Karlsruher Institut für Technologie

Reconstructing phenylpropanoid pathway of a living fossil "*Cephalotaxus hainanensis*" in *Nicotiana tabacum*

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

(Dr. rer. nat.)

von der KIT-Fakultät für Chemie und Biowissenschaften
des
Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

M. Sc. Nasim Reshadinejad

Referent: Prof. Dr. Peter Nick

Korreferent: Prof. Dr. Tilman Lamparter

Tag der mündlichen Prüfung: 17.02.2023

Eidesstattliche Erklärung

Hiermit erkläre ich, die vorliegende Dissertation, von der Verwendung der angegebenen Hilfsmittel abgesehen, selbständig verfasst zu haben.

Alle Stellen, die gemäß Wortlaut oder Inhalt aus anderen Arbeiten entnommen sind, wurden durch Angabe der Quelle kenntlich gemacht.

Diese Dissertation liegt in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vor.

Karlsruhe, Januar 2023

Nasim Reshadinejad

Die vorliegende Arbeit wurde am Botanischen Institut des Karlsruher Instituts für Technologie (KIT), Lehrstuhl I für molekulare Zellbiologie, im Zeitraum von Mai 2018 bis Januar 2023 angefertigt.

Acknowledgment

I feel a great debt of appreciation to all of the people who supported me in pursuing my doctorate. The help I've received from so many individuals, both directly and indirectly, has made this process simpler than I anticipated.

First of all, I owe a huge debt of gratitude to my benevolent PhD supervisor, Professor Dr. Peter Nick, who has been supporting me not only in science with his perseverance, inspiration, and vast scientific knowledge, but also in my personal life by teaching me how to deal with the most difficult situations in the simplest way. I appreciate him giving me the opportunity to work in his lab, as well as his advice, which was extremely useful throughout the entire research and thesis-writing process. I could not have asked for a better supervisor for my doctoral studies.

Furthermore, I'd like to express my gratitude to Dr. Michael Riemann and Dr. Jathish Ponnu for their precious advice and feedback throughout this study.

I would like to express my appreciation to Dr. Sergi Afonin from the department of organic chemistry, for all of his assistance and support with relation to my HPLC analysis.

I consider myself extremely fortunate to have worked with such wonderful people in Nick-Lab such as Sabine and Ernst, and to have met wonderful friends like Eman, Toranj, Lucas, Noemi, Natalie, Christian, Islam and Kai who have made the last five years much more enjoyable and kept me happy throughout the entire process.

Last and foremost, I would like to thank and dedicate this work to;

My entire life, my love, my friend and my husband, Iloush, who brightens and warms my heart and life. For all of his patience, support, and encouragement not only throughout these five years of study, but also during the entire experience of being together and enjoying each other's presence known as love. I couldn't be more grateful and glad to have you as a partner and friend to share my world with, during this long journey named life, because I believe nothing makes sense without you. I love you to the moon and back.

My mother, for giving me life, for never giving up on me, for always encouraging me to do my best, for showing me how to fight for what I want, for teaching me how to be a good person and how to love for having a meaningful life.

My father who passed away during this study; I know how important it was to be with you during your last moments but I hope you accept this humble effort as an apology, I am sure it will make you happy and proud. I miss you a lot.

Nasim

January 2023

Abstract

Cephalotaxine alkaloids have been shown to impede the growth of several human tumor cell lines. Homoharringtonine, an esterified form of Cephalotaxine approved in 2012 for the treatment of chronic myeloid leukemia, is an extremely expensive pharmaceutical compound due to the scarcity of its raw material, the plant genus *Cephalotaxus* (Cephalotaxaceae), which is the only source of this type of natural product. The principal source of Homoharringtonine is *Cephalotaxus hainanensis,* an endangered species of the *Cephalotaxus* genus. Because of its broad range of herbal medicinal capabilities, this extremely slow-growing tree, which is so prominent in traditional Chinese medicine, is on the danger of extinction due to illegal over-exploitation. Due to the limited availability of Homoharringtonine and the rising demand for this drug, alternate methods of synthesis are required. A suitable alternative for producing this pricey compound would be metabolic engineering and heterologous production of this natural substance in a suitable biological host. However, in order to logically design a plant metabolic pathway for future applications or the metabolic engineering of 1-phenethylisoquinoline scaffold (the general and unique precursor for Cephalotaxine alkaloids), a thorough understanding of the structural and functional characterization of enzymes involved in this pathway, metabolite quantification, and signal transduction pathway discovery in vivo would be required. As a result, we narrowed down our investigation on the sub-cellular localisation of the enzymes involved in the 1-phenethylisoquinoline scaffold synthesis process. The sub-cellular localisation results of this study revealed the presence of two distinct pathways for the production of Cinnamoyl CoA, one of which occurs exclusively in plastids and the other of which takes place between the cytosol, ER, and peroxisomes. The effectiveness of the mechanism by which metabolites from this chemical are guided into various secondary metabolite production pathways would be increased as a result of these two unique routes. The findings of this study also demonstrated that compartmentalization is a mechanism to steer the pathway into different directions notwithstanding of transient metabolon formation for facilitating the transportation among different cellular compartment. Therefore, to proof of this concept we tried to redirect the production of *Nicotiana tabacum* cells overexpressing *Ch.*PAL by altering the cellular compartment localization. Then we presented evidences regarding modifying cellular compartmentalization in *Nicotiana tabacum* cells overexpressing *Ch.*PAL genes using Brefeldin A which resulted in lowering the concentration of side metabolites like Ferulic acid and conducting the flux towards a desired pathway. This study's findings about the sub-cellular localisation of the enzymes participating in this pathway will undoubtedly aid future research into how to direct this extremely dynamic metabolic flux toward preferred side routes or potential intermediates.

