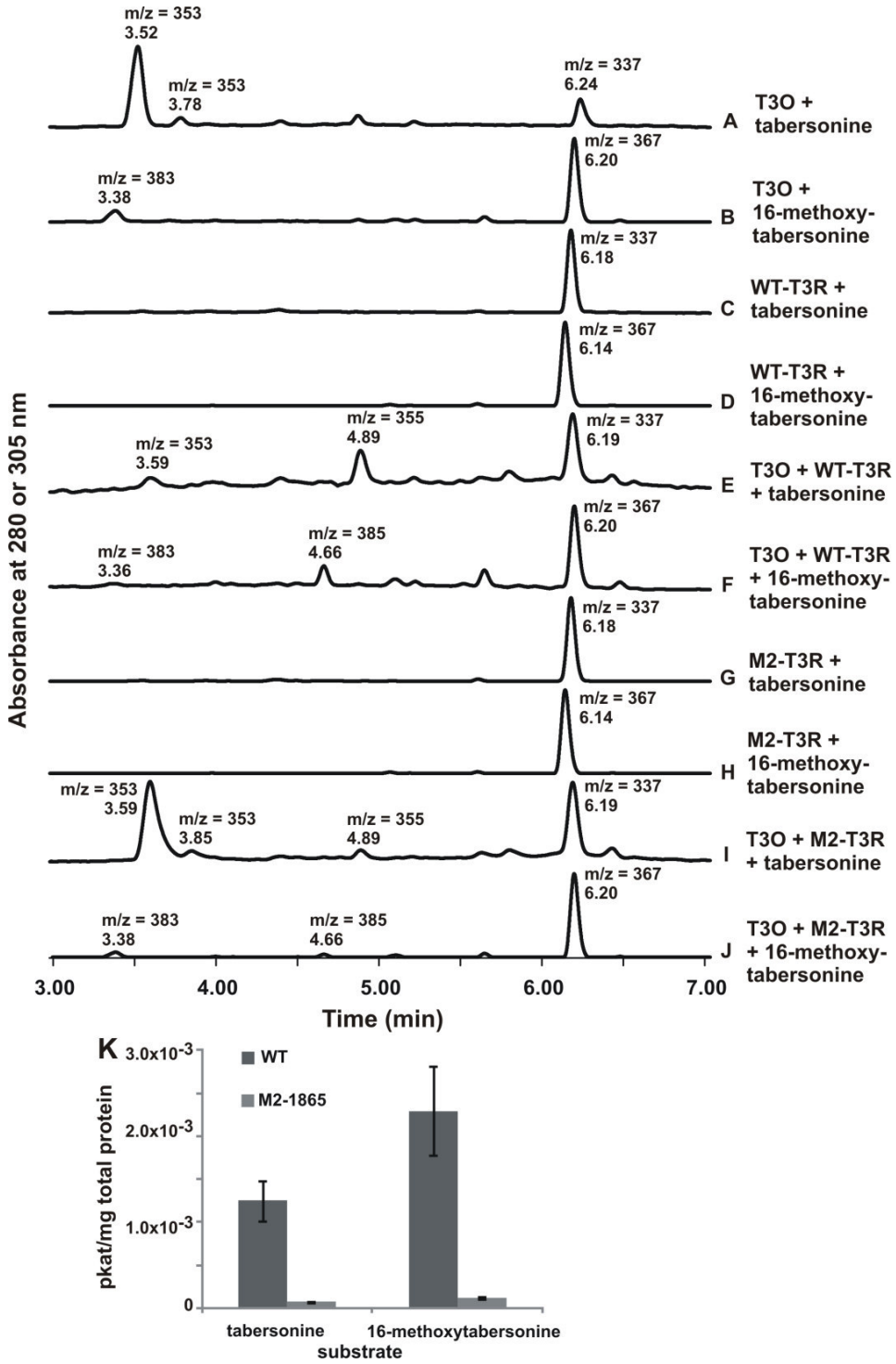


Figure 17: Tabersonine and 16-methoxytabersonine are effectively converted to 2,3-dihydro-3-hydroxy derivatives in assays containing recombinant WT-T3R but very poorly in assays containing recombinant M2-T3R. (**A-J**) LC profiles of assays containing 1 mM NADPH, +/- 50 µg recombinant yeast microsomal T3O, +/- recombinant bacterial protein from *E. coli* expressing *WT-T3R* (50 µg) or *M2-T3R* (500 µg), and 25 µM tabersonine or 16-methoxytabersonine as substrate. Tabersonine and 16-

methoxytabersonine reaction products were detected at 280 nm and 305 nm, respectively [tabersonine (RT 6.18-6.24 min, m/z 337), 16-methoxytabersonine (RT 6.14-6.20 min, m/z 367), 2,3-epoxytabersonine (RT 3.52 and 3.78 min, m/z 353), 16-methoxytabersonine (RT 3.38 min, m/z 383), 2,3-dihydro-3-hydroxytabersonine (RT 4.89 min, m/z 355), 16-methoxy-2,3dihydro-3-hydroxytabersonine (RT 4.66 min, m/z 385)]. (**K**) 2,3-dihydro-3-hydroxy product formation by T3O and bacterial recombinant proteins from *E. coli* expressing *WT-T3R* or *M2-T3R* in triplicate assays.

18B). The F2 generation exhibited a wide range in MIA content as seen in the levels of vindoline and 2,3-epoxides. The 16 plants classified as homozygous mutants contained an average of 324.80 ± 157.79 µg/gfw of 2,3-epoxides, with minimum and maximum levels of 99.65 and 696.87 µg/gfw, respectively. All homozygous mutants contained reduced levels of vindoline, with an average of 59.66 ± 28.14 µg/gfw and minimum and maximum levels of 10.88 and 114.37 µg/gfw, respectively. The 44 plants classified as homozygous WT and heterozygous contained no significant levels of 2,3-epoxides but displayed a large variation in vindoline levels, with an average of 197.67 ± 82.43 µg/gfw and minimum and maximum levels of 74.77 and 384.79 µg/gfw, respectively.

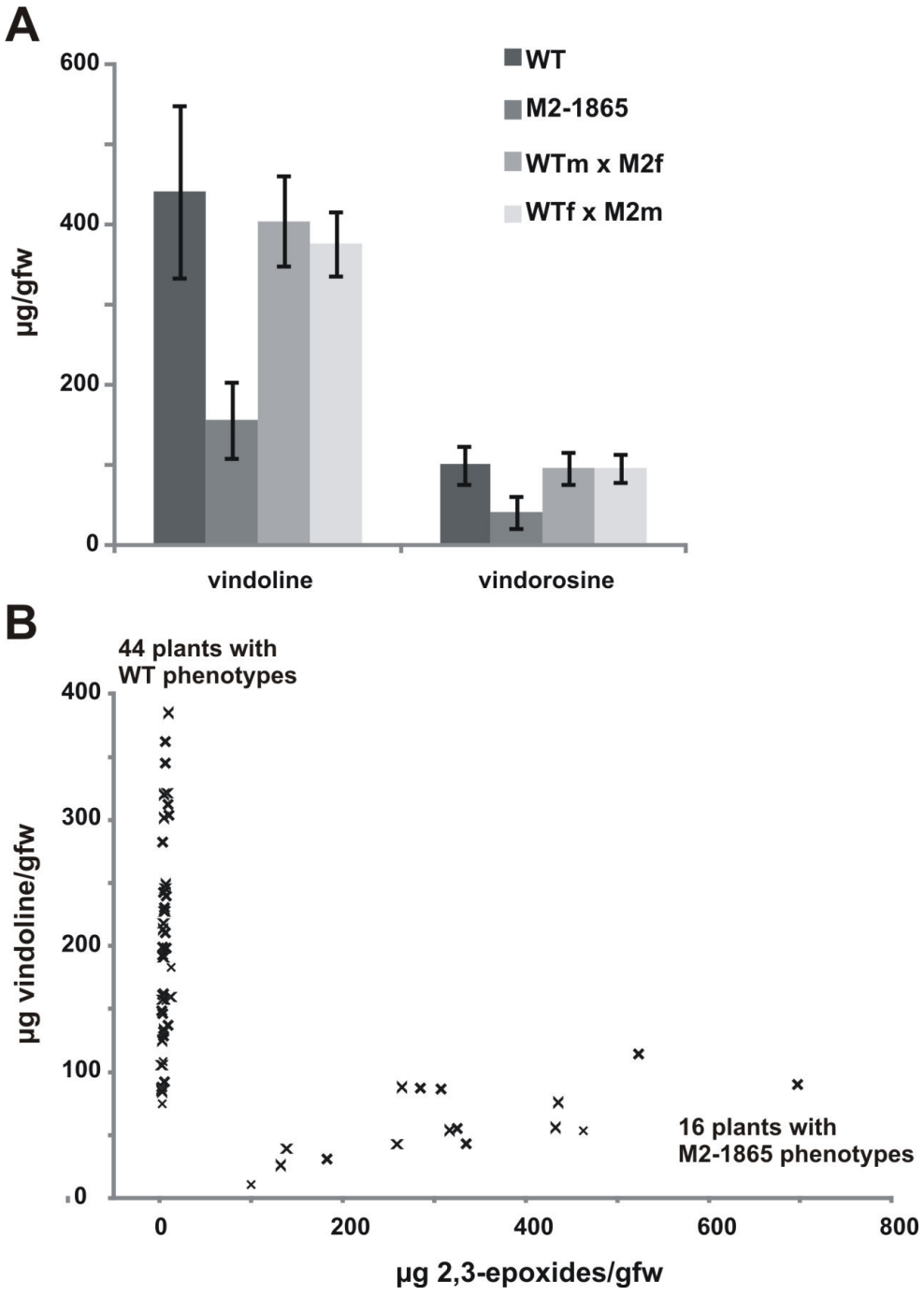


Figure 18: Reciprocal crossings between WT and M2-1865 plant lines show that the mutant phenotype is recessive and exhibits standard Mendelian inheritance. (**A**) The F1 generation plants (Wtm x M2f n = 11; WTf x M2m n = 4) show similar levels of vindoline and vindorosine in LP2 compared to WT. (**B**) Phenotype distribution of 60 F2 generation plants using levels of tabersonine 2,3-epoxide derivatives and vindoline contained in LP2. Homozygous mutant offspring (n = 16) were identified by the presence of 2,3-epoxides on the leaf surface while heterozygous and/or homozygous WT offspring (n = 44) contained no significant levels of 2,3-epoxides.

D. Discussion

D.1. A single gene mutation affecting enzyme activity of T3R is responsible for the M2-1865 phenotype.

The EMS-mutagenized *C. roseus* M2-1865 plant line has an altered MIA profile: it contains reduced levels of vindoline, vindorosine, and vindoline-containing dimers within the leaves and accumulates 2,3-epoxide derivatives of tabersonine on the leaf surface. The altered MIA profile in the M2-1865 line could have resulted due to one of many different types of mutations that may directly or indirectly affect biosynthesis or expression in the vindorosine/vindoline pathway. For example, a mutation in a regulatory gene could affect expression of several genes in a signal transduction pathway that in turn may suppress the expression of the key *T3R* step in vindoline biosynthesis. Since M2-1865 did not display any differences in morphology, there were likely no mutations in central regulators of plant development, although mutations could have occurred in regulatory elements specifically involved in MIA biosynthesis. However, the results obtained in this study indicate that the M2-1865 phenotype is the result of a single functional gene mutation causing reduced enzymatic activity.

Vindoline and catharanthine are both derived from the same MIA precursor (Dewick, 2009; El-Sayed et al., 2004; De Luca et al., 2012). Since M2-1865 contains similar levels of catharanthine to WT, only the tabersonine to vindorosine/vindoline portion of the pathway has been affected. This is further supported by the accumulation of the tabersonine 2,3-epoxide derivatives produced by T3O in the absence of T3R (Qu et al., 2015), indicating either reduced *T3R* gene expression or enzymatic activity. Since none of the MIA biosynthetic genes, including *T3R*, exhibited changes in expression

between M2-1865 and WT plants, alterations to either functional gene expression or MIA regulatory elements are therefore not responsible for the M2-1865 phenotype. Cloning and sequencing of the WT and M2-1865 *T3R* genes indicated two G/C to A/T pair nucleotide mutations, as expected of EMS (Greene et al., 2003; Haughn and Somerville, 1987), resulting in one silent and one missense mutation to produce the H189Y amino acid substitution. Coupled *in vitro* assays with recombinant T3O and protein extracts of *E. coli* expressing *WT-T3R* or *M2-T3R* indicated ~95% reduction in T3R activity towards tabersonine and 16-methoxytabersonine of the mutant protein compared to WT, supporting its role in the M2-1865 phenotype. In addition, the mutant phenotype displays standard Mendelian inheritance. The mutation is recessive, as indicated by the WT phenotype displayed by the F1 generation, and a single functional copy of the *T3R* gene is thus sufficient to restore vindoline biosynthesis. The F2 generation displays an approximately 3:1 ratio of inheritance, with 44 plants displaying the WT phenotype and 16 plants displaying the mutant phenotype. This 3:1 ratio is the expected pattern of single-gene segregation in Mendelian inheritance (Hartwell et al., 2011). Although the 44 plants that did not accumulate 2,3-epoxides exhibited a wide range of vindoline levels, the F2 generation did not display a pattern that would be expected for multi-gene inheritance; for example, two-gene segregation would display a 9:3:3:1 phenotype ratio (Hartwell et al., 2011), where approximately 11 plants would be expected to display moderate levels of both vindoline and 2,3-epoxides. As this is not the case, and because levels of 2,3-epoxides among the 16 homozygous mutants also vary widely, this range is likely the result of natural variation. Previous studies have indicated that levels of MIAs, particularly catharanthine, vindoline, and 3',4'-anhydrovinblastine, can vary widely

among individual mature plants and especially seedlings (Costa et al., 2008; Magnotta et al., 2006), which supports this conclusion. The H189Y mutation is therefore solely responsible for the M2-1865 phenotype due to highly-reduced T3R activity as a result of changes in structure, catalysis, or binding of substrate/cofactors.

D.2. The H189Y mutation may affect cofactor binding.