Zusammenfassung

Cephalotaxin-Alkaloide hemmen das Wachstum mehrerer menschlichen Tumorzelllinien. Homoharringtonin, eine veresterte Form von Cephalotaxin, die 2012 zur Behandlung von chronischer myeloischer Leukämie zugelassen wurde, ist aufgrund der Knappheit seines Rohstoffs *Cephalotaxus*, der die einzige Quelle für diese Art von Naturprodukt darstellt, ein äußerst teurer pharmazeutischer Wirkstoff. Die Hauptquelle von Homoharringtonin ist *Cephalotaxus hainanensis*, eine vom Aussterben bedrohte Art der Gattung *Cephalotaxus*. Dieser extrem langsam wachsende Baum, der in der traditionellen chinesischen Medizin so prominent ist, ist aufgrund seines breiten Spektrums an pflanzlichen medizinischen Fähigkeiten aufgrund illegaler Übernutzung vom Aussterben bedroht. Aufgrund der begrenzten Verfügbarkeit von Homoharringtonin und der steigenden Nachfrage nach diesem Medikament sind alternative Synthesemethoden erforderlich. Eine geeignete Alternative zur Herstellung dieser teuren Verbindung wäre das Metabolic Engineering und die heterologe Herstellung dieser natürlichen Substanz in einem geeigneten biologischen Wirt. Um jedoch einen pflanzlichen Stoffwechselweg für zukünftige Anwendungen oder das metabolische Engineering des 1-Phenethylisochinolin-Gerüsts (der allgemeinen und einzigartigen Vorstufe für Cephalotaxin-Alkaloide) logisch zu entwerfen, ist ein gründliches Verständnis der strukturellen und funktionellen Charakterisierung von Enzymen, die an diesem Weg beteiligt sind, Quantifizierung von Metaboliten und die Entdeckung des Signaltransduktionswegs in vivo wären erforderlich. Als Ergebnis haben wir unsere Untersuchung auf die subzelluläre Lokalisierung der Enzyme eingegrenzt, die am 1-Phenethylisochinolin-Gerüstsyntheseprozess beteiligt sind. Die subzellulären Lokalisierungsergebnisse dieser Studie zeigten das Vorhandensein von zwei unterschiedlichen Wegen für die Produktion von Cinnamoyl-CoA, von denen einer ausschließlich in Plastiden auftritt und der andere zwischen dem Cytosol, ER und Peroxisomen stattfindet. Die Wirksamkeit des Mechanismus, durch den Metaboliten dieser Chemikalie in verschiedene sekundäre Stoffwechselwege geleitet werden, würde als Ergebnis dieser beiden einzigartigen Wege erhöht werden. Die Ergebnisse dieser Studie zeigten auch, dass die Kompartimentierung ein Mechanismus ist, um den Weg ungeachtet der vorübergehenden Metabolonbildung in verschiedene Richtungen zu lenken, um den Transport zwischen verschiedenen Zellkompartimenten zu erleichtern. Um dieses Konzept zu beweisen, versuchten wir daher, die Produktion von *Nicotiana tabacum*-Zellen, die *Ch*.PAL überexprimieren, umzuleiten, indem wir die Lokalisierung des Zellkompartiments verändern. Dann präsentierten wir Beweise zur Modifizierung der zellulären Kompartimentierung in Nicotiana tabacum-Zellen, die *Ch*.PAL-Gene unter Verwendung von Brefeldin A überexprimieren, was zu einer Verringerung der Konzentration von Nebenmetaboliten wie Ferulasäure und zur Lenkung des Flusses zu einem gewünschten Weg

führte. Die Ergebnisse dieser Studie über die subzelluläre Lokalisierung der an diesem Stoffwechselweg beteiligten Enzyme werden zweifellos die zukünftige Forschung darüber unterstützen, wie dieser äußerst dynamische Stoffwechselfluss auf bevorzugte Nebenwege oder potenzielle Zwischenprodukte gelenkt werden kann.

Table of Contents

Abbreviations

PAL: Phenylalanine ammonia-lyase

C4H: Cinnamate 4-hydroxylase

4CL: 4-Coumarate CoA ligase

DBR: NADPH-dependent double bond reductases

CCR: Cinnamoyl CoA reductase

TyDC: Tyrosine decarboxylase

DODC: DOPA decarboxylase

PPO: Polyphenol oxidase

Pr10: Pathogenesis-related 10/Bet v1 proteins

PSS: Phenethylisoquinoline scaffold synthase

C3H: 4-Coumarate 3-hydroxylase

COMT: Caffeic acid 3-O-methyltransferase

4-HDCA: 4-hydroxydihydrocinnamaldehyde

Ch: *Cephalotaxus hainanensis*

HPLC: High pressure liquid chromatography

EF: Elongation Factor

L-Phe: L-Phenylalanine

BY-2 WT: *Nicotiana tabacum* wild type

1. Introduction

1.1. Medicinal plants

Medicinal plants have a history of healing as old as the human being. Numerous sources provide sufficient evidence which show this search for drugs in nature origins in far past, and is a response to years of tackling illnesses. Medicinal herbs can commonly be found in barks, seeds, fruit bodies, and other parts of a plant. Contemporary science, or more specifically pharmacotherapy, is included a wide range of drugs with plant origin which were used by ancient civilizations (Petrovska, 2012). China and India are among the countries with a long history of plant medicines as the core system of treatment for thousands of years. Even nowadays, medicinal plants comprise a promising source of essential drugs for the majority of people on Earth. One may refer to herbs as a commonplace method of treating illnesses in societies with less industrialized movement (Sofowora et al., 2013). According to a report by the World Health Organization (WHO), 80 percent of people in most Asian and African countries refer to herbal medicines as primary health care system (Oyebode et al., 2016). As opposed to the high expense of pharmaceuticals, half of the people in the world who live on less than $2 per day may find it more affordable to either grow medicinal plants or gather them from nature (Sofowora et al., 2013). Many modern-day medicines have roots in herbal remedies such as opium, digitalis, and quinine. However, the quest for drugs and supplements with plant origin has accelerated during the last decade. According to a report by WHO about 25% of current drugs in the U.S. have been extracted from plants (Fabricant & Farnsworth, 2001).

1.2. Plant secondary metabolites

The combinations of secondary metabolites which are taxonomically distinct and are unique to particular plant species leads to the medicinal effects of each plant (John, 2000). Around10^6 specific known secondary metabolites (Saito & Matsuda, 2010). which are generated in plants can have many various roles such as fighting herbivory, pathogen attack, inter-plant competition, and attracting organisms such as pollinators or symbionts (Becker, 2000)and therefore, are pharmaceutically active (Qiao et al., 2014). These chemically diverse secondary products are generally classified into three groups of phenolics, terpenes and steroids, and alkaloids (Bourgaud et al., 2001), which among all, the last group “alkaloids” is the subject of this study.

1.3. Alkaloids

Alkaloids are naturally occurring organic compounds containing at least one nitrogen atom. They are secondary metabolites that are pharmaceutically active and contain at least one nitrogen atom (Ziegler & Facchini, 2008).

They are mainly derived from amino acids like Phenylalanine, Tyrosine, Tryptophan, Lysine and Arginine. Approximately 20% of the flowering plants contain alkaloids; more than 12,000 of those belong to over 150 families. Interesting to know that alkaloids are not unique to plant kingdom but they also could be found in nearly all kinds of creatures, including frogs, ants, butterflies, bacteria, sponges, fungi, spiders, beetles, and mammals but it doesn´t specifically mean that alkaloids present in animals are always being generated there. They could be sometimes taken up by other organisms, for instance insects employ plant alkaloids as a source of attractants, pheromones, and defensive compounds. Some animals, including a frog named *Bufo marinus* create poisonous alkaloids in its skin or secretory glands (Conn, 1981).

The role of alkaloids in plants are not yet well understood. Despite evidence to the contrary, it has been proposed that they are merely metabolic byproducts of plants´ metabolic processes but scientific evidence showed that they may have a role in plants specialized biological activities since in some plants, alkaloids concentration rises immediately before seed development and subsequently decreases once the seed is mature (Mann & Mudd, 1963). Some plants may also be protected by alkaloids from being harmed by specific insect species. it is also likely that they deter fungal, bacterial, or insect infestations in some cases (Conn, 1981).

A significant number of alkaloids are identified as natural compounds that provide a rich source of diverse, potent therapeutic chemical scaffolds with a broad range of biological activities (Herrendorff et al., 2016; K.M. Chaves et al., 2016). Plant alkaloids or their semi-synthetic derivatives have led to a substantial number of drugs in modern medicine (Amirkia & Heinrich, 2014). For instance, the antimitotic and antigout colchicine from *Gloriosa superba*, the anticancer vinblastine and vincristine from *Catharanthus roseus*, the analgesic morphine from *Papaver somniferum*, the antidiabetic chelerythrine from *Chelidonium majus*, and the antimalarial quinine from *Cinchona officinalis*. (Achan et al., 2011; Beaudoin & Facchini, 2014; Capistrano et al., 2016; Kuo et al., 2015; Sivakumar, 2013; Yu et al., 2015; Zheng et al., 2015). In Table 1.1 you will see a list of some physiologically active alkaloids in plants and their use in modern medicine.

Alkaloid	Plant source	Use
Ajmaline	*Rauwolfia serpentina*	Antiarrhythmic that functions by inhibiting glucose uptake by heart tissue mitochondria
Atropine, (±) hyoscyamine	*Hyoscyamus niger*	Anticholinergic, bronchodilator
Caffeine	*Coffea arabica*	Widely used central nervous system stimulant
Camptothecin	*Camptotheca acuminata*	Potent anticancer agent
Cocaine	*Erythroxylon coca*	Topical anesthetic, potent central nervous system stimulant, and adrenergic blocking agent; drug of abuse
Codeine	*Papaver somniferum*	Relatively nonaddictive analgesic and antitussive
Coniine	*Conium maculatum*	First alkaloid to be synthesized; extremely toxic, causes paralysis of motor nerve endings, used in homeopathy in small doses
Emetine	*Uragoga ipecacuanha*	Orally active emetic, amoebicide
Morphine	*Papaver somniferum*	Powerful narcotic analgesic, addictive drug of abuse
Nicotine	*Nicotiana tabacum*	Highly toxic, causes respiratory paralysis, horticultural insecticide; drug of abuse
Pilocarpine	*Pilocarpus jaborandi*	Peripheral stimulant of the parasympathetic system, used to treat glaucoma
Quinine	*Cinchona officinalis*	Traditional antimalarial, important in treating *Plasmodium falciparum* strains that are resistant to other antimalarials
Sanguinarine	*Eschscholzia californica*	Antibacterial showing antiplaque activity, used in toothpastes and oral rinses
Scopolamine	*Hyoscyamus niger*	Anticholinergic, effective against motion sickness
Strychnine	*Strychnos nuxvomica*	Violent tetanic poison, rat poison, used in homeopathy
(+)-Tubocurarine	*Chondrodendron tomentosm*	Nondepolarizing muscle relaxant producing paralysis, used as an adjuvant to anesthesia
Vincristine	*Catharanthus roseus*	Antineoplastic used to treat childhood leukemia and other cancers.

Table 1.1. Example of some physiologically active alkaloids in plants and their use in in modern medicine

1.4. Alkaloids´classifications

Alkaloids can be divided into groups based on different parameters. One of them is the biological system in which they are found. Opium alkaloids, for example, are found in the opium poppy (*Papaver somniferum*). The most important and frequent parameter for categorizing alkaloids is their chemical structures. Based on that they can be divided into groups according to the chemical structures of their common central intermediates in the biosynthetic pathways and their biosynthetic precursors, such as the amino acids that supply nitrogen atoms in their skeleton. This has led to the development of various subcategories of alkaloids, including Monoterpenoid indole alkaloids (MIA), Benzylisoquinoline alkaloids (BIA), Tropane alkaloids and Nicotine, Purine alkaloids, Pyrrolizidine alkaloids, and Phenlyethylisoquinoline alkaloids (PIAs), whose members share similar biosynthetic starting precursors as well as structurally similar central intermediates.