T3R is closely-related to zinc-dependent alcohol dehydrogenases (ADHs) of the medium-chain dehydrogenase/reductase (MDR) superfamily (Qu et al., 2015). Specifically, it is part of the cinnamyl/sinapyl alcohol dehydrogenase (CAD/SAD) family, which also includes 10HGO but is primarily made up of genes involved in monolignol biosynthesis (Guo et al., 2010). A phylogenetic analysis including 71 putative or probable CAD/SADs from various plants (Figure 19) indicates that T3R and 10HGO, which share 61% amino acid sequence similarity, are grouped with the Class II A-SADs (Guo et al., 2010). This group includes *Nicotiana tabacum* NtaSAD, *Populus tremuloides* PtrSAD, and *Arabidopsis thaliana* AthCAD2-3,6-9. Other classes include the Class I *bona fide* CADs as exemplified by AthCAD4-5, *Pinus taeda* PtaCAD, *Eucalyptus gunnii* EguCAD2, *Brachypodium distachyon* BdiCAD5, and *Oryza sativa* OsaCAD2; the Class II G-SADs such as *Metasequoia glyptostroboides* MglSAD, *Abies holophylla* AhoSAD1-2, and *Tsuga dumosa* TduSAD1-2; and the Class IIIs such as AthCAD1 and OsaCAD1,4 (Guo et al., 2010). *Eriobotrya japonica* EjaCAD1-2, *Camellia sinensis* CsiCAD2, *Medicado truncatula* MtrCAD2, EguCAD1, and NtaCAD1 form an outgroup with low sequence similarity to the other CAD/SADs. Characterizations of these proteins have identified them as atypical CADs belonging to the short-chain dehydrogenase/reductase superfamily (Shan et al., 2008; Deng et al., 2013; Pan et al., 2014; Goffner et al., 1992;

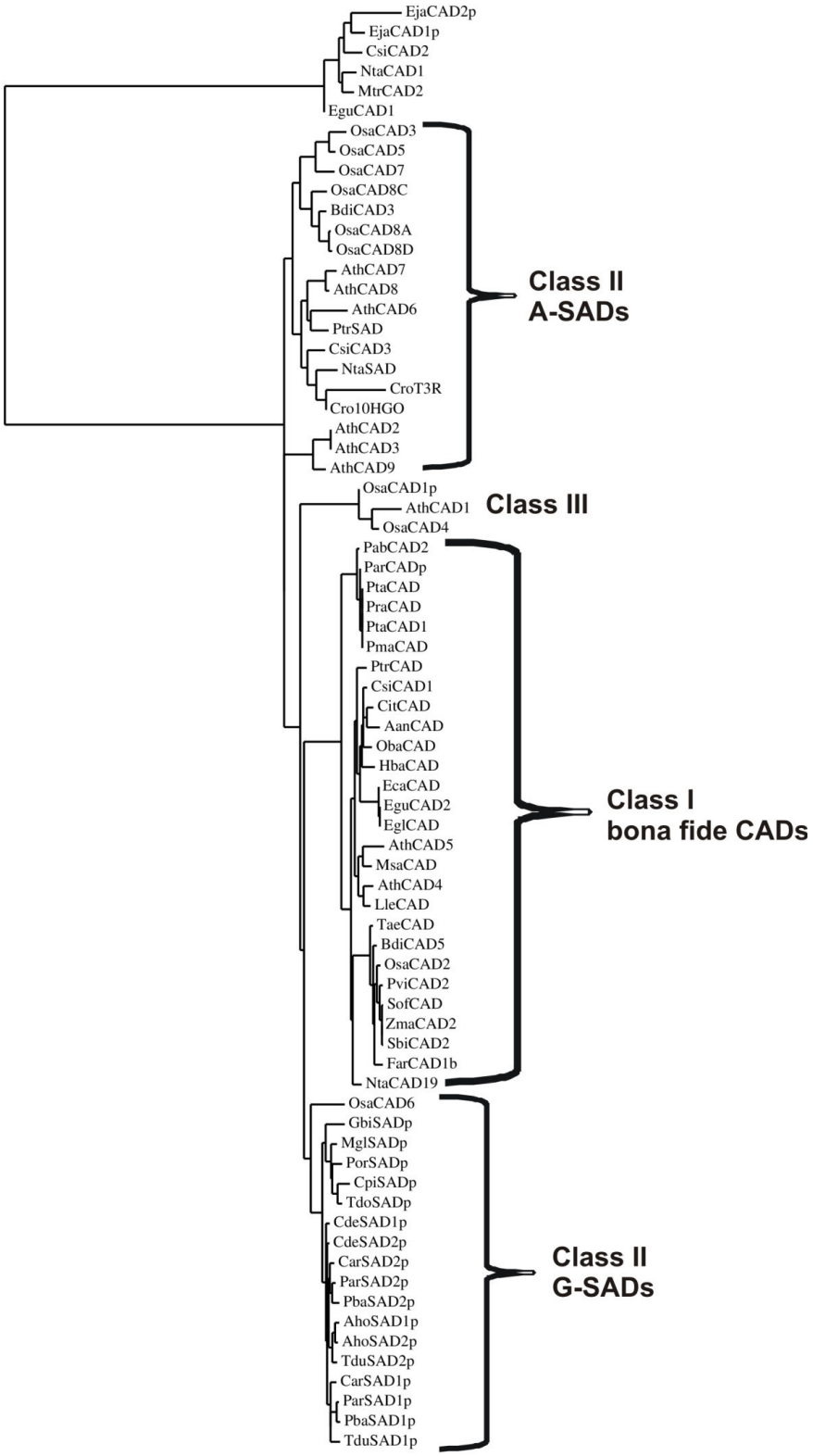


Figure 19: Phylogenetic analysis of *C. roseus* T3R with putative and probable cinnamyl/sinapyl alcohol dehydrogenases (CAD/SADs) shows that T3R is closely-related to Class II A-SADs such as *P. tremuloides* SAD. Species and accessions are listed in Supplementary Table 2.

Goffner et al., 1998; Damiani et al., 2005). CAD/SADs reduce hydroxycinnamyl aldehydes such as coniferaldehyde and sinapaldehyde into their corresponding alcohols (Gross et al., 1973; Mansell et al., 1974), while 10HGO catalyzes the oxidation of 10-hydroxygeraniol and other primary alcohols to their corresponding aldehydes and can also catalyze the reverse reductions (Ikeda et al., 1991; Miettinen et al., 2014). The function of T3R as suggested by Qu et al. (2015) is somewhat different, as it is only proposed to reduce the unstable intermediate produced by T3O and does not significantly display reductase activity towards the 2,3-epoxide side products.

Since crystal structures and characterizations are available of PtrSAD [PDB: 1YQD] (Bomati and Noel, 2005) and AthCAD5 [PDB: 2CF5] (Youn et al., 2006), an amino acid alignment between these two proteins as well as T3R and 10HGO was performed in order to investigate potential consequences of the H189Y mutation (Figure 20). T3R exhibits 59% and 44% amino acid sequence identity with PtrSAD and AthCAD5, respectively, while 10HGO exhibits 78% and 52% similarity. PtrSAD and AthCAD5 consist of two domains: the nucleotide-binding domain (Figure 20, highlighted in yellow; residues 170-302 in PtrSAD and residues 163-301 in AthCAD5) and the catalytic or substrate-binding domain (Figure 20, highlighted in red; residues 1-169 and 303-362 in PtrSAD and residues 1-162 and 302-357 in AthCAD5). Both proteins coordinate two Zn^{2+} ions, one catalytic and one structural, as characteristic of zinc-dependent MDRs (Bomati and Noel, 2005; Youn et al., 2006). The catalytic Zn^{2+} is coordinated by residues C50, H70, and C166 in PtrSAD and by residues C47, H69, E70, and C163 in AthCAD5; tetrahedral coordination is usually completed via a water molecule in PtrSAD and other MDRs rather than E70 (Bomati and Noel, 2005; Colonna-

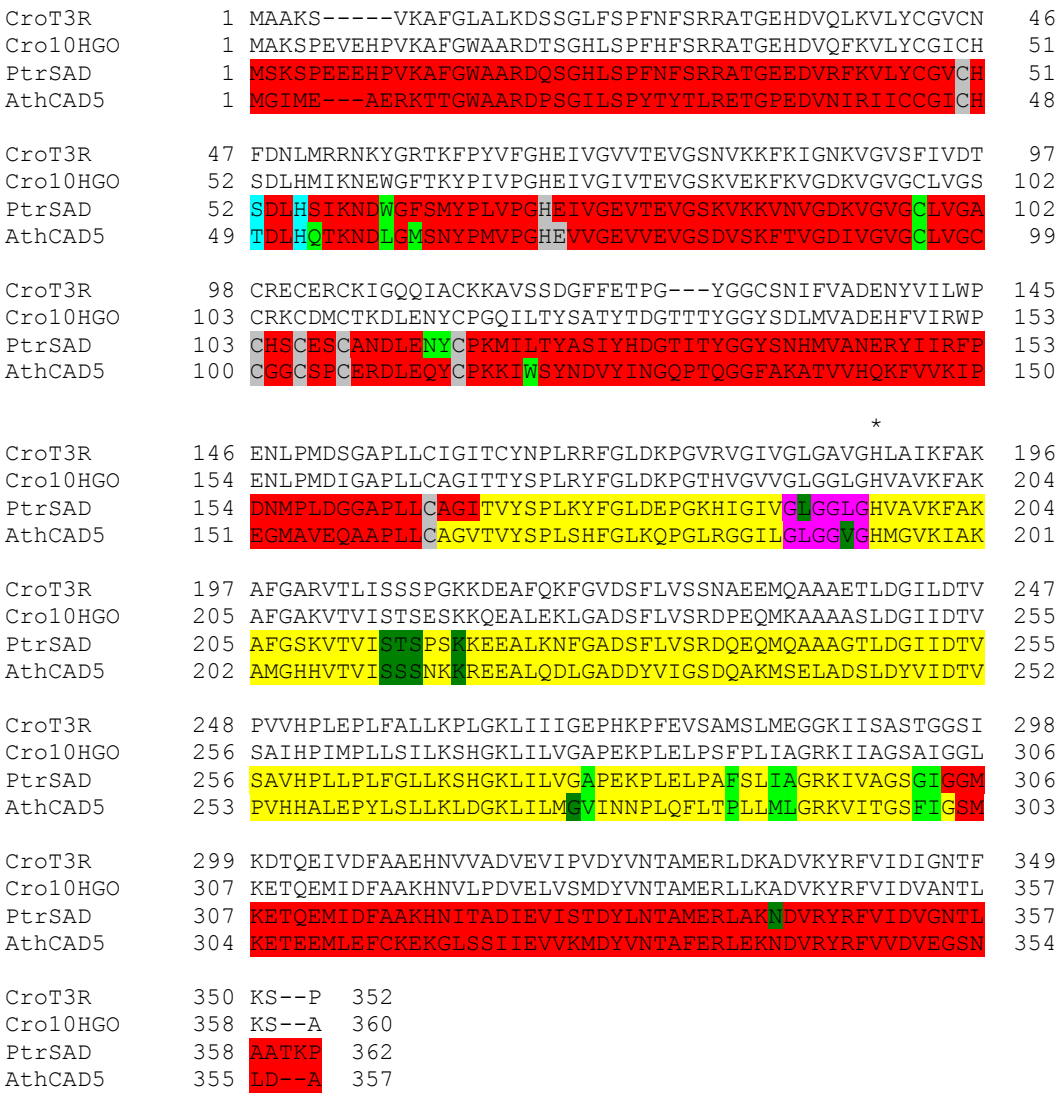


Figure 20: Amino acid alignment of *C. roseus* T3R and 10HGO, *P. tremuloides* SAD, and *A. thaliana* CAD5 shows that T3R contains conserved residues involved in cofactor binding and Zn^{2+} coordination but does not contain conserved catalytic or active site residues. Domains and residues identified by crystal structures (Bomati and Noel, 2005; Youn et al., 2006) are highlighted as follows: catalytic or substrate-binding domain (red), nucleotide-binding domain (yellow), Zn^{2+}-coordinating residues (gray), catalytic residues (turquoise), active site residues (light green), NADPH-stabilizing residues (dark green), and the glycine-rich motif (fuschia). H189 of T3R, which is substituted for tyrosine in mutant T3R, is indicated with an asterisk (*). PDB IDs: P. tremuloides SAD [1YQD], A. thaliana CAD [2CF5].

Cesari et al., 1986; Eklund and Brändén, 1987). The structural Zn^{2+} is coordinated by C103, C106, C109, and C117 in PtrSAD and the same corresponding cysteines in AthCAD5. All of the Zn^{2+}-coordinating residues (Figure 20, highlighted in gray) are conserved in both T3R and 10HGO, supporting their inclusion in the zinc-dependent ADH family.

Many zinc-dependent ADHs are proposed to utilize a proton shuttle in the active site in order to facilitate hydride transfer to or from the nicotinamide cofactor (Clarke and Daffron, 1998). The highly-conserved catalytic residues (Figure 20, highlighted in blue) for this shuttle in typical CAD/SADs are a serine or threonine (S52 in PtrSAD, T49 in AthCAD5) and a histidine (H55 and H52 in PtrSAD and AthCAD5, respectively) (Bomati and Noel, 2005; Youn et al., 2006). Notably, these catalytic residues are conserved in 10HGO but not in T3R, which instead contains phenylalanine and leucine at these positions. Of the typical CAD/SADs used in the phylogenetic analysis in Figure 19, these catalytic serine/threonine and histidine residues are both completely conserved in all but PtaCAD1,2, AthCAD1,7, NtaSAD, OsaCAD1,4, *Picea abies* PabCAD2, *Pinus armandii* ParCAD, *Pinus radiata* PraCAD, and *Pinus massoniana* PmaCAD, most of which have substitutions to nonpolar residues like alanine, valine, or isoleucine. However, only AthCAD1,7 and NtaSAD functions have been investigated. Kim et al. (2007) observed that AthCAD7 displayed minor activity and similar expression relative to *bona fide* CAD, whereas AthCAD1 did not exhibit CAD activity or expression patterns and has an unknown function. Similarly, NtaSAD was identified as a SAD-like protein not likely to display functional CAD/SAD activity by Barakate et al. (2011), who identified other functional putative SADs in *N. tabacum*. These proteins, like T3R, may

therefore display altered activities and catalytic mechanisms compared to typical CAD/SADs.