1.4.1. Monoterpenoid indole alkaloids (MIA)

In monoterpenoid indole alkaloids (MIA) which represents one of the largest classes of natural products, tryptophane is the common amino acid precursor. Their complex synthesis involves the assembly of tryptamine (TDC/from tryptophan) with a secologanine to produce a central intermediate named strictosidine which is the common biosynthetic intermediate for all monoterpene indole alkaloid enzymatic pathways. *Catharanthus roseus*, *Rauvolfia serpentine*, and *Cinochona officinalis* are the main sources of monoterpenoid indole alkaloids (MIA). This category includes four well-known alkaloids: vinblastine, vincristine, ajmaline, and quinine.(Lazar, 2005)

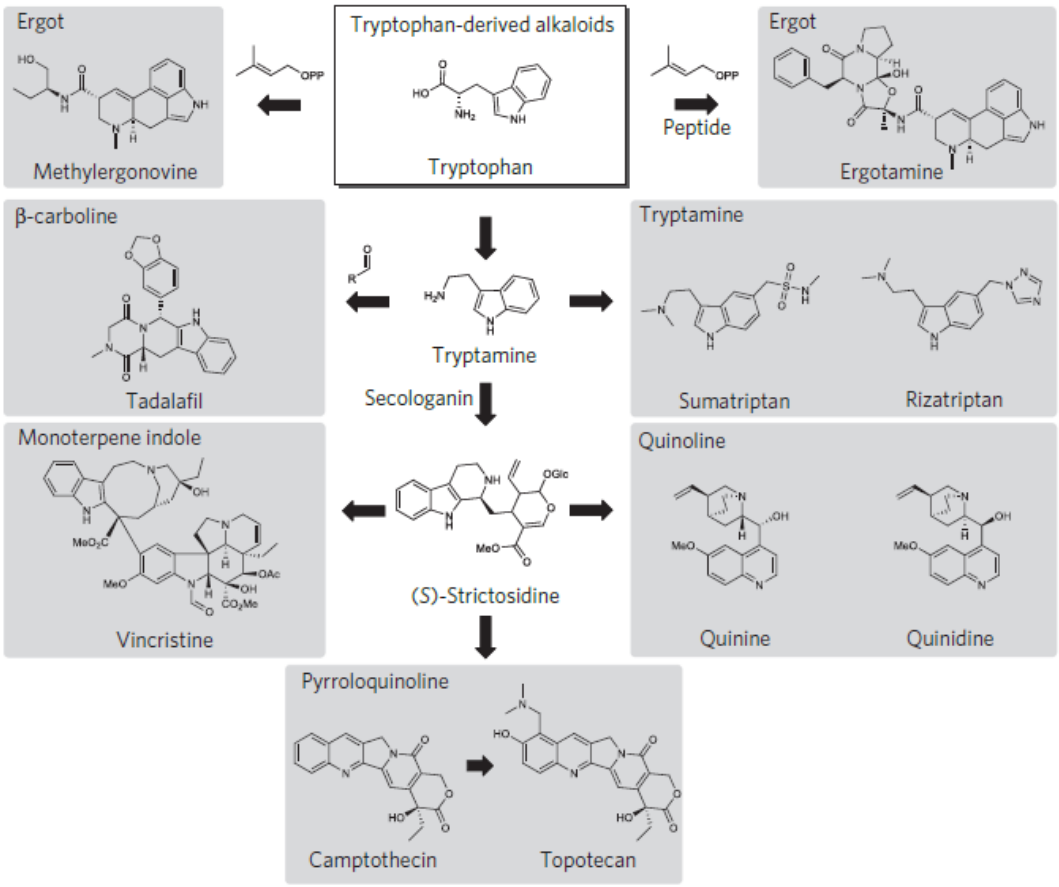


Figure 1.1. Representative tryptophan-derived FDA-approved alkaloid-based pharmaceuticals and key biosynthetic intermediates discussed in this perspective.(Ehrenworth & Peralta-Yahya, 2017)

1.4.2. Benzylisoquinoline alkaloids´ (BIA)

In Benzylisoquinoline alkaloids´ (BIA) feeding experiments, on the other hand, the amino acid (S)-Tyrosine was identified as the primary metabolic precursor by isotope labeling of predicted

intermediates. Tyramine and L-Dopa were two Tyrosine derivatives that were condensed to create Dopamine during the production of BIA. The precursor for the Isoquinoline moiety is Dopamine, while the Benzyl component is 4-Hydroxyphenylacetaldehyde, which is produced when Tyramine is deaminated. The first common step in the biosynthesis of the Benzylisoquinoline alkaloid (BIA) is (S)-Norcoclaurine, which gains two O-methyl groups on its way to the key intermediate (S)-reticuline which is the common branchpoint intermediate in the biosynthesis of BIAs. Most often, the *Papaveraceae*, *Ranunculaceae*, *Berberidaceae*, and *Menispermaceae* families contain Benzylisoquinoline alkaloids. This category of alkaloids includes narcotic painkillers like codeine and morphine, the cough suppressant papaverine, and the antimicrobials sanguinarine and berberine.

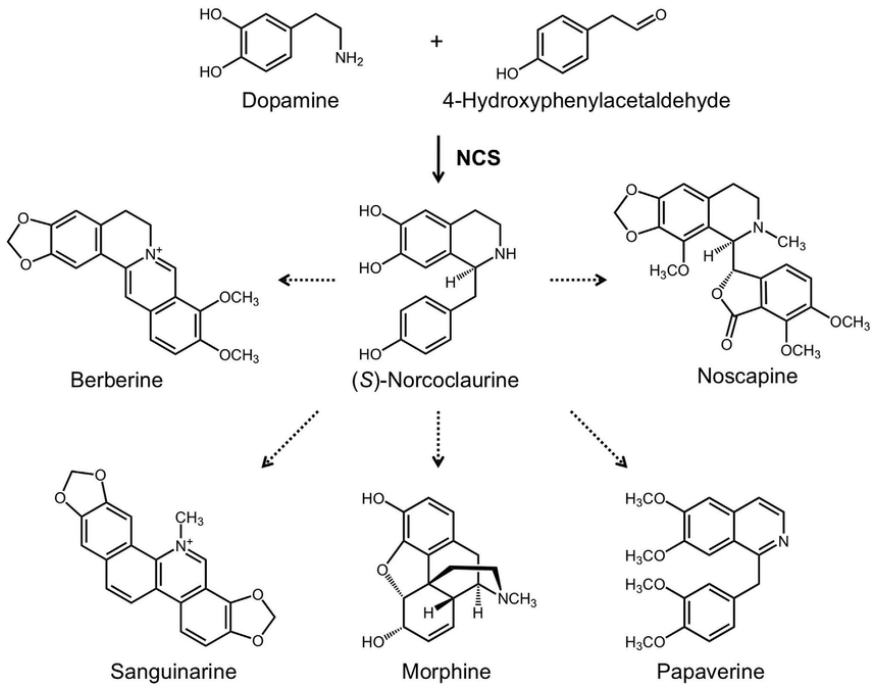


Figure 1.2. The biosynthesis of structurally diverse Benzylisoquinoline alkaloids such as berberine, sanguinarine, morphine, noscapine, and papaverine begins with the condensation of Dopamine and 4-Hydroxyphenylacetaldehyde by NCS (Norcoclaurine synthase) which synthetize (S)-Norcoclaurine, the first common step in the biosynthesis of the Benzylisoquinoline alkaloid (BIA)(J. Li et al., 2016)