Residues lining the active site (Figure 20, highlighted in light green) are important in determining substrate specificity even if they are not directly involved in the catalytic mechanism. In AthCAD5, the active site is made up of Q53, L58, M60, C95, W119, V276, P286, M289, L290, F299, and I300, in addition to the catalytic and Zn^{2+}-coordinating residues. Youn et al. (2006) observed that nine of these residues are completely conserved in all *bona fide* CADs and that M289 is partially-conserved, sometimes substituted as isoleucine. The active site residues of PtrSAD are W61, C98, N115, Y116, A279, F289, I292, A293, G302, and I303 (Bomati and Noel, 2005). AthCAD5 and PtrSAD thus share only two active site residues, which likely determines their different substrate specificities: Youn et al. (2006) noted that W119 and F299 reduced the size of the binding pocket and entrance in AthCAD5 compared to PtrSAD, which contains leucine and glycine at these positions, while the substitution from V276 in AthCAD5 to A293 in PtrSAD further increases the size of the active site and results in broader substrate specificity by PtrSAD. Similarly, C95 of AthCAD5 (C98 of PtrSAD), which corresponds to F93 in horse liver ADH, has been proposed to be involved in determining substrate specificity in mammalian ADHs and in *Sulfolobus solfataricus* ADH based on steric hindrance (Esposito et al., 2003; Stone et al., 1989). All but one of the PtrSAD active site residues are conserved in 10HGO, which substitutes an alanine for G302 and which has been reported to display fairly broad substrate specificity (Miettinen et al., 2014). Conversely, none of the active site residues of PtrSAD are conserved in T3R with the exception of M284, the partially-conserved residue which corresponds to M289

in AthCAD5 and I292 in PtrSAD. Based on these observations, T3R may exhibit a different active site topology and catalytic mechanism than more typical CAD/SADs, resulting in its novel function for involvement in vindoline biosynthesis.

Although these observations indicate how T3R may have gained different functionality from CAD/SADs, they do not explain the consequences of the H189Y mutation, as the histidines corresponding to H189 in PtrSAD and AthCAD5 were not assigned any role in substrate binding or catalysis. Therefore, cofactor binding was subsequently investigated. Stabilization of the nicotinamide cofactor involves several residues (Figure 20, highlighted in dark green) located in three flexible loops (Min et al., 2003; Youn et al., 2005). In PtrSAD, S214, T215, S216, and K219 of the first loop, as well as N343 interact with the 2'-phosphate of $NADP^+$/NADPH; Bomati and Noel (2005) also state that N192 is involved, but this position actually contains a leucine rather than asparagine. AthCAD5 substitutes T215 with a serine (S213) but otherwise contains the same conserved residues as PtrSAD. Youn et al. (2006) also determined that the amide group of $NADP^+$ interacts with the backbone of G275, a residue in the third loop. With the exception of N343, these residues are all conserved in both T3R and 10HGO, as expected of $NAD(P)^+$/NAD(P)H-dependent proteins. The serine/threonine corresponding to PtrSAD T215 and AthCAD5 S213 is believed to be integral in determining cofactor specificity for $NADP^+$/NADPH, as NAD^+/NADH-dependent ADHs contain an aspartic acid residue at this position, which would cause charge repulsion with the phosphate group of $NADP^+$/NADPH (Barbosa Pereira et al., 2001; Powell et al., 2000; Lauvergeat et al., 1995; Bomati and Noel, 2005). Of particular interest is the second loop, a tight turn between the ßA and αB secondary structures, which contains a conserved glycine-rich

(Gx(x)GxxG) motif (Figure 20, highlighted in fuschia; residues 191-196 in PtrSAD and 188-193 in AthCAD5). This motif is typical of $NAD(P)^+$/NAD(P)H-dependent ADHs and is directly involved in cofactor binding via stabilization of the pyrophosphate oxygens by the helix macrodipole (Min et al., 2003; Youn et al., 2005; Barbosa Pereira et al., 2001; Hol et al., 1978; Stone et al., 1989; Eklund et al., 1984). Residue 192 of PtrSAD and V192 of AthCAD5 within this motif were both determined to be directly involved in hydrogen bonding to the cofactor (Bomati and Noel, 2005; Youn et al., 2006). H189 of T3R (Figure 20, asterisk) is directly adjacent to this glycine-rich motif and is conserved in all typical CAD/SADs used in the phylogenetic analysis. Although it does not appear to have a direct role in cofactor binding, this histidine likely contributes to the local polar environment and/or secondary structure. Substitution of this residue to tyrosine in the mutant T3R would therefore affect a vital portion of the protein required for cofactor binding. Studies on a pyrophosphate-stabilizing residue in NAD^+-dependent human $ß_1ß_1$ alcohol dehydrogenase (R47) indicated that substitutions of this arginine residue to histidine, lysine, glutamine, or glycine resulted in lower affinity/increased K_M for NAD^+ as well as poorer dissociation/increased K_d for NADH (Hurley et al., 1990). Dissociation of NADH from NAD^+-dependent ADHs is believed to be rate-limiting (Bosron et al., 1983; Dalziel and Dickinson, 1966), so poorer cofactor binding and/or dissociation, as well as structural rearrangements, could be possible consequences of the H189Y mutation in T3R. Detailed modelling of the T3R active site and NADPH-binding domain would be beneficial in determining the residues involved in substrate binding and catalysis and in further investigating the consequences of the H189Y mutation on cofactor interactions.

D.3. M2-1865 supports the proposed T3O-T3R mechanism in vindoline biosynthesis.

Two different pathways have been proposed in regards to T3O activity in *C. roseus*. Biotransformation experiments by Kellner et al. (2015) suggested that T3O produces a 2,3-epoxide that spontaneously forms an imine alcohol which, in the absence of the reductase, undergoes a structural rearrangement to a vincamine-type compound. However, VIGS experiments and the characterization of both T3O and T3R by Qu et al. (2015) indicated that the 2,3-epoxide is a stable by-product formed in the absence of T3R, as these epoxides accumulate on the leaf surface of *T3R*-silenced plants and in enzymatic assays with recombinant T3O. M2-1865 supports the findings by Qu et al. (2015). This plant line is impaired in T3R activity and accumulates the same 2,3-epoxides on the leaf surface, rather than the by-products observed by Kellner et al. (2015). The rearrangements observed by Kellner et al. may therefore be an artifact due to the assay conditions or MIA extraction procedures used in the study.

D.4. EMS mutagenesis and screening are valuable tools for targeting secondary metabolic pathways.

The M2-1865 phenotype was generated as a result of chemical mutagenesis with EMS. EMS mutagenesis has been widely employed in model plant systems and/or important crop species (Greene et al., 2003; Botticella et al., 2011; Cooper et al., 2008; Rowland, 1991; Till et al., 2007; Xin et al., 2008; Okabe et al., 2011; Zhu et al., 2012) with only few studies in *C. roseus* (Verma and Singh, 2010; Kulkarni et al., 1999; Rai et al., 2003; Kumari et al., 2013), primarily with the aim of investigating genes involved in plant development and important agronomical traits. The EMS-induced high-ajmalicine-accumulating *C. roseus* line (Czerwinski, 2011; Thamm, 2014) demonstrated the viability

of utilizing EMS mutagenesis for discovery of genes and regulatory factors involved in MIA biosynthesis. This study further supports the usefulness of the EMS method in gene discovery and characterization. Whereas the high-ajmalicine phenotype was the result of changes in expression in several regulatory elements involved in MIA biosynthesis (Thamm, 2014), the M2-1865 phenotype was generated due to a single amino acid substitution in a functional gene. *T3R* was first identified by Qu et al. (2015) using a bioinformatics approach and VIGS. As such, M2-1865 did not result in the identification of a novel gene, but nevertheless correlates the *in vitro* findings by Qu et al. with the *in vivo* function of T3O/T3R. Taken together, these results demonstrate the viability of integrated approaches combining mutagenesis and bioinformatics for gene identification and characterization.

E. Future Directions

In the short term, this work will continue in the following directions: 1) purification of recombinant M2-T3R via alternative methods, such as Co-NTA or Strep-tag affinity purification, for complete biochemical characterization; 2) site-directed mutagenesis of H189 to other residues such as glycine and arginine to assess its role in T3R and to assess the role of the H189Y mutation in the reduction of T3R activity; and 2) modelling of WT-T3R to determine residues involved in the active site/catalysis and to further assess the consequences of the H189Y mutation, which may affect NADPH-binding.

Pathway engineering is a long-term goal of MIA gene characterization. To this end, M2-1865 supports the roles of T3O and T3R in vindoline biosynthesis in *C. roseus*. With the tabersonine to vindoline pathway fully identified and proof-of concept studies demonstrating the potential for genetically engineering MIA biosynthesis in yeast (Qu et al., 2015), the next step towards pathway engineering is the identification of genes involved in converting strictosidine to tabersonine and catharanthine. As exemplified in this work, EMS mutagenesis combined with bioinformatics would be a valuable tool in this endeavour.

Lastly, the M2-1865 plant line will provide a valuable resource in studying the spatial separation of the MIA pathway in *C. roseus*. The 2,3-epoxides that accumulate on the leaf surface are indicative of MIA transport to the leaf surface, but the catharanthine transporter does not transport tabersonine (Yu and De Luca, 2013) and as such is not expected to transport tabersonine derivatives. Other intermediates of the epidermal portion of the pathway likewise appear on the leaf surface in VIGS experiments (Qu et

al., 2015) but unlike transient VIGS phenotypes, the M2-1865 phenotype is the result of a stable mutation. This mutant line will therefore be a useful tool for VIGS experiments in assessing other candidate transporters.

F. Summary and Conclusions

The EMS-induced M2-1865 phenotype, which has reduced levels of vindoline and accumulates 2,3-epoxide derivatives of tabersonine on the leaf surface, was determined to be the result of a single amino acid substitution in the functional MIA biosynthetic gene, *T3R*. This conclusion was drawn based on the following results:

1) The MIAs that accumulate on the leaf surface of M2-1865 showed identical mass and retention time as the 2,3-epoxides produced by T3O from tabersonine and 16-methoxytabersonine in the absence of T3R (Qu et al., 2015).

2) *T3R* gene expression was unchanged between WT and M2-1865 plants, and expression of other MIA biosynthetic genes was also similar, suggesting no mutations affecting regulatory elements.

3) Sequencing of cloned *T3R* indicated a single missense mutation in the mutant gene resulting in an amino acid substitution, H189Y, between the WT and M2 proteins. Subsequent *in vitro* assays with the recombinant bacterial proteins indicated that mutant T3R activity was reduced by ~95% compared to WT-T3R using either tabersonine or 16-methoxytabersonine as substrate.

4) Reciprocal crossings with WT plants showed that the M2-1865 phenotype was recessive and inherited in a 3:1 ratio in the F2 generation, exhibiting standard Mendelian single-gene inheritance.

In addition, comparison of amino acid sequences of T3R and characterized alcohol dehydrogenases in the same protein family suggested that the H189Y substitution may affect NADPH binding, as this residue is directly adjacent to a conserved glycine-rich motif involved in cofactor stabilization (Min et al., 2003; Youn et al., 2005; Barbosa

Pereira et al., 2001; Hol et al., 1978; Stone et al., 1989; Eklund et al., 1984).

Overall, this study highlights the usefulness of EMS mutagenesis in targeting secondary metabolic pathways for novel gene discovery and in correlating *in vitro* studies with *in vivo* observations.

G. References

Aerts, R.J., De Luca, V. (1992). Phytochrome is involved in the light-regulation of vindoline biosynthesis in *Catharanthus*. *Plant Physiol* **100**:1029-1032.

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

Anisimova, M., Gascuel, O. (2006). Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. *Syst Biol* **55**:539-552.

Arvy, M.P., Imbault, N., Naudascher, F., Thiersault, M., Doireau, P. (1994). 2,4-D and alkaloid accumulation in periwinkle cell suspensions. *Biochimie* **76**:410-416.

Asada, K., Salim, V., Masada-Atsumi, S., Edmunds, E., Nagatoshi, M., Terasaka, K., Mizukami, H., De Luca, V. (2013). A 7-deoxyloganetic acid glucosyltransferase contributes a key step in secologanin biosynthesis in Madagascar periwinkle. *Plant Cell* **25**:4123-4134.

Balsevich, J., De Luca, V., Kurz, W.G.W. (1986). Altered alkaloid pattern in dark grown seedlings of *Catharanthus roseus*. The isolation of characterization of desacetoxyvindoline: a novel indole alkaloid and proposed precursor of vindoline. *Heterocycles* **24**:2415-2421.

Balsevich, J., Bishop, G. (1989). Distribution of catharanthine, vindoline and 3'4'-anhydrovinblastine in the aerial parts of some *Catharanthus roseus* plants and the significance thereof in relation to alkaloid production in cultured cells. In *Primary and Secondary Metabolism of Plant Cell Cultures*. Kurz, W.G.W (Ed). Berlin: Springer-Verlag, pp 149-153.

Barakate, A., Stephens, J., Goldie, A., Hunter, W.N., Marshall, D., Hancock, R.D., Lapierre, C., Morreel, K., Boerjan, W., Halpin, C. (2011). Syringyl lignin is unaltered by severe sinapyl alcohol dehydrogenase suppression in tobacco. *Plant Cell* **23**:4492-4506.

Barbosa Pereira, P.J., Macedo-Ribeiro, S., Párraga, A., Pérez-Luque, R., Cunningham, O., Darcy, K., Mantle, T.J., Coll, M. (2001). Structure of human biliverdin IXß reductase, an early fetal bilirubin IXß producing enzyme. *Nat Struct Biol* **8**:215-220.

Battersby, A.R., Brown, R.T., Knight, J.A., Martin, J.A., Plunkett, A.O. (1966). Biosynthesis of the indole alkaloids from a monoterpene. *Chem Commun* 346-347.

Battersby, A.R., Kapil, R.S., Martin, J.A., Mo, L. (1968). Loganin as precursor of the indole alkaloids. *Chem Commun* 133-134.

Battersby, A.R., Burnett, A.R., Parsons, P.G. (1969). Alkaloid biosynthesis, part XV. Partial synthesis and isolation of vincoside and isovincoside: biosynthesis of the three major classes of indole alkaloids from vincoside. *J Chem Soc (C)* 1193-1200.

Battersby, A.R., Burnett, A.R., Parsons, P.G. (1970a). Preparation and isolation of

deoxyloganin: its role as precursor of loganin and the indole alkaloids. *Chem Commun* 826-827.

Battersby, A.R., Brown, S.H., Payne, T.G. (1970b). Biosynthesis of loganin and the indole alkaloids from hydroxygeraniol-hydroxynerol. *Chem Commun* 827-828.

Baxter, C., Slaytor, M. (1972). Partial purification and some properties of tryptophan decarboxylase from *Phalaris tuberosa*. *Phytochemistry* **11**:2763-2766.

Baxter, R.L., Dorschel, C.A., Lee, S-L., Scott, A.I. (1979). Biosynthesis of the antitumour catharanthus alkaloids. Conversion of anhydrovinblastine into vinblastine. *Chem Commun* 257.

Berlin, J. (1984). Plant cell cultures – a future source of natural products? *Endeavour* **8**:5-8.