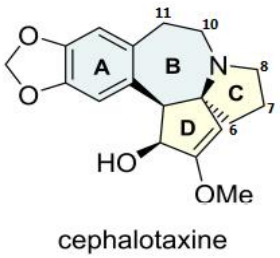
1.4.3. Phenylethylisoquinoline alkaloids (PIA)

In Phenylethylisoquinoline alkaloids (PIA) the biosynthetic precursors are Phenylalanine and Tyrosine and their common synthetic scaffold intermediate is 1-phenethylisoquinoline. Alkaloids that fall under this category have a strong pharmacological potential, particularly in the treatment of

cancer, inflammation, and cardiovascular disease. Well-known substances, such as colchicine, provide as examples of their therapeutic effects. Since more than 2000 years ago, the famous alkaloid colchicine, which is mostly found in plants of the Liliaceae family, has been utilized in traditional medicine (Hartung, 1954; Stander et al., 2021) and today is a well-known substance used to treat acute gout and familial Mediterranean fever, a hereditary inflammatory condition, as well as for its antimitotic properties. It is Commercially separated from the closely related ornamental plants flame lily (*Gloriosa superba L.*) and autumn crocus (*Colchicum autumnale L.*)(Larsson & Ronsted, 2013).

Cephalotaxine, a class of alkaloids exclusive to the family Cephalotaxaceae, is another class of alkaloids generated from phenethylisoquinoline precursors (Abdelkafi & Nay, 2012).

The biosynthesis of Cephalotaxine from two molecules of Tyrosine and Phenylalanine was confirmed according to comprehensive isotope labeling studies with Carbone-14 conducted by Parry and colleagues in 1981 (Abdelkafi & Nay, 2012; Parry et al., 1980). They provided proof that Cephalotaxine belongs to the family of Phenylethylisoquinoline alkaloids, as do Schelhammera and Colchicum alkaloids and that C-6, C-7 and C-8 of Cephalotaxine originated from C-3, C-2 and C-1, respectively, of the Phenylalanine side chain and that C-10 and C-11 of the alkaloid was derived from C-2 and C-3 of the Tyrosine side chain. They also proposed that ring D may have been derived from the aromatic ring of Phenylalanine by ring contraction (Abdelkafi & Nay, 2012; Parry et al., 1980). Natural products within the Cephalotaxine-type alkaloid class are characterized by promising synthetic and biological properties.



cephalotaxine

Figure 1.3. Cephalotaxine structure

1.5. Cephalotaxine and its esterified form alkaloids

Cephalotaxine was the first natural product from this family to be reported (Figure 1.3.). it is the maternal alkaloid of Cephalotaxine-type alkaloids with a unique structure, because of their intriguing pentacyclic structures, and great pharmaceutical background. Cephalotaxine has now been proven to have no biological activity itself but it´s esterified forms indicate great antileukemic properties. Cephalotaxine was initially isolated in 1963 from *Cephalotaxus fortunei* and then *Cephalotaxus harringtonia* which are two species of genus *Cephalotaxus* (Paudler et al., 1963). The structure of its ester alkaloids, most noticeably Harringtonine, Isoharringtonine, Deoxyharringtonine, and Homoharringtonine (HHT) from *Cephalotaxus harringtonia* extract was

identified in 1969. (Powell et al., 1970). To date, at least 70 compounds from this family have been isolated and identified that can be divided into two categories (Cephalotaxine-type and Homoerythrine-type) (Abdelkafi & Nay, 2012).

1.6. Homoharringtonine, one of the esterified forms of Cephalotaxine

Most recently, it was determined that Homoharringtonine-based induction therapies were superior than daunorubicin and cytarabine, which were formerly regarded as the gold standard for induction chemotherapy, in treating acute myeloid leukemia (Jin et al., 2013). In 2012, the US Food and Drug Administration (FDA) authorized the use of Homoharringtonine (omacetaxine mepesuccinate, Synribo®) for the treatment of patients with chronic myelogenous leukemia (CML), whose cancer had grown resistant to at least two Tyrosine kinase inhibitors (TKIs). Homoharringtonine's ability to stop the initial elongation step of protein synthesis served as the basis for its mode of action. It binds to the 60S ribosome's ribosomal A-site and prevents aminoacyl-tRNAs from approaching the ribosome (FRESNO et al., 1977; Huang, 1975). As a result, numerous oncogenic short-lived proteins are rapidly depleted, which induces Apoptosis in cancer cells. Additionally, several *Cephalotaxus* alkaloids have been shown to be effective against human parasites such Plasmodium falciparum and Leishmania major (Morita et al., 2010) as well as epidermoid carcinoma (Morita et al., 2002), lymphoma (Yoshinaga et al., 2004), and nasopharynx carcinoma (Bocar et al., 2003).