Besseau, S., Kellner, F., Lanoue, A., Thamm, A.M.K., Salim, V., Schneider, B., Geu-Flores, F., Höfer, R., Guirimand, G., Guihur, A., Oudin, A., Glevarec, G., Foureau, E., Papon, N., Clastre, M., Giglioli-Guivarc'h, N., St-Pierre, B., Werck-Reichhart, D., Burlat, V., De Luca, V., O'Connor, S.E., Courdavault, V. (2013). A pair of tabersonine 16-hydroxylases initiates the synthesis of vindoline in an organ-dependent manner in *Catharanthus roseus*. *Plant Physiol* **163**:1792-1803.

Bhadra, R., Vani, S., Shanks, J.V. (1993). Production of indole alkaloids by selected hairy root lines of *Catharanthus roseus*. *Biotechnol Bioeng* **41**:581-592.

Bomati, E.K., Noel, J.P. (2005). Structural and kinetic basis for substrate selectivity in *Populus tremuloides* sinapyl alcohol dehydrogenase. *Plant Cell* **17**:1598-1611.

Bosron, W.F., Magnes, L.J., Li, T.K. (1983). Kinetic and electrophoretic properties of native and recombined isoenzymes of human liver alcohol dehydrogenase. *Biochemistry* **22**:1852-1857.

Botticella, E., Sestili, F., Hernandez-Lopez, A., Phillips, A., Lafiandra, D. (2011). High resolution melting analysis for the detection of EMS induced mutations in wheat *SbeIIa* genes. *BMC Plant Biol* **11**:156.

Burch-Smith, T.M., Anderson, J.C., Martin, G.B., Dinesh-Kumar, S.P. (2004). Applications and advantages of virus-induced gene silencing for gene function studies in plants. *Plant J* **39**:734-746.

Burlat, V., Oudin, A., Courtois, M., Rideau, M., St-Pierre, B. (2004). Co-expression of three MEP pathway genes and *geraniol 10-hydroxylase* in internal phloem parenchyma of *Catharanthus roseus* implicates multicellular translocation of intermediates during the biosynthesis of monoterpenoid indole alkaloids and isoprenoid-derived primary metabolites. *Plant J* **38**:131-141.

Cacace, S., Schröder, G., Wehinger, E., Strack, D., Schmidt, J., Schröder, J. (2003). A flavonol *O*-methyltransferase from *Catharanthus roseus* performing two sequential methylations. *Phytochemistry* **62**:127-137.

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* **17**:540-552.

Chen, F., Tholl, D., Bohlmann, J., Pichersky, E. (2011). The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. *Plant J* **66**:212-229.

Chevenet, F., Brun, C., Bañuls, A.L., Jacq, B., Christen, R. (2006). TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinformatics* **7**:439.

Clarke, A.R., Daffron, T.R. (1998). *Nicotinamide cofactor-dependent enzymes*. San Diego: Academic Press.

Collu, G., Unver, N., Peltenburg-Looman, A.M.G., van der Heijden, R., Verpoorte, R., Memelink, J. (2001). Geraniol 10-hydroxylase, a cytochrome P450 enzyme involved in terpenoid indole alkaloid biosynthesis. *FEBS Lett* **508**:215-220.

Collu, G., Garcia, A.A., van der Heijden, R., Verpoorte, R. (2002). Activity of the cytochrome P450 enzyme geraniol 10-hydroxylase and alkaloid production in plant cell cultures. *Plant Sci* **162**:165-172.

Colonna-Cesari, F., Perahia, D., Karplus, M., Eklund, H., Brädén, C.I., Tapia, O. (1986). Interdomain motion in liver alcohol dehydrogenase. Structural and energetic analysis of the hinge bending mode. *J Biol Chem* **261**:15273-15280.

Constabel, F., Gaudet-Laprairie, P., Kurz, W.G.W., Kutney, J.P. (1982). Alkaloid production in *Catharanthus roseus* cell cultures. XII. Biosynthetic capacity of callus from original explants and regenerated shoots. *Plant Cell Rep* **1**:139-142.

Contin, A., van der Heijden, R., Lefeber, A.W.M., Verpoorte, R. (1998). The iridoid glucoside secologanin is derived from the novel triose phosphate/pyruvate pathway in a *Catharanthus roseus* cell culture. *FEBS Lett* **434**:413-416.

Cooper, J.L., Till, B.J., Laport, R.G., Darlow, M.C., Kleffner, J.M., Jamai, A., El-Mellouki, T., Liu, S., Ritchie, R., Nielsen, N., Bilyeu, K.D., Meksem, K., Comai, L., Henikoff, S. (2008). TILLING to detect induced mutations in soybean. *BMC Plant Biol* **8**:9.

Cordell, G.A. (1974). The biosynthesis of indole alkaloids. *Lloydia* **37**:219-298.

Cordell, G.A., Farnsworth, N.R. (1976). *Catharanthus* alkaloids XXXII: isolation of alkaloids from *Catharanthus trichophyllus* roots and structure elucidation of cathaphylline. *J Pharm Sci* **65**:366-369.

Costa, M., Hilliou, F., Duarte, P., Pereira, L.G., Almeida, I., Leech, M., Memelink, J., Ros Barceló, A., Sottomayor, M. (2008). Molecular cloning and characterization of a vacuolar class III peroxidase involved in the metabolism of anticancer alkaloids in *Catharanthus roseus*. *Plant Physiol* **146**:403-417.

Czerwinski, M. (2011). Preliminary characterization of an ajmalicine accumulating mutant of *Catharanthus roseus* generated by EMS mutagenesis. BSc thesis, Brock University, St. Catharines, Ontario, Canada.

Dalziel, K., Dickinson, F.M. (1966). The kinetics and mechanism of liver alcohol dehydrogenase with primary and secondary alcohols as substrates. *Biochem J* **100**:34-46.

Damiani, I., Morreel, K., Danoun, S., Goeminne, G., Yahiaoui, N., Marque, C., Kopka, J., Messens, E., Goffner, D., Boerjan, W., Boudet, A.M., Rochange, S. (2005). Metabolite profiling reveals a role for atypical cinnamyl alcohol dehydrogenase CAD1 in the synthesis of coniferyl alcohol in tobacco xylem. *Plant Mol Biol* **59**:753-769.

Davioud, E., Kan, C., Hamon, J., Tempé, J., Husson, H-P. (1989). Production of indole alkaloids by *in vitro* root cultures from *Catharanthus trichophyllus*. *Phytochemistry* **28**:2675-2680.

De Carolis, E., Chan, F., Balsevich, J., De Luca, V. (1990). Isolation and characterization of a 2-oxoglutarate dependent dioxygenase involved in the second-to-last step in vindoline biosynthesis. *Plant Physiol* **94**:1323-1329.

De Carolis, E., De Luca, V. (1993). Purification, characterization, and kinetic analysis of a 2-oxoglutarate-dependent dioxygenase involved in vindoline biosynthesis from *Catharanthus roseus*. *J Biol Chem* **268**: 5504-5511.

De Carolis, E. (1994). Enzymology of vindoline biosynthesis: purification, characterization and molecular cloning of a 2-oxoglutarate dependent dioxygenase involved in the second to last step in vindoline biosynthesis from *Catharanthus roseus*. PhD thesis, Université de Montréal, Montréal, Quebec, Canada.

De Luca, V., Balsevich, J., Kurz, W.G.W. (1985). Acetyl coenzyme A: deacetylvindoline *O*-acetyltransferase, a novel enzyme from *Catharanthus*. *J Plant Physiol* **121**:417-428.

De Luca, V., Balsevich, J., Tyler, R.T., Eilert, U., Panchuk, B.D., Kurz, W.G.W. (1986). Biosynthesis of indole alkaloids: developmental regulation of the biosynthetic pathway from tabersonine to vindoline in *Catharanthus roseus*. *J Plant Physiol* **125**:147-156.

De Luca, V., Balsevich, J., Tyler, R.T., Kurz, W.G.W. (1987). Characterization of a novel N-methyltransferase (NMT) from *Catharanthus roseus* plants. *Plant Cell Rep* **6**:458-461.

De Luca, V., Cutler, A.J. (1987). Subcellular localization of enzymes involved in indole alkaloid biosynthesis in *Catharanthus roseus*. *Plant Physiol* **85**:1099-1102.

De Luca, V., Fernandez, J.A., Campbell, D., Kurz, W.G.W. (1988). Developmental regulation of enzymes of indole alkaloid biosynthesis in *Catharanthus roseus*. *Plant Physiol* **86**:447-450.

De Luca, V., Kurz, W.G.W. (1988). Monoterpene indole alkaloids (*Catharanthus* alkaloids). In *Cell Culture and Somatic Cell Genetics of Plants, Vol 5*. Constabel, F., Vasil I.K. (Eds). New York: Academic Press, pp 385-401.

De Luca, V., Marineau, C., Brisson, N. (1989). Molecular cloning and analysis of cDNA encoding a plant tryptophan decarboxylase: Comparison with animal dopa decarboxylases. *Proc Natl Acad Sci USA* **86**:2582-2586.

De Luca, V., Laflamme, P. (2001). The expanding universe of alkaloid biosynthesis. *Curr Opin Plant Biol* **4**:225-233.

De Luca, V., Salim, V., Levac, D., Masada-Atsumi, S., Yu, F. (2012). Discovery and functional analysis of monoterpenoid indole alkaloid pathways in plants. *Methods Enzymol* **515**:207-229.

De Luca, V., Salim, V., Thamm, A., Masada-Atsumi, S., Yu, F. (2014). Making iridoids/secoiridoids and monoterpenoid indole alkaloids: progress on pathway elucidation. *Curr Opin Plant Biol* **19**:35-42.

Decendit, A., Liu, D., Ouelhazi, L., Doireau, P., Mèrillon, J.M., Rideau, M. (1992). Cytokinin-enhanced accumulation of indole alkaloids in *Catharanthus roseus* cell cultures—the factors affecting the cytokinin response. *Plant Cell Rep* **11**:400-403.

Decendit, A., Petit, G., Andreu, F., Doireau, P., Mèrillon, J.M., Rideau, M. (1993). Putative sites of cytokinin action during their enhancing effect on indole alkaloid accumulation in periwinkle cell suspensions. *Plant Cell Rep* **12**:710-712.

Degenhardt, J., Köllner, T.G., Gershenzon, J. (2009). Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. *Phytochemistry* **70**:1621-1637.

Deng, W.W., Zhang, M., Wu, J.Q., Jiang, Z.Z., Tang, L., Li, Y.Y., Wei, C.L., Jiang, C.J., Wan, X.C. (2013). Molecular cloning, functional analysis of three cinnamyl alcohol dehydrogenase (CAD) genes in the leaves of tea plant, *Camellia sinensis*. *J Plant Physiol* **170**:272-282.

Dereeper, A., Audic, S., Claverie, J.M., Blanc, G. (2010). BLAST-EXPLORER helps you building datasets for phylogenetic analysis. *BMC Evol Biol* **10**:8.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.F., Guindon, S., Lefort, V., Lescot, M., Claverie, J.M., Gascuel, O. (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* **36**:W465-469.

Deus-Neumann, B., Zenk, M.H. (1984). A highly selective alkaloid uptake system in vacuoles of higher plants. *Planta* **162**:250-260.

Deus-Neumann, B., Stöckijt, J., Zenk, M.H. (1987). Radioimmunoassay for the quantitative determination of catharanthine. *Planta Med* **53**:184-188.

Dewick, P.M. (2009). *Medicinal Natural Products: A Biosynthetic Approach, 3rd Edition*. West Sussex: John Wiley & Sons, Ltd.

Di Tommaso, P., Moretti, S., Xenarios, I., Orobitg, M., Montanyola, A., Chang, J.M., Taly, J.F., Notredame, C. (2011). T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Res* **39**:13-17.

Do, C.B., Mahabhashyam, M.S., Brudno, M., Batzoglou, S. (2005). ProbCons: Probabilistic consistency-based multiple sequence alignment. *Genome Res* **15**:330-340.

Edge, A. (2013). Biochemical characterization of recombinant loganic acid *O*-methyltransferase from *Lonicera japonica*. BSc thesis, Brock University, St. Catharines, Ontario, Canada.

Eilert, U., Constabel, F., Kurz, W.G.W. (1986). Elicitor-stimulation of monoterpene indole alkaloid formation in suspension cultures of *Catharanthus roseus*. *J Plant Physiol* **126**:11-22.

Eilert, U., De Luca, V., Constabel, F., Kurz, W.G. (1987a). Elicitor-mediated induction of tryptophan decarboxylase and strictosidine synthase activities in cell suspension cultures of *Catharanthus roseus*. *Arch Biochem Biophys* **254**:491-497.

Eilert, U., De Luca, V., Kurz, W.G.W., Constabel, F. (1987b). Alkaloid formation by habituated and tumorous cell suspension cultures of *Catharanthus roseus*. *Plant Cell Rep* **6**:271-274.

Eklund, H., Brändén, C.I. (1987). Alcohol dehydrogenase. In *Biological Macromolecules and Assemblies, Vol 3*. Jurnak, F.A., McPherson, A. (Eds). London: John Wiley & Sons, Inc, pp 75-142.

Eklund, H., Samama, J.P., Jones, T.A. (1984). Crystallographic investigations of nicotinamide adenine dinucleotide binding to horse liver alcohol dehydrogenase. *Biochemistry* **23**:5982-5996.

El-Sayed, M., Verpoorte, R. (2002). Effect of phytohormones on growth and alkaloid accumulation by a *Catharanthus roseus* cell suspension culture fed with alkaloid precursors tryptamine and loganin. *Plant Cell Tissue Organ Cult* **68**:265-270.

El-Sayed, M., Choi, Y.H., Frédérich, M., Roytrakul, S., Verpoorte, R. (2004). Alkaloid accumulation in *Catharanthus roseus* cell suspension fed with stemmadenine. *Biotechnol Lett* **26**:793-798.

El-Sayed, M., Verpoorte, R. (2007). *Catharanthus* terpenoid indole alkaloids: biosynthesis and regulation. *Phytochem Rev* **6**:277-305.

Endo, T., Goodbody, A., Vukovic, J., Misawa, M. (1988). Enzymes from *Catharanthus roseus* cell suspension cultures that couple vindoline and catharanthine to form 3',4'-anhydrovinblastine. *Phytochemistry* **27**:2147-2149.

Escher, S., Loew, P., Arigoni, D. (1970). The role of hydroxygeraniol and hydroxynerol in the biosynthesis of loganin and indole alkaloids. *Chem Commun* 823-825.

Esposito, L., Bruno, I., Sica, F., Raia, C.A., Giordano, A., Rossi, M., Mazzarella, L., Zagari, A. (2003). Crystal structure of a ternary complex of the alcohol dehydrogenase from *Sulfolobus solfataricus*. *Biochemistry* **42**:14397-14407.