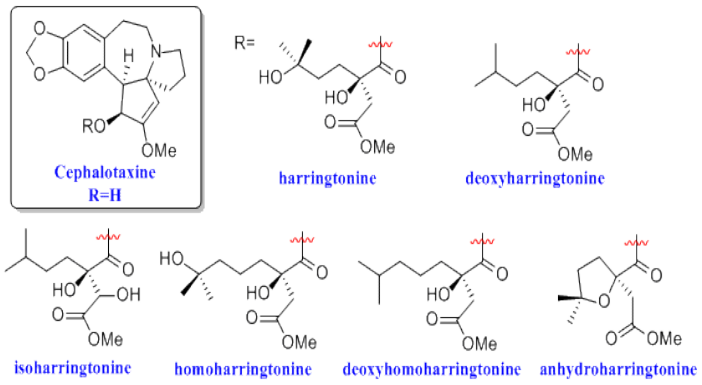


Figure 1.4. Structure of Cephalotaxine and various R groups in different Cephalotaxine esterified forms

1.7. Genus *Cephalotaxus*

Cephalotaxine alkaloids as said in 1.5 are only found in the plant genus *Cephalotaxus*. As members of the Cephalotaxaceae family, *Cephalotaxus*, also known as plum yews, are dioecious, evergreen coniferous trees or shrubs (formerly included in the Taxaceae) with six to twelve

species and variations, depending on the taxonomic descriptions. This genus is typically regarded as the sole representative of its botanical family. One of the members of this family, *Cephalotaxus harringtonia* (formerly *Cephalotaxus drupacea*; Japanese name: Inukaya), was discovered in Japan and transported to Europe in 1829 by the scientist Von Siebold, the writer of Flora Japonica book. These species of this family comprise 1–10 meter tall (in some cases to 20 meter) evergreen shrubs which are widespread throughout southern and eastern Asia, specifically in China, including the islands of Taiwan and Hainan, as well as in India, Japan, Korea, Laos, Myanmar, Thailand, and Vietnam. Due to a combination of significant human impact as a result of excessive exploitation, a lengthy seed development period (up to two years), or a poor rate of wild regeneration, some species have been classified as endangered, particularly in China (Abdelkafi & Nay, 2012).

Two-ranked needles with a similar form to those of real yews and fleshy covered seeds with the size and shape of an olive or a tiny plum, which turn from blue-green to purple-brown due to ripening, are common morphological traits of the Cephalotauxus plants. Given that many of these morphological characteristics are expected to change with plant age and environment, the nomenclature of this genus is particularly confusing and is likely to remain a challenge. This explains why, despite the fact that genomic analysis may enhance classification, the exact number of species is still up for debate.(Miao et al., 2012; Wang et al., 2016). The seven species listed below however, were defined by Lang et al in 2013: *Cephalotaxus* fortunei Hooker (1850: 76) (as 'fortuni'), *Cephalotaxus griffithii* Hooker (1888: 648), *Cephalotaxus alpina* (Li) Fu (1984: 282), *Cephalotaxus oliveri* Masters (1898: 270), *Cephalotaxus nana Nakai* (1919: 193), *Cephalotaxus harringtonii* (Knight ex J.Forbes) Koch (1873: 102), *Cephalotaxus hainanensis* Li (1953: 164) (LANG et al., 2013).

1.8. *Cephalotaxus hainanensis*

As a member of *Cephalotaxus* genus, *Cephalotaxus hainanensis*, is exclusive to the tropical island of Hainan and is an evergreen conifer tree with a 20–25 m height range. Due to its antileukemia properties, *Cephalotaxus hainanensis* is a common herbal treatment in China (Han, 1994). This species was threatened by the commercial exploitation of the bark and leaves. Its decline has historically been attributed to logging, but since logging is no longer permitted in Hainan, the pressure is now principally coming from illegal harvesting for Chinese medicines (Yi et al., 2013).

That is the reason that nowadays this plant is regarded as an endangered species in China due to its slow growth and excessive exploitation for lumber and medicinal uses (Qiao et al., 2014).

1.9. Obstacles in providing Homoharringtonine for consumers

The demand for Homoharringtonine on a global scale, as well as its resource, the *Cephalotaxus* species, have led to a crisis due to rising consumer populations and declining renewable resources. This could explain the overwhelming demand for *Cephalotaxus* biomass due to the process used to prepare its derivative pharmaceuticals. On the other hand, the traditional process for the purification of Homoharringtonine has caused significant environmental harm, and factors like the difficulty of production, the unreliable source of supply, the toxicity of the original dose schemes, and the requirement for a sizable quantity of the tree (*Cephalotaxus* trees, which are scarcely found in China), have all prevented its progress. Except that, the traditional extraction is limited in several other ways, including low efficiency, high prices, and being time-consuming, while chemical synthesis also faces difficulties due to structural complexity and the use of environmentally harmful chemicals. Although it is primarily conceivable to chemically synthesize this uncommon alkaloid (Hudlicky & W. Reed, 2007), this method has not been actively pursued, perhaps due to its poor efficiency.

Nowadays in industrial scale, Homoharringtonine synthesis mostly relies on plant extraction or semi-chemical synthesis employing plant extracts as precursors. Since 1999, it has proven possible to produce Homoharringtonine semi-synthetically by esterifying Cephalotaxine from dried *Cephalotaxus* leaves, which eliminates the need for the pricy barks and reduces the amount of biomass needed by a factor of 70 while producing a purer end product (Robin et al., 1999). Omacetaxine mepesuccinate, also known as ceflatonin, CGX-653, or Myelostat, is the name given to this semi-synthetic Homoharringtonine by ChemGenex Pharmaceuticals Ltd. (Menlo Park, California, USA), in cooperation with StragenPharma (Geneva, Switzerland) (Pérard-Viret et al., 2017). However, the manufacture of semi-synthetic Homoharringtonine from needle biomass faces increasing challenges due to the rising demand for Homoharringtonine. Therefore, a substitute source is required to make up for the uncommon, slow-growing *Cephalotaxus* and meet the expanding demand of the pharmaceutical market for this compound (Lü & Wang, 2014). In order to provide alternate in vivo and in vitro strategies for the production of Homoharringtonine, extensive investigations on pathway elucidation and enzyme discovery would be required.

1.10. Synthetic biology and metabolic engineering of natural products in recent years

Synthetic biology and metabolic engineering have emerged in recent years as a promising option for the efficient large-scale manufacture of targeted medicinal natural products at low cost. Such a strategy depends on bacteria' heterologous expression of metabolic gene clusters to provide novel synthesis abilities. Along with producing the two top-selling drugs hydrocortisone and artemisinic acid (Hommel, 2008; Szczebara et al., 2003), heterologous syntheses of some other alkaloids like strictosidine, noscapine, or morphine in recombinant yeast factories have also been documented (Courdavault et al., 2020a, 2020b). The full identification of the respective biosynthetic pathways at the gene/enzyme levels was thus necessary for these successful implementations of natural product metabolisms, as recently illustrated for vinblastine/vincristine or morphine (Caputi et al., 2018; Hagel & Facchini, 2010; Stander et al., 2021). To enable the alternative manufacture of this priceless molecule, such as heterologous production of this natural product in an appropriate biological host, a complete understanding of the Cephalotaxine metabolic production route would be necessary.

1.11. First steps to the Cephalotaxine alkaloids´ biosynthesis pathway

Cephalotaxine alkaloids´ biosynthesis pathway has received very little attention up until this point. To start investigating about the biosynthetic pathway of a new group of alkaloids at the gene/enzyme level the first step is to look deeply through the production pathway and common intermediates of other alkaloids which were classified under the same group.

As described before Cephalotaxine alkaloids are belong to the group of alkaloids named Phenylethylisoquinoline alkaloids (PIAs). This is why reviewing earlier researches on the colchicine production pathway as a member of the PIA family is definitely worth considering in order to obtain a better understanding of the PIA's biosynthesis pathway.

Colchicine is a well-known drug used to treat acute gout and familial Mediterranean fever, a hereditary inflammatory illness, which found in and named after the autumn crocus (*Colchicum autumnale*, Colchicaceae) is extracted commercially from two closely related ornamental species, the flame lily (*Gloriosa superba*) and autumn crocus (*Colchicum autumnale*) (Larsson & Ronsted, 2013b). In an outstanding new article, Sattely's team has illuminated a significant portion of the biosynthetic pathway for colchicine (Nett et al., 2020). According to isotope labelling experiments (Abdelkafi & Nay, 2012; Parry et al., 1980) they proposed that the biosynthetic pathway of

colchicine was started by the conversion of Phenylalanine to 4-hydroxydihydrocinnamaldehyde (4-HDCA) and Tyrosine to Dopamine. The 1-phenethylisoquinoline scaffold is created by coupling these two products through a Pictet-Spengler reaction. This indicates that the pathway for both PIAs and Benzylisoquinoline alkaloids (BIAs) are sharing common primary steps (Abdelkafi & Nay, 2012; Sheng & Himo, 2019; Xu et al., 2021) but split off from one another at their own shared intermediate scaffold. The difference between two Pictet-Spengler condensation in BIAs and PIAs is coming from the fact that for formation of the intermediate in the case of BIAs, Dopamine and 4-hydroxyphenylacetylaldehyde (4-HPAA) are joined by (S)-Norcoclaurine synthase (NCS) through a Picket-Spengler condensation, whereas Dopamine is joined with 4-HDCA in the case of PIAs to form their common scaffold which is 1-phenethylisoquinoline (Figure 1.4) (Nett et al., 2020; Y. Wang et al., 2022)

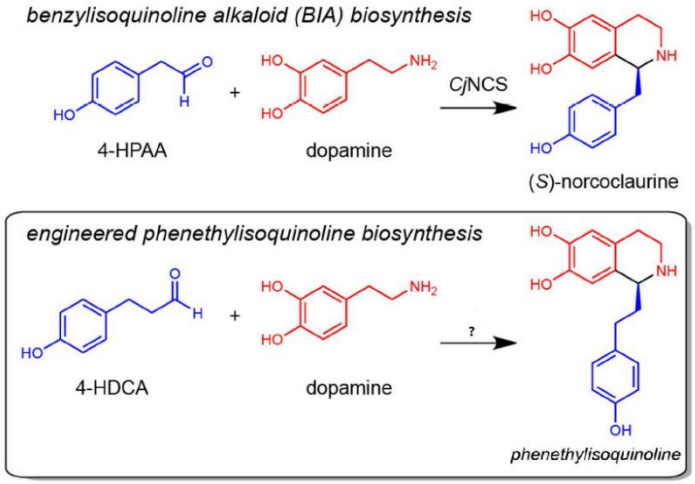


Figure 1.5. Comparison of the native function of *Coptis japonica* NCS (*Cj*NCS) in Benzylisoquinoline alkaloid (BIA) biosynthesis to the reaction required in colchicine alkaloid biosynthesis as a member of PIAs (Nett et al., 2020).

1.12. PSS the responsible enzyme for the Picket-Spengler condensation of PIAs in *Cephalotaxus*

The respective enzyme for the Picket-Spengler condensation of PIAs which led to 1-phenethylisoquinoline was remained elusive until recently and the problem was that the investigations for the homologs of NCS in *Coptis japonica* (CjNCS) producing S-Norcoclaurine in BIAs family, offered up no candidates in the transcriptomes of *Gloriosa superba* (producing Colchicine) (Nett et al., 2020) or *Cephalotaxus hainanensis* transcriptome (NCBI accession No. PRJNA767117), indicating that the Picket-Spengler condensation in those species may be driven by enzymes influenced by evolutionary processes.

By assuming that the biosynthesis of 4-HPAA and 4-HDCA should be comparable due to their structural similarities, and that their fusing into the 1-phenethylisoquinoline scaffold should likewise be accomplished by a similar enzymatic Picket-Spengler condensation driven by an NCS-like protein, Zhao et al. 2022, used a novel strategy, by focusing on critical three-dimensional structure features rather than overall sequence homology in the first place, to identify this enzyme.

Since NCS is a member of the Pathogenesis-related 10/Bet v1 protein family (Lee & Facchini, 2010), Zhao et al. 2022 hypothesized that the unknown enzyme was also a member of the same protein family. They isolated 32 candidates from the newly constructed *C. hainanensis* transcriptome (Qiao et al., 2014) that met the criteria for Pr 10-like proteins and discovered 22 members of the Pr 10/Bet v1 family among them. They used a working model that had been created to identify the NCS of the Ranunculaceae member, *Thalictrum flavum*, based on quantum chemical calculations, crystal structures derived from X-ray diffraction, as well as in-vitro assays using recombinant enzyme tailored by site-specific mutagenesis to filter out the most pertinent candidate for the Pictet-Spengler reaction in PIA´s biosynthetic pathway (Bonamore et al., 2010; Ilari et al., 2009; Xu et al., 2021). These researches discovered a glycine-rich loop between the β2 and β3 sheets in the 3D protein structure that is conserved in all reported NCSs and is associated with its enzymatic activity (Morris et al., 2021). Only six of the Pr 10/Bet v1 candidates from *C. hainanensis* transcriptome were discovered to display this glycine-rich loop pattern. These six protein candidates were examined in greater detail by heterologous expression in *E. coli* as they constituted the most likely candidates for the putative Pictet-Spengler enzyme that is thought to be responsible for the production of the PIA backbone. Only one of these candidates was able to produce a product from the substrates after being fed Dopamine and 4-HDCA. Because this enzyme appears to be involved in the synthesis of phenethylisoquinoline scaffolds in *C. hainanensis*, it was named as phenethylisoquinoline scaffold synthase (*Ch*PSS) (Zhao et al. 2022).