Facchini, P.J., Huber-Allanach, K.L., Tari, L.W. (2000). Plant aromatic L-amino acid decarboxylases: evolution, biochemistry, regulation, and metabolic engineering applications. *Phytochemistry* **54**:121-138.

Facchini, P.J., Bohlmann, J., Covello, P.S., De Luca, V., Mahadevan, R., Page, J.E., Ro, D.K., Sensen, C.W., Storms, R., Martin, V.J. (2012). Synthetic biosystems for the production of high-value plant metabolites. *Trends Biotechnol* **30**:127-131.

Fahn, W., Laußermair, E., Deus-Neumann, B., Stöckigt, J. (1985a). Late enzymes of vindoline biosynthesis. *Plant Cell Rep* **4**:337-340.

Fahn, W., Gundlach, H., Deus-Neumann, B., Stöckigt, J. (1985b). Late enzymes of vindoline biosynthesis. Acetyl-CoA:17-*O*-deacetylvindoline 17-*O*-acetyl-transferase. *Plant Cell Rep* **4**:333-336.

Furuya, T., Sakamoto, K., Iida, K., Asada, Y., Yoshikawa, T., Sakai, S-T., Aimi, N. (1992). Biotransformation of tabersonine in cell suspension cultures of *Catharanthus roseus*. *Phytochemistry* **31**:3065-3068.

Gantet, P., Imbault, N., Thiersault, M., Doireau, P. (1999). Necessity of a functional octadecanoic pathway for indole alkaloid biosynthesis by *Catharanthus roseus* cell suspensions cultured in an auxin-starved medium. *Plant Cell Physiol* **39**:220-225.

Gavidia, I., Tarrio, R., Rodriguez-Trelles, F., Perez-Bermudez, P., Seitz, H.U. (2007). Plant progesterone 5β-reductase is not homologous to the animal enzyme. Molecular evolutionary characterization of P5βR from *Digitalis purpurea*. *Phytochemistry* **68**:853-864.

Geerlings, A., Martinez-Lozano Ibiñez, M., Memelink, J., van der Heijden, R., Verpoorte, R. (2000). Molecular cloning and analysis of strictosidine *β*-D-glucosidase, an enzyme in terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. *J Biol Chem* **275**:3051-3056.

Geisler, K., Hughes, R.K., Sainsbury, F., Lomonossoff, G.P., Rejzek, M., Fairhurst, S., Olsen, C.E., Motawia, M.S., Melton, R.E., Hemmings, A.M., Bak, S., Osbourn, A. (2013). Biochemical analysis of a multifunctional cytochrome P450 (CYP51) enzyme required for synthesis of antimicrobial triterpenes in plants. *Proc Natl Acad Sci USA* **110**:E3360-E3367.

Geu-Flores, F., Sherden, N.H., Courdavault, V., Burlat, V., Glenn, W.S., Wu, C., Nims, E., Cui, Y., O'Connor, S.E. (2012). An alternative route to cyclic terpenes by reductive cyclization in iridoid biosynthesis. *Nature* **492**:138-142.

Gibson, R.A., Barrett, G., Wightman, F. (1972). Biosynthesis and metabolism of indol-3yl-acetic acid III: Partial purification and properties of a tryptamine-forming L-

tryptophan decarboxylase from tomato shoots. *J Exp Bot* **23**:775-786.

Giddings, L.A., Liscombe, D.K., Hamilton, J.P., Childs, K.L., DellaPenna, D., Buell, C.R., O'Connor, S.E. (2011). A stereoselective hydroxylation step of alkaloid biosynthesis by a unique cytochrome P450 in *Catharanthus roseus*. *J Biol Chem* **286**:16751-16757.

Gigant, B., Wang, C., Ravelli, R.B., Roussi, F., Steinmetz, M.O., Curmi, P.A., Sobel, A., Knossow, M. (2005). Structural basis for the regulation of tubulin by vinblastine. *Nature* **435**:519-522.

Goeggel, H., Arigoni, D. (1965). The mevalonoid nature of vindoline and reserpinine. *Chem Commun* 538-539.

Goffner, D., Joffroy, I., Grima-Pettenati, J., Halpin, C., Knight, M.E., Schuch, W. (1992). Purification and characterization of isoforms of cinnamyl alcohol dehydrogenase from *Eucalyptus* xylem. *Planta* **188**:48-53.

Goffner, D., Van Doorsselaere, J., Yahiaoui, N., Samaj, J., Grima-Pettenati, J., Boudet, A.M. (1998). A novel aromatic alcohol dehydrogenase in higher plants: molecular cloning and expression. *Plant Mol Biol* **36**:755-765.

Góngora-Castillo, E., Fedewa, G., Yeo, Y., Chappell, J., DellaPenna, D., Buell, C.R. (2012a). Chapter seven – genomic approaches for interrogating the biochemistry of medicinal plant species. *Methods Enzymol* **517**:139-159.

Góngora-Castillo, E., Childs, K.L., Fedewa, G., Hamilton, J.P., Liscombe, D.K., Magallanes-Lundback, M., Mandadi, K.K., Nims, E., Runguphan, W., Vaillancourt, B., Varbanova-Herde, M., Dellapenna, D., McKnight, T.D., O'Connor, S., Buell, C.R. (2012b). Development of transcriptomic resources for interrogating the biosynthesis of monoterpene indole alkaloids in medicinal plant species. *PLoS One* **7**:e52506.

Greene, E.A., Codomo, C.A., Taylor, N.E., Henikoff, J.G., Till, B.J., Reynolds, S.H., Enns, L.C., Burtner, C., Johnson, J.E., Odden, A.R., Comai, L., Henikoff, S. (2003). Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in *Arabidopsis*. *Genetics* **164**:731-740.

Gross, G.G., Stöckigt, J., Mansell, R.L., Zenk, M.H. (1973). Three novel enzymes involved in the reduction of ferulic acid to coniferyl alcohol in higher plants: ferulate:CoA ligase, feruloyl-CoA reductase and coniferyl alcohol oxidoreductase. *FEBS Lett* **31**:283-286.

Grosse, W., Klapheck, S. (1979). Enzymatische Untersuchungen über die Biosynthese von Serotonin und ihre Regulation in Samen von *Juglans regia* L.: Enzymic studies on the biosynthesis of serotonin and the regulation in walnuts (*Juglans regia* L.). *Z Pflanzenphysiologie* **93**:359-363.

Guindon, S., Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52**:696-704.

Guirimand, G., Burlat, V., Oudin, A., St-Pierre, B., Courdavault, V. (2009). Optimization of the transient transformation of *Catharanthus roseus* cells by particle bombardment and its application to the subcellular localization of hydroxymethylbutenyl 4-diphosphate synthase and geraniol 10-hydroxylase. *Plant Cell Rep* **28**:1215-1234.

Guirimand, G., Courdavault, V., Lanoue, A., Mahroug, S., Guihur, A., Blanc, N., Giglioli-Guivarc'h, N., St-Pierre, B., Burlat, V. (2010). Strictosidine activation in Apocynaceae: towards a "nuclear time bomb"? *BMC Plant Biol* **10**:182.

Guirimand, G., Guihur, A., Ginis, O., Poutrain, P., Héricourt, F., Oudin, A., Lanout, A., St-Pierre, B., Burlat, V., Courdavault, V. (2011a). The subcellular organization of strictosidine biosynthesis in *Catharanthus roseus* epidermis highlights several trans-tonoplast translocations of intermediate metabolites. *FEBS J* **278**:749-763.

Guirimand, G., Guihur, A., Poutrain, P., Héricourt, F., Mahroug, S., St-Pierre, B., Burlat, V., Courdavault, V. (2011b). Spatial organization of the vindoline biosynthetic pathway in *Catharanthus roseus*. *J Plant Physiol* **168**:549-557.

Guo, D-M., Ran, J-H., Wang, X-Q. (2010). Evolution of the cinnamyl/sinapyl alcohol dehydrogenase (CAD/SAD) gene family: the emergence of real lignin is associated with the origin of *bona fide CAD*. *J Mol Evol* **71**:202-218.

Hall, E.S., McCapra, F., Money, T., Fukumoto, K., Hanson, J.R., Mootoo, B.S., Phillips, G.T., Scott, A.I. (1966). Concerning the terpenoid origin of indole alkaloids: biosynthetic mapping by direct mass spectrometry. *Chem Commun* 348-350.

Hallahan, D.L., West, J.M., Wallsgrove, R.M., Smiley, D.W.M., Dawson, G.W., Pickett, J.A., Hamilton, J.G.C. (1995). Purification and characterization of an acyclic monoterpene primary alcohol:$NADP^+$ oxidoreductase from catmint (*Nepeta racemosa*). *Arch Biochem Biophys* **318**:105-112.

Hartwell, L.H., Hood, L., Goldberg, M.L., Reynolds, A.E., Silver, L.M. (2011). Mendel's Principles of Heredity. In *Genetics: From Genes to Genome, 4th Edition*. New York: McGraw-Hill, pp 13-42.

Haughn, G., Somerville, C.R. (1987). Selection for herbicide resistance at the whole-plant level. In *Applications of Biotechnology to Agricultural Chemistry*. LeBaron, H.M., Mumma, R.O., Honeycutt, R.C., Duesing, J.H. (Eds). Easton: American Chemical Society, pp 98-107.

Hemscheidt, T., Zenk, M.H. (1980). Glucosidases involved in indole alkaloid biosynthesis of *Catharanthus* cell cultures. *FEBS Lett* **110**:187-191.

Hol, W.G., van Duijnen, P.T., Berendsen, H.J. (1978). The alpha-helix dipole and the properties of proteins. *Nature* **273**:443-446.

Hong, S-B., Hughes, E.H., Shanks, J.V., San, K-Y., Gibson, S.I. (2003). Role of the non-

mevalonate pathway in indole alkaloid production by *Catharanthus roseus* hairy roots. *Biotechnol Prog* **19**:1105-1108.

Hurley, T.D., Edenberg, H.J., Bosron, W.F. (1990). Expression and kinetic characterization of variants of human $\beta_1\beta_1$ alcohol dehydrogenase containing substitutions at amino acid 47. *J Biol Chem* **265**:16366-16372.

Ikeda, H., Esaki, N., Nakai, S., Hashimoto, K., Uesato, S., Soda, K., Fujita, T. (1991). Acyclic monoterpene primary alcohol:$NADP^+$ oxidoreductase of *Rauwolfia serpentina* cells: the key enzyme in biosynthesis of monoterpene alcohols. *J Biochem* **109**:341-347.

Inouye, H., Ueda, S., Aoki, Y,. Takeda, Y. (1969). Zur Biosynthese der Iridoidglucoside. *Tetrahedron Lett* **10**:2351-2354.

Inouye, H. (1978). Neuere Ergebnisse über die Biosynthese der Glucoside der Iridoidreihe. *Planta Med* **33**:193–216.

Irmler, S., Schröder, G., St-Pierre, B., Crouch, N.P., Hotze, M., Schmidt, J., Strack, D., Matern, U., Schröder, J. (2000). Indole alkaloid biosynthesis in *Catharanthus roseus*: new enzyme activities and identification of cytochrome P450 CYP72A1 as secologanin synthase. *Plant J* **24**:797-804.

Jordan, M.A., Wilson, L. (2004). Microtubules as a target for anticancer drugs. *Nat Rev Cancer* **4**:253-265.

Katano, N., Yamamoto, H., Iio, R., Inoue, K. (2001). 7-Deoxyloganin 7-hydroxylase in *Lonicera japonica* cell cultures. *Phytochemistry* **58**:53-58.

Kellner, F., Geu-Flores, F., Sherden, N.H., Brown, S., Foureau, E., Courdavault, V., O'Connor, S.E. (2015). Discovery of a P450-catalyzed step in vindoline biosynthesis: a link between the aspidosperma and eburnamine alkaloids. *Chem Commun* **51**:7626-7628.

Kim, S.J., Kim, K.W., Cho, M.H., Franceschi, V.R., Davin, L.B., Lewis, N.G. (2007). Expression of cinnamyl alcohol dehydrogenases and their putative homologues during *Arabidopsis thaliana* growth and development: lessons for database annotations? *Phytochemistry* **68**:1957-1974.

Kim, Y., Schumaker, K.S., Zhu, J.K. (2006). Methods in Molecular Biology, Vol 323: Arabidopsis Protocols, 2nd Edition. Salinas, J., Sanchez-Serrano, J.J. (Eds.). Totowa: Human Press, Inc.

Kulkarni, R.N., Baskaran, K., Chandrashekara, R.S., Kumar, S. (1999). Inheritance of morphological traits of periwinkle mutants with modified contents and yields of leaf and root alkaloids. *Plant Breeding* **118**:71-74.

Kumari, R., Sharma, V., Sharma, V., Kumar, S. (2013). Pleiotropic phenotypes of the salt-tolerant and cytosine hypomethylated leafless inflorescence, evergreen dwarf and irregular leaf lamina mutantsof *Catharanthus roseus* possessing Mendelian inheritance. *J Genet* **92**:369-394.

Kurowska, M., Daszkowska-Golec, A., Gruszka, D., Marzec, M., Szurman, M., Szarejko, I., Maluszynski, M. (2011). TILLING – a shortcut in functional genomics. *J Appl Genet* **52**:371-390.

Kurz, W.G.W., Chatson, K.B., Constabel, F., Kutney, J.P., Choi, L.S.L., Kolodziejczyk, P, Sleigh, S.K., Stuart, K.L., Worth, B.R. (1980). Alkaloid production in *Catharanthus roseus* cell cultures: initial studies on cell lines and their alkaloid content. *Phytochemistry* **19**:2583-2587.

Kurz, W.G.W., Constabel, F. (1985). Aspects of affecting biosynthesis and biotransformation of secondary metabolites in plant cell cultures. *CRC Crit Rev Biotechnol* **2**:105-118.

Kutchan, T.M., Hampp, N., Lottspeich, F., Beyreuther, K., Zenk, M.H. (1988). The cDNA clone for strictosidine synthase from *Rauvolfia serpentina* DNA sequence determination and expression in *Escherichia coli*. *FEBS Lett* **237**:40-44.

Kutney, J.P., Choi, L.S.L., Kolodziejczyk, P., Sleigh, S.K., Stuart, K.L., Worth, B.R., Kurz, W.G.W., Chatson, K.B., Constabel, F. (1980). Alkaloid production in *Catharanthus roseus* cell cultures: Isolation and characterization of alkaloids from one cell line. *Phytochemistry* **19**:2589-2595.

Kutney, J.P., Choi, L.S.L., Honda, T., Lewis, N.G., Sato, T., Stuart, K.L., Worth, B.R. (1982). Biosynthesis of the indole alkaloids. Cell-free systems from *Catharanthus roseus* plants. *Helv Chim Acta* **65**:2088-2101.

Kwok, P.Y. (2001). Methods for genotyping single nucleotide polymorphisms. *Annu Rev Genomics Hum Genet* **2**:235-258.

Laflamme, P., St-Pierre, B., De Luca, V. (2001). Molecular and biochemical analysis of a Madagascar periwinkle root-specific minovincinine-19-hydroxy-*O*-acetyltransferase. *Plant Physiol* **125**:189-198.

Lauvergeat, V., Kennedy, K., Feuillet, C., McKie, J.H., Gorrichon, L., Baltas, M., Boudet, A.M., Grima-Pettenati, J., Douglas, K.T. (1995). Site-directed mutagenesis of a serine residue in cinnamyl alcohol dehydrogenase, a plant NADPH-dependent dehydrogenase, affects the specificity for the coenzyme. *Biochemistry* **34**:12426-12434.

Levac, D., Murata, J., Kim, W.S., De Luca, V. (2008). Application of carborundum abrasion for investigating the leaf epidermis: molecular cloning of *Catharanthus roseus* 16-hydroxytabersonine-16-*O*-methyltransferase. *Plant J* **53**:225-236.

Liscombe, D.K., Usera, A.R., O'Connor, S.E. (2010). Homolog of tocopherol *C* methyltransferases catalyzes *N* methylation in anticancer alkaloid biosynthesis. *Proc Natl Acad Sci USA* **107**:18793-18798.

Liscombe, D.K., O'Connor, S.E. (2011). A virus-induced gene silencing approach to understanding alkaloid metabolism in *Catharanthus roseus*. *Phytochemistry* **72**:1969-1977.

Loew, P., Goeggel, H., Arigoni, D. (1966). A monoterpene precursor in the biosynthesis of indole alkaloids. *Chem Commun* 347-348.

Loew, P., Arigoni, D. (1968). The biological conversion of loganin into indole alkaloids. *Chem Commun* 137-138.

Luijendijk, T.J.C., Stevens, L.H., Verpoorte, R. (1998). Purification and characterisation of strictosidine β-D-glucosidase from *Catharanthus roseus* cell suspension cultures. *Plant Physiol Biochem* **36**:419-425.

Madyastha, K.M., Guarnaccia, R., Coscia, C.J. (1971). Enzymic synthesis of loganin by carboxyl group methylation of loganic acid. *FEBS Lett* **14**:175-177.

Madyastha, K.M., Guarnaccia, R., Baxter, C., Coscia, C.J. (1973). *S*-Adenosyl-L-methionine: loganic acid methyltransferase. *J Biol Chem* **248**:2497-2501.

Madyastha, K.M., Meehan, T.D., Coscia, C.J. (1976). Characterization of a cytochrome P-450 dependent monoterpene hydroxylase from the higher plant *Vinca rosea*. *Biochemistry* **15**:1097-1102.

Madyastha, K.M., Ridgway, J.E., Dwyer, J.G., Coscia, C.J. (1977). Subcellular localization of a cytochrome P-450-dependent monogenase in vesicles of the higher plant *Catharanthus roseus*. *J Cell Biol* **72**:302-313.

Mahroug, S., Burlat, V., St-Pierre, B. (2007). Cellular and sub-cellular organisation of the monoterpenoid indole alkaloid pathway in *Catharanthus roseus*. *Phytochem Rev* **6**:363-381.

Magnotta, M., Murata, J., Chen, J., De Luca, V. (2006). Identification of a low vindoline accumulating cultivar of *Catharanthus roseus* (L.) G. Don by alkaloid and enzymatic profiling. *Phytochemistry* **67**:1758-1764.

Magnotta, M., Murata, J., Chen, J., De Luca, V. (2007). Expression of deacetylvindoline-4-*O*-acetyltransferase in *Catharanthus roseus* hairy roots. *Phytochemistry* **68**:1922-1931.

Mangold, U., Eichel, J., Batschauer, A., Lanz, T., Kaiser, T., Spangenberg, G., Werck-Reichhart, D., Schröder, J. (1994). Gene and cDNA for plant cytochrome P450 proteins (CYP72 family) from *Catharanthus roseus*, and transgenic expression in tobacco and *Arabidopsis thaliana*. *Plant Sci* **96**:129-136.

Mann, J., Davidson, R.S., Hobbs, J.B., Banthorpe, D.V., Harborne, J.B. (1994). *Natural products: their chemistry and biological significance*. Harlow: Longman Scientific and Technical, p 423.

Mansell, R.L.G., Gross, G.G., Stöckigt, J., Franke, H., Zenk, M.H. (1974). Purification and properties of cinnamyl alcohol dehydrogenase from higher plants involved in lignin biosynthesis. *Phytochemistry* **13**:2427-2435.

McCapra, F., Money, T., Scott, A.I., Wright, I.G. (1965). Biosynthesis of the indole alkaloids: vindoline. *Chem Commun* 537-538.

McFarlane, J., Madyastha, K.M., Coscia, C.J. (1975). Regulation of secondary metabolism in higher plants. Effect of alkaloids on a cytochrome P-450 dependent monooxygenase. *Biochem Biophys Res Commun* **66**:1263-1269.

McKnight, T.D., Roessner, C.A., Devagupta, R., Scott, A.I., Nesslier, C.L. (1990). Nucleotide sequence of a cDNA encoding the vacuolar protein strictosidine synthase from *Catharanthus roseus*. *Nucl Acids Res* **18**:4939.

McKnight, T.D., Bergey, D.R., Burnett, R.J., Nessler, C.L. (1991). Expression of enzymatically active and correctly targeted strictosidine synthase in transgenic tobacco plants. *Planta* **185**:148-152.

McLauchlan, W.R., Hasan, M., Baxter, R.L. (1983). Conversion of anhydrovinblastine to vinblastine by cell-free homogenates of *Catharanthus* roseus cell suspension cultures. *Tetrahedron* **39**:3777-3780.

Meehan, T.D., Coscia, C.J. (1973). Hydroxylation of geraniol and nerol by a monooxygenase. *Biochem Biophys Res Commun* **53**:1043-1048.

Meijer, A.H., Verpoorte, R., Hoge, J.H.C. (1993a). Regulation of enzymes and genes involved in terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. *J Plant Res* (special issue) **3**:145-164.

Meijer, A.H., De Waal, A., Verpoorte, R. (1993b). Purification of the cytochrome P-450 enzyme geraniol 10-hydroxylase from cell cultures of *Catharanthus roseus*. *J Chromatogr* **635**:237-249.

Memelink, J., Verpoorte, R., Kijne, J.W. (2001). ORCAnisation of jasmonate-responsive gene expression in alkaloid metabolism. *Trends Plant Sci* **6**:212-219.

Menke, F.L.H., Champion, A., Kijne, J.W., Memelink, J. (1999). A novel jasmonate- and elicitor-responsive element in the periwinkle secondary metabolite biosynthetic gene *Str* interacts with a jasmonate- and elicitor-inducible AP2-domain transcription factor, ORCA2. *EMBO J* **18**:4455-4463.

Miettinen, K., Dong, L., Navrot, N., Schneider, T., Burlat, V., Pollier, J., Woittiez, L., van der Krol, S., Lugan, R., Ilc, T., Verpoorte, R., Oksman-Caldentey, K-M., Martinoia, E., Bouwmeester, H., Goossens, A., Memelink, J., Werck-Reichhard, D. (2014). The seco-iridoid pathway from *Catharanthus roseus*. *Nat Commun* **5**:3606.

Min, S., Kasahara, H., Bedgar, D.L., Youn, B., Lawrence, P.K., Gang, D.R., Halls, S.C., Park, H., Hilsenbeck, J.L., Davin, L.B., Lewis, N.G., Kang, C. (2003). Crystal structures of pinoresinol-lariciresinol and phenylcoumaran benzylic ether reductases and their relationship to isoflavone reductases. *J Biol Chem* **278**:50714-50723.

Misawa, M., Endo, T., Goodbody, A., Vukovic, J., Chapplet, C., Choi, L., Kutney, J.P. (1988). Synthesis of dimeric indole alkaloids by cell free extracts from cell suspension cultures of *Catharanthus roseus*. *Phytochemistry* **27**:1355-1359.

Mizukami, H., Nordlöv, H., Lee, S-L., Scott, A.I. (1979). Purification and properties of strictosidine synthase (an enzyme condensing tryptamine and secologanin) from *Catharanthus roseus* cultured cells. *Biochemistry* **18**:3760-3763.

Moreno, P.R.H., Poulsen, C., van der Heijden, R., Verpoorte, R. (1996). Effects of elicitation on different secondary metabolic pathways in *Catharanthus roseus* cell suspension cultures. *Enzyme Microb Technol* **18**:99-107.

Moretti, S., Armougom, F., Wallace, I.M., Higgins, D.G., Jongeneel, C.V., Notredame, C. (2007). The M-Coffee web server: a meta-method for computing multiple sequence alignments by combining alternative alignment methods. *Nucleic Acids Res* **35**:645-648.

Morgan, J.A., Shanks, J.V. (1999). Inhibitor studies of tabersonine metabolism in *C. roseus* hairy roots. *Phytochemistry* 51:61-68.

Morgan, J.A., Shanks, J.V. (2000). Determination of metabolic rate-limitations by precursor feeding in *Catharanthus roseus* hairy root cultures. *J Biotechnol* **79**:137-145.

Moza, B.K., Trojánek, J. (1963). On alkaloids. VII. New alkaloids from *Catharanthus roseus* G. Don. *Collect Czech Chem Commun* **28**:1419-1426.

Murata, J., De Luca, V. (2005). Localization of tabersonine 16-hydroxylase and 16-OH tabersonine-16-*O*-methyltransferase to leaf epidermal cells defines them as a major site of precursor biosynthesis in the vindoline pathway in *Catharanthus roseus*. *Plant J* **44**:581-594.

Murata, J., Roepke, J., Gordon, H., De Luca, V. (2008). The leaf epidermome of *Catharanthus roseus* reveals its biochemical specialization. *Plant Cell* **20**:524-542.

Nagatoshi, M., Terasaka, K., Nagatsu, A., Mizukami, H. (2011). Iridoid-specific glucosyltransferase from *Gardenia jasminoides*. *J Biol Chem* **286**:32866-32874.

Noé, W., Mollenschott, C., Berlin, J. (1984). Tryptophan decarboxylase from *Catharanthus roseus* cell suspension cultures: purification, molecular and kinetic data of the homogenous protein. *Plant Mol Biol* **3**:281-288.

Notredame, C., Higgins, D.G., Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* **302**:205-217.

O'Connor, S.E., Maresh, J.J. (2006). Chemistry and biology of monoterpene indole alkaloid biosynthesis. *Nat Prod Rep* **23**:532-547.

O'Keefe, B.R., Mahady, G.B., Gills, J.J., Beecher, C.W.W. (1997). Stable vindoline production in transformed cell cultures of *Catharanthus roseus*. *J Nat Prod* **60**:261-264.

Okabe, Y., Asamizu, E., Saito, T., Matsukura, C., Ariizumi, T., Brès, C., Rothan, C., Mizoguchi, T., Ezura, H. (2011). Tomato TILLING technology: development of a reverse genetics tool for the efficient isolation of mutants from Micro-Tom mutant libraries. *Plant Cell Physiol* **52**:1994-2005.

Palazón, J., Cusidó, R.M., Gonzalo, J., Bonfill, M., Morales, C., Piñol, M.T. (1998). Relation between the amount of *rolC* gene product and indole alkaloid accumulation in *Catharanthus roseus* transformed root cultures. *J Plant Physiol* **153**:712-718.

Pan, H., Zhou, R., Louie, G.V., Mühlemann, J.K., Bomati, E.K., Bowman, M.E., Dudareva, N., Dixon, R.A., Noel, J.P., Wang, X. (2014). Structural studies of cinnamoyl-CoA reductase and cinnamyl-alcohol dehydrogenase, key enzymes of monolignol biosynthesis. *Plant Cell* **26**:3709-3727.

Papadopoulou, K., Melton, R.E., Leggett, M., Daniels, M.J., Osbourn, A.E. (1999). Compromised disease resistance in saponin-deficient plants. *Proc Natl Acad Sci USA* **96**:12923-12928.

Pasquali, G., Goddijn, O.J.M., de Waal, A., Verpoorte, R., Schilperoort, R.A., Hoge, J.H.C., Memelink, J. (1992). Coordinated regulation of two indole alkaloid biosynthetic genes from *Catharanthus roseus* by auxin and elicitors. *Plant Mol Biol* **18**:1121-1131.

Powell, A.J., Read, J.A., Banfield, M.J., Gunn-Moore, F., Yan, S.D., Lustbader, J., Stern, A.R., Stern, D.M., Brady, R.L. (2000). Recognition of structurally diverse substrates by type II 3-hydroxyacyl-CoA dehydrogenase (HADH II)/amyloid-ß binding alcohol dehydrogenase (ABAD). *J Mol Biol* **303**:311-327.

Power, R., Kurz, W.G.W., De Luca, V. (1990). Purification and characterization of acetylcoenzyme A:deacetylvindoline 4-*O*-acetyltransferase from *Catharanthus roseus*. *Arch Biochem Biophys* **279**:370-376.

Qin, B., Eagles, J., Mellon, F.A., Mylona, P., Peña-Rodriguez, L., Osbourn, A.E. (2010). High throughput screening of mutants of oat that are defective in triterpene synthesis. *Phytochemistry* **71**:1245-1252.

Qu, Y., Easson, M.L.A.E., Froese, J., Simionescu, R., Hudlicky, T., De Luca, V. (2015). Completion of the seven-step pathway from tabersonine to the anticancer drug precursor vindoline and its assembly in yeast. *Proc Natl Acad Sci USA* **112**:6224-6229.

Qureshi, A.A., Scott, A.I. (1968). Biosynthesis of indole alkaloids: sequential precursor formation and biological conversion in *Vinca rosea*. *Chem Commun* 948-950.

Rai, S.P., Luthra, R., Kumar, S. (2003). Salt-tolerant mutants in glycophytic salinity response (GSR) genes in *Catharanthus roseus*. *Theor Appl Genet* **106**:221-230.

Rodriguez, S., Compagnon, V., Crouch, N.P., St-Pierre, B., De Luca, V. (2003). Jasmonate-induced epoxidation of tabersonine by a cytochrome P-450 in hairy root cultures of *Catharanthus roseus*. *Phytochemistry* **64**:401-409.

Roepke, J., Salim, V., Wu, M., Thamm, A.M.K., Murata, J., Ploss, K., Boland, W., De Luca, V. (2010). Vinca drug components accumulate exclusively in leaf exudates of Madagascar periwinkle. *Proc Natl Acad Sci USA* **107**:15287-15292.

Roewer, I.A., Cloutier, N., Nessler, C.L., De Luca, V. (1992). Transient induction of tryptophan decarboxylase (TDC) and strictosidine synthase (SS) genes in cell suspension

cultures of *Catharanthus roseus*. *Plant Cell Rep* **11**:86-89.

Rowland, G.G. (1991). An EMS-induced low-linolenic-acid mutant in McGregor flax (*Linum usitatissimum* L.). *Can J Plant Sci* **71**:393-396.

Salim, V., Yu, F., Altarejos, J., De Luca, V. (2013). Virus-induced gene silencing identifies *Catharanthus roseus* 7-deoxyloganic acid-7-hydroxylase, a step in iridoid in monoterpene indole alkaloid biosynthesis. *Plant J* **76**:754-765.

Salim, V., Wiens, B., Masada-Atsumi, S., Yu, F., De Luca, V. (2014). 7-Deoxyloganetic acid synthase catalyzes a key 3 step oxidation to form 7-deoxyloganetic acid in *Catharanthus roseus* iridoid biosynthesis. *Phytochemistry* **101**:23-31.

Schröder, G., Unterbusch, E., Kaltenbach, M., Schmidt, J., Strack, D., De Luca, V., Schröder, J. (1999). Light-induced cytochrome P450-dependent enzyme in indole alkaloid biosynthesis: tabersonine 16-hydroxylase. *FEBS Lett* **458**:97-102.

Schröder, G., Wehinger, E., Schröder, J. (2002). Predicting the substrates of cloned plant *O*-methyltransferases. *Phytochemistry* **59**:1-8.

Schröder, G., Wehinger, E., Lukacin, R., Wellmann, F., Seefelder, W., Schwab, W., Schröder, J. (2004). Flavonoid methylation: a novel 4'-*O*-methyltransferase from *Catharanthus roseus*, and evidence that partially methylated flavanones are substrates of four different flavonoid dioxygenases. *Phytochemistry* **65**:1085-1094.

Scott, A.I., Reichardt, P.B., Slaytor, M.B., Sweeny, J.G. (1971). Mechanisms of indole alkaloid biosynthesis. Recognition of intermediacy and sequence by short-term incubation. *Bioorg Chem* **1**:157-173.

Scott, A.I., Mizukami, H., Hirato, T., Lee, S.L. (1980). Formation of catharanthine, akuammicine and vindoline in *Catharanthus roseus* suspension cells. *Phytochemistry* **19**:488-489.

Shan, L.L., Li, X., Wang, P., Cai, C., Zhang, B., Sun, C.D., Zhang, W.S., Xu, C.J., Ferguson, I., Chen, K.S. (2008). Characterization of cDNAs associated with lignification and their expression profiles in loquat fruit with different lignin accumulation. *Planta* **227**:1243-1254.

Shanks, J.V., Bhadra, R., Morgan, J., Rijhwani, S., Vani, S. (1998). Quantification of metabolites in the indole alkaloid pathways of *Catharanthus roseus*: implications for metabolic engineering. *Biotechnol Bioeng* **58**:333-338.

Simkin, A.J., Miettinen, K., Claudel, P., Burlat, V., Guirimand, G., Courdavault, V., Papon, N., Meyer, S., Godet, S., St-Pierre, B., Giglioli-Guivarc'h, N., Fischer, M.J., Memelink, J., Clastre, M. (2013). Characterization of the plastidial geraniol synthase from Madagascar periwinkle which initiates the monoterpenoid branch of the alkaloid pathway in internal phloem associated parenchyma. *Phytochemistry* **85**:36-43.

Sottomayor, M., de Pinto, M.C., Salema, R., DiCosmo, F., Pedreño, M.A., Ros Barcelo, A. (1996). The vacuolar localization of a basic peroxidase isoenzyme responsible for the

synthesis of α-3',4'-anhydrovinblastine in *Catharanthus roseus* (L.) G. Don leaves. *Plant Cell Environ* **19**:761-767.

Sottomayor, M., López-Serrano, M., DiCosmo, F., Ros Barceló, A. (1998). Purification and characterization of alpha-3',4'-anhydrovinblastine synthase (peroxidase-like) from *Catharanthus roseus* (L.) G. Don. *FEBS Lett* **428**:299-303.

Smith, J.I., Smart, N.J., Kurz, W.G.W., Misawa, M. (1987). Stimulation of indole alkaloid production in cell suspension cultures of *Catharanthus roseus* by abscisic acid. *Planta Med* **53**:470-474.

Stevens, L.H., Blom, T.J.M., Verpoorte, R. (1993). Subcellular localization of tryptophan decarboxylase, strictosidine synthase and strictosidine glucosidase in suspension cultured cells of *Catharanthus roseus* and *Tabernaemontana divaricata*. *Plant Cell Rep* **12**:573-576.

Stöckijt, J., Soll, H.J. (1980). Indole alkaloids from cell suspension cultures of *Catharanthus roseus* and *C. ovalis*. *Planta Med* **40**:22-30.

Stöckigt, J., Zenk, M.H. (1977a). Isovincoside (strictosidine), the key intermediate in the enzymatic formation of indole alkaloids. *FEBS Lett* **79**:233-237.

Stöckigt, J., Zenk, M.H. (1977b). Strictosidine (isovincoside): the key intermediate in the biosynthesis of monoterpenoid indole alkaloids. *Chem Commun* 646-648.

Stöckigt, J., Pfitzner, A., Firl, J. (1981). Indole alkaloids from cell suspension cultures of *Rauwolfia serpentina* benth. *Plant Cell Rep* **1**:36-39.

Stone, C.L., Li, T.K., Bosron, W.F. (1989). Stereospecific oxidation of secondary alcohols by human alcohol dehydrogenases. *J Biol Chem* **264**:11112-11116.

St-Pierre, B., De Luca, V. (1995). A cytochrome P-450 monooxygenase catalyzes the first step in the conversion of tabersonine to vindoline in *Catharanthus roseus*. *Plant Physiol* **109**:131-139.

St-Pierre, B., Laflamme, P., Alarco, A.M., 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 J* **14**:703-713.

St-Pierre, B., Vazquez-Flota, F.A., De Luca, V. (1999). Multicellular compartmentation of *Catharanthus roseus* alkaloid biosynthesis predicts intercellular translocation of a pathway intermediate. *Plant Cell* **11**:887-900.

Svoboda, G.H., Blake, D.A. (1975). The phytochemistry and pharmacology of *Catharanthus roseus* (L.) G. Don. In *The Catharanthus Alkaloids: Botany, Chemistry, Pharmacology, and Clinical Uses*. Taylor, W.I., Farnsworth, N.R. (Eds). New York: Marcel Dekker, Inc., pp 45-83.

Tanahashi, T., Nagakura, N., Inouye, H., Zenk, M.H. (1984). Radioimmunoassay for the determination of loganin and the biotransformation of loganin to secologanin by plant

cell cultures. *Phytochemistry* **23**:1917-1922.

Thamm, A.M.K. (2014). Identification and characterization of a *Catharanthus roseus* mutant altered in monoterpenoid indole alkaloid biosynthesis. PhD thesis, Brock University, St. Catharines, Ontario, Canada.

Till, B.J., Cooper, J., Tai, T.H., Colowit, P., Greene, E.A., Henikoff, S., Comai, L. (2007). Discovery of chemically induced mutations in rice by TILLING. *BMC Plant Biol* **7**:9.

Treimer, J.F., Zenk, M.H. (1979). Purification and properties of strictosidine synthase, the key enzyme in indole alkaloid formation. *Eur J Biochem* **101**:225-233.

Tyler, R.T., Kurz, W.G.W., Panchuk, B.D. (1986). Photoautotrophic cell suspension cultures of periwinkle (*Catharanthus roseus* (L.) G. Don): transition from heterotrophic to photoautotrophic growth. *Plant Cell Rep* **5**:195-198.

Uesato, S., Matsuda, S., Inouye, H. (1984a). Mechanism for iridane skeleton formation from acyclic monoterpenes in the biosynthesis of secologanin and vindoline in *Catharanthus roseus* and *Lonicera morrowii*. *Chem Pharm Bull* **32**:1671-1674.

Uesato, S., Matsuda, S., Iida, A., Inouye, H., Zenk, M.H. (1984b). Intermediacy of 10-hydroxygeraniol, 10-hydroxynerol and iridodial in the biosynthesis of ajmaline and vomilenine in *Rauwolfia serpentina* suspension cultures. *Chem Pharm Bull (Tokyo)* **32**:3764–3767.

Uesato, S., Kanomi, S., Iida, A., Inouye, H., Zenk, M.H. (1986a). Mechanism for iridane skeleton formation in the biosynthesis of secologanin and indole alkaloids in *Lonicera tatarica*, *Catharanthus roseus*, and suspension cultures of *Rauwolfia serpentina*. *Phytochemistry* **25**:839-842.

Uesato, S., Ogawa, Y., Inouye, H., Saiki, K., Zenk, M.H. (1986b). Synthesis of iridodial by cell free extracts from *Rauwolfia serpentina* cell suspension cultures. *Tetrahedron Lett* **27**:2893-2896.

Uesato, S., Ikeda, H., Fujita, T., Inouye, H., Zenk, M.H. (1987). Elucidation of iridodial formation mechanism – partial purification and characterization of the novel monoterpene cyclase from *Rauwolfia serpentina* cell suspension cultures. *Tetrahedron Lett* **28**:4431-4434.

van der Fits, L., Memelink, J. (2000). ORCA3, a jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism. *Science* **289**:295-297.

van der Fits, L., Memelink, J. (2001). The jasmonate-inducible AP2/ERF-domain transcription factor ORCA3 activates gene expression via interaction with a jasmonate-responsive promoter element. *Plant J* **25**:43-53.

van der Heijden, R., Verpoorte, R., Ten Hoopen, J.G. (1989). Cell and tissue cultures of *Catharanthus roseus* (L.) G. Don: a literature survey. *Plant Cell Tissue Organ Cult*

18:231-280.

van der Heijden, R., Jacobs, D.I., Snoeijer, W., Hallard, D., Verpoorte, R. (2004). The *Catharanthus* alkaloids: pharmacognosy and biotechnology. *Curr Med Chem* **11**:607-628.

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

Vazquez-Flota, F.A., De Luca, V. (1998a). Developmental and light regulation of desacetoxyvindoline 4-hydroxylase in *Catharanthus roseus* (L.) G. Don. *Plant Physiol* **117**:1351-1361.

Vazquez-Flota, F.A., De Luca, V. (1998b). Jasmonate modulates development- and light-regulated alkaloid biosynthesis in *Catharanthus roseus*. *Phytochemistry* **49**:395-402.

Vázquez-Flota, F., Carillo-Pech, M., Mincero-García, Y., de Lourdes Miranda-Ham, M. (2004). Alkaloid metabolism in wounded *Catharanthus roseus* seedlings. *Plant Physiol Biochem* **42**:623-628.

Verma, A.K., Singh, R.R. (2010). Induced dwarf mutant in *Catharanthus roseus* with enhanced antibacterial activity. *Indian J Pharm Sci* **72**:655-657.

Verpoorte, R., van der Heijden, R., Moreno, P.R.H. (1997). Biosynthesis of terpenoid indole alkaloids in *Catharanthus roseus* cells. In *The Alkaloids, Vol 49*. Cordell, G.A. (Ed). San Diego: Academic, pp 221-299.

Vetter, H-P., Mangold, U., Schröder, G., Marner, F-J., Werck-Reichhart, D., Schröder, J. (1992). Molecular analysis and heterologous expression of an inducible cytochrome P-450 protein from periwinkle (*Catharanthus roseus* L.). *Plant Physiol* **100**:998-1007.

Vogt, T., Jones, P. (2000). Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. *Trends Plant Sci* **5**:380-386.

Wallace, I.M., O'Sullivan, O., Higgins, D.G., Notredame, C. (2006). M-Coffee: Combining multiple sequence alignment methods with T-Coffee. *Nucleic Acids Res* **34**:1692-1699.

Westkemper, P., Wieczorek, U., Gueritte, F., Langlois, N., Langlois, Y., Potier, P., Zenk, M.H. (1980). Radioimmunoassay of the indole alkaloid vindoline in *Catharanthus*. *Planta Med* **39**:24-37.

Xiao, M., Zhang, Y., Chen, X., Lee, E-J., Barber, C.J.S., Chakrabarty, R., Desgangé-Penix, I., Haslam, T.M., Kim, Y-B., Liu, E., MacNevin, G., Masada-Atsumi, S., Reed, D.W., Stout, J.M., Zerbe, P., Zhang, Y., Bohlmann, J., Covello, P.S., De Luca, V., Page, J.E., Ro, D-K., Martin, V.J.J., Facchini, P.J., Sensen, C.W. (2013). Transcriptome analysis based on next-generation sequencing of non-model plants producing specialized metabolites of biotechnological interest. *J Biotech* **166**:122-134.

Xin, Z., Wang, M.L., Barkley, N.A., Burow, G., Franks, C., Pederson, G., Burke, J. (2008). Applying genotyping (TILLING) and phenotyping analyses to elucidate gene function in a chemically induced sorghum mutant population. *BMC Plant Biol* **8**:103.

Yamamoto, H., Katano, N., Ooi, A., Inoue, K. (1999). Transformation of loganin and 7-deoxyloganin into secologanin by *Lonicera japonica* cell suspension cultures. *Phytochemistry* **50**:417-422.

Yamamoto, H., Katano, N., Ooi, A., Inoue, K. (2000). Secologanin synthase which catalyzes the oxidative cleavage of loganin into secologanin is a cytochrome P450. *Phytochemistry* **53**:7-12.

Yang, H., Ganguly, A., Cabral, F. (2010). Inhibition of cell migration and cell division correlates with distinct effects of microtubule inhibiting drugs. *J Biol Chem* **285**:32242-50.

Youn, B., Camacho, R., Moinuddin, S.G.A., Lee, C., Davin, L.B., Lewis, N.G., Kang, C. (2006). Crystal structures and catalytic mechanism of the *Arabidopsis* cinnamyl alcohol dehydrogenases AtCAD5 and AtCAD4. *Org Biomol Chem* **4**:1687-1697.

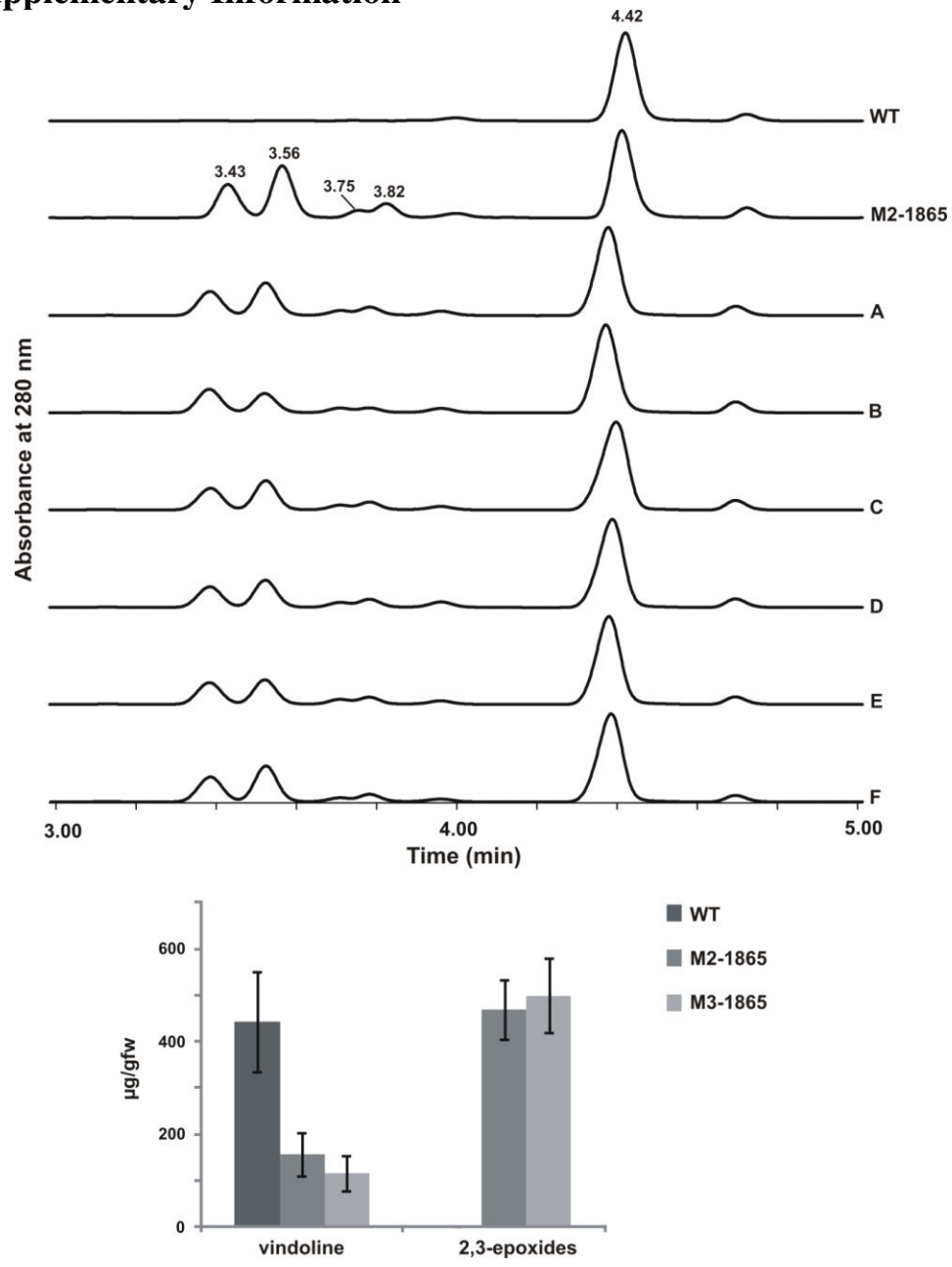
Youn, B., Moinuddin, S.G.A., Davin, L.B., Lewis, N.G., Kang, C. (2005). Crystal structures of apo-form and binary/ternary complexes of *Podophyllum* secoisolariciresinol dehydrogenase, an enzyme involved in formation of health-protecting and plant defense lignans. *J Biol Chem* **280**:12917-12926.

Yu, F., De Luca, V. (2013). ATP-binding cassette transporter controls leaf surface secretion of anticancer drug components in *Catharanthus roseus*. *Proc Natl Acad Sci USA* **110**:15830-15835.

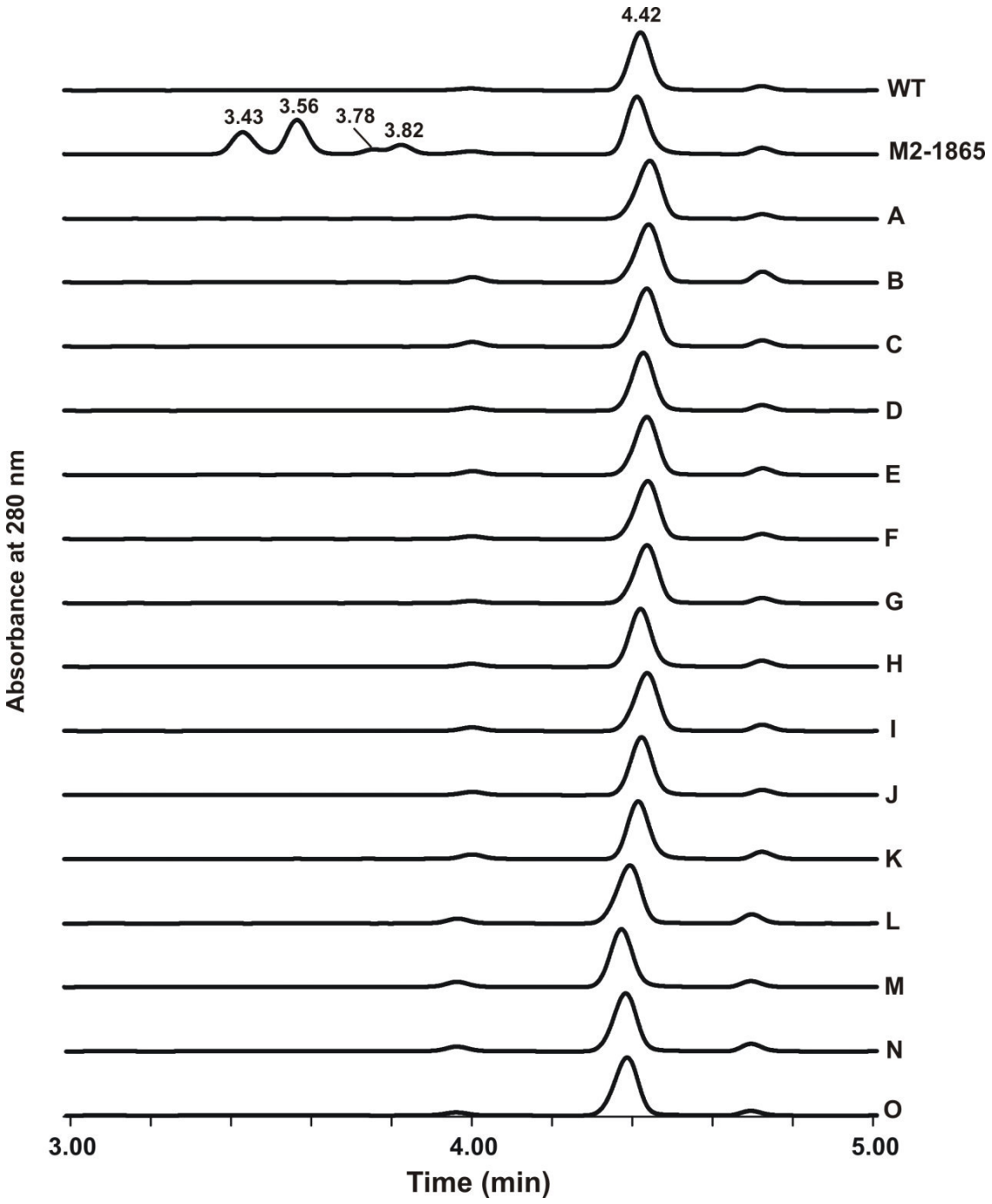
Zhao, J., Zhu, W., Hu, Q. (2001). Enhanced catharanthine production in *Catharanthus roseus* cell culture by combined elicitor treatment in shake flasks and bioreactors. *Enzyme Microb Technol* **28**:673-681.

Zhu, Q., Smith, S.M., Ayele, M., Yang, L., Jogi, A., Chaluvadi, S.R., Bennetzen, J.L. (2012). High-throughput discovery of mutations in tef semi-dwarfing genes by next-generation sequencing analysis. *Genetics* **192**:819-829.

H. Supplementary Information



Supplementary Figure 1: MIA profiles of the M3 generation obtained by self-fertilization of the M2 generation show that M2-1865 is homozygous for the mutation. The LC profiles of LP2 leaf surface extracts from six individual M3 generation plants (**A-F**) compared to WT and M2-1865 plants show accumulation of the tabersonine 2,3-epoxide derivatives in all M3 offspring [16-methoxy-2,3-epoxytabersonine (RT 3.43 and 3.75 min, m/z 383), 2,3-epoxytabersonine (RT 3.56 and 3.82 min, m/z 353), catharanthine (RT 4.42 min m/z 337)]. Quantification of vindoline and the 2,3-epoxides shows that all plants in the M3 generation accumulate similar levels of 2,3-epoxides and similar reduced levels of vindoline as the M2-1865 line.



Supplementary Figure 2: The F1 generation plants of reciprocal crossings between WT and M2-1865 plant lines exhibit WT phenotypes. The leaf surface extract of the M2-1865 line shows accumulation of tabersonine 2,3-epoxide derivatives (RT = 3.43 and 3.78 min, m/z = 383; RT = 3.56 and 3.82 min, m/z = 353) that do not accumulate in leaf surface extract of the WT line or in 15 individual F1 generation offspring of WTm x M2f (**A-K**) or WTf x M2m (**L-O**) crossings [catharanthine (4.42 min)].

Supplementary Table 1: Primer sequences used for real-time PCR.

Target	**Forward (5' - 3')**	**Reverse (5' - 3')**
60S	TCTTAGTTGGAATGTTCAGCACCTG	CAAGGTTGGAGCCCCTGCTCGTGTT
STR	CCAAGATGGCCGAGTTATCA	TTTTCTCTGGATCGGTGCTG
SGD	ATTTGCACCAGGAAGAGGTG	TATGAACCATCCGAGCATGA
T16H	GCTTCATCCACCAGTTCCAT	CCGGACATATCCTTCTTCCA
16OMT	GGATTCTCCATGACTGGAAC	GCCCATTCTTTCTCACATCT
T3O	CCCTGCAGTTCCATTTCAATCTCG	TGGCGTAATAGGCAAGCACGATTC
T3R	CGCGAGTACGGGTGGAAGTATAAA	CGGGGATAACCTCAACATCTGCAA
NMT	GGTGCTGTATCACTGTGTTCT	ATTTTCGTCCCGTAACTACA
D4H	TTGACATTTGGGACAAAGCAA	CCAAAAGCAACAGCAACAGA
DAT	GTGCGTATCCGTTGGTTTCT	CGAACTCAATTCCATCGTCA

Supplementary Table 2: Species and accession IDs of proteins used in the phylogenetic analysis (Figure 18).

Species	Protein	NCBI Accession
Catharanthus roseus	CroT3R	
	Cro10HGO	Q6V4H0
Abies holophylla	AhoSAD1	ADP68543.1
	AhoSAD2	ADP68544.1
Arabidopsis thaliana	AthCAD1	AAP68279.1
	AthCAD2	NP_179765.1
	AthCAD3	NP_179780.1
	AthCAD4	AY302081.1
	AthCAD5	2CF5
	AthCAD6	NP_195510.1
	AthCAD7	AAK93608.1
	AthCAD8	NP_195512.1
	AthCAD9	NP_195643.1
Artemisia annua	AanCAD	EU417964.1
Brachypodium distachyon	BdiCAD3	AFK80371.1
	BdiCAD5	AFK80372.1
Camellia sinensis	CsiCAD1	AEE69007.1
	CsiCAD2	ADO51749.1
	CsiCAD3	AEE69008.1
Cathaya argyrophylla	CarSAD1	ADP68545.1
	CarSAD2	ADP68546.1
Cedrus deodara	CdeSAD1	ADP68547.1
	CdeSAD2	ADP68548.1
Chamaecyparis pisifera	CpiSAD	ADP68554.1
Citrus sinensis	CitCAD	ABM67695.1
Eriobotrya japonica	EjaCAD1	ABV44810.1
	EjaCAD2	ABV44811.1
Eucalyptus camaldulensis	EcaCAD	ACX68560.1
Eucalyptus globulus	EglCAD	AAC07987.1
Eucalyptus gunnii	EguCAD1	CAA61275.1
	EguCAD2	X65631.1
Festuca arundinacea	FarCAD1b	AAK97809.1
Gingko biloba	GbiSAD	ADP68557.1
Hibiscus cannabinus	HbaCAD	ADK24218.1
Leucaena leucocephala	LleCAD	CAJ90658.1
Medicago sativa	MsaCAD	CAA79625.1

Medicago truncatula	MtrCAD2	4QTZ
Metasequoia glyptostroboides	MglSAD	ADP68558.1
Nicotiana tabacum	NtaCAD1	AY911854.1
	NtaCAD19	X62344.1
	NtaSAD	ABD73280.1
Ocimum basilicum	ObaCAD	AAX83108.1
Oryza sativa	OsaCAD1	NP_001064283.2
	OsaCAD2	Q6ZHS4.1
	OsaCAD3	NP_001064659.2
	OsaCAD4	NP_001068303.2
	OsaCAD5	Q0J6T3.2
	OsaCAD6	NP_001052290.1
	OsaCAD7	Q0JA75.1
	OsaCAD8A	Q6ERX1.1
	OsaCAD8C	Q6ERW7.2
	OsaCAD8D	Q6ERW5.1
Panicum virgatum	PviCAD2	ADO01602.1
Picea abies	PabCAD2	CAA05095.1
Pinus armandii	ParCAD	ADP68542.1
	ParSAD1	ADP68549.1
	ParSAD2	ADP68550.1
Pinus banksiana	PbaSAD1	ADP68556.1
	PbaSAD2	ADP68551.1
Pinus massoniana	PmaCAD	AHL67653.1
Pinus radiata	PraCAD	AAC31166.1
Pinus taeda	PtaCAD2	CAA86072.1
	PtaCAD1	CAA86073.1
Platycladus orientalis	PorSAD	ADP68559.1
Populus tremuloides	PtrSAD	1YQD
	PtrCAD	AF217957.1
Saccharum officinarum	SofCAD	O82056.1
Sorghum bicolor	SbiCAD2	ACL80889.1
Thujopsis dolabrata	TdoSAD	ADP68555.1
Triticum aestivum	TaeCAD	ADI59734.1
Tsuga dumosa	TduSAD1	ADP68552.1
	TduSAD2	ADP68553.1
Zea mays	ZmaCAD2	CAA06687.1