1.13. Upstream biosynthesis pathway of 1-phenethylisoquinoline scaffold

Zhao et al. 2022 were able to create the upstream biosynthetic pathway of *Ch*PSS´s substrates (4-HDCA and Dopamine) by combining pathway modeling, *Cephalotaxus* transcriptome mining, and experimental verification of recombinantly expressed candidates by precursor feeding in vitro. In this proposed pathway, the PIA´s central intermediate, 1-phenethylisoquinoline scaffold (Compound number 15), is produced by combining the products of two pathways, one commencing from L-Phenylalanine (L-Phe) and the other from Tyrosine. Figure 1.6 shows the proposed up-stream biosynthesis pathway of PIAs.

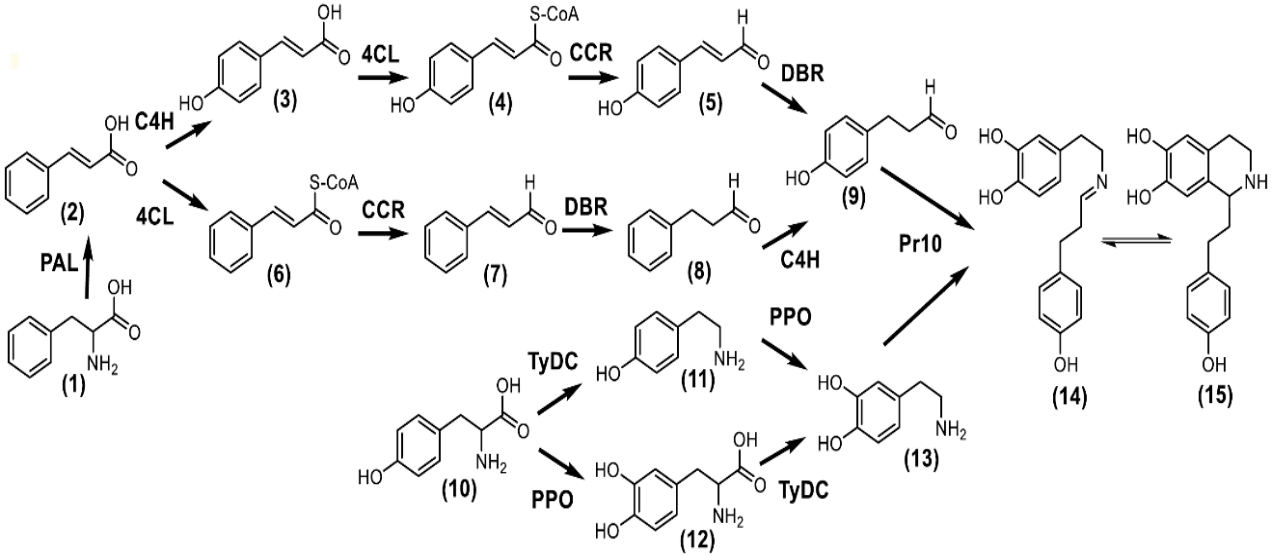


Figure 1.6. The proposed up-stream biosynthesis pathway of PIAs (Zhao et al. 2022).

Compounds: **(1)** Phenylalanine; **(2)** Cinnamic acid; **(3)** *p*-Coumaric acid; **(4)** *p*-Coumaroyl-CoA; **(5)** *p*-Coumaroyl aldehyde; **(6)** Cinnamoyl-CoA; **(7)** Cinnamaldehyde; **(8)** Phenylpropyl aldehyde; **(9)** 4-hydroxydihydrocinnamaldehyde (4-HDCA); **(10)** L-Tyrosine; **(11)** Tyramine; **(12)** L-DOPA; **(13)** Dopamine; **(14)** 6,7-dihydroxy-1-(4-hydroxyphenylethyl)-1,2,3,4-tetrahydroisoquinoline intermediate; **(15)** 6,7-dihydroxy-1-(4-hydroxyphenylethyl)-1,2,3,4-tetrahydroisoquinoline (1-Phenethylisoquinoline scaffold).

Enzyme abbreviation	Enzyme full name
PAL	Phenylalanine ammonia-lyase
C4H	Cinnamate 4-hydroxylase
4CL	4-Coumarate CoA ligase
DBR	NADPH-dependent double bond reductases
CCR	Cinnamoyl CoA reductase
TyDC/DODC	Tyrosine/DOPA decarboxylase
PPO	Polyphenoloxidase
Pr10	Pathogenesis-related 10/Bet v1 proteins (*Ch*PSS)

Table 1.3. Summary of Enzymes involved in the proposed upstream pathway of PIAs by Zhao et al. 2022

1.14. Common precursors in biosynthesis pathway of 1-phenethylisoquinoline scaffold and general phenylpropanoid pathway

Phenolic compounds are secondary metabolites found in plants that share an aromatic ring which contains one or more hydroxyl groups. Over 8000 natural phenolic compounds have been identified. Simple Phenols, Flavonoids, Lignins and Lignans, Tannins, Xanthones, and Coumarins are examples of phenolic compounds isolated from plant sources. Phenylalanine and Tyrosine which are the precursors for biosynthesis of Phenethylisoquinoline scaffolds are also obvious candidates for primary metabolites that could lead to the production of plant phenolics (Anantharaju et al., 2016). Numerous labeling studies, in which plants were "fed" radioactive labeled or stable isotopically labeled compounds to track phenolic production and labeling compound incorporation rates, have demonstrated that L-Phe is typically the favored precursor rather than Tyrosine (Lazar, 2005). The aromatic amino acids Phenylalanine and Tyrosine are derivatives of the shikimic–chorismic acid pathway.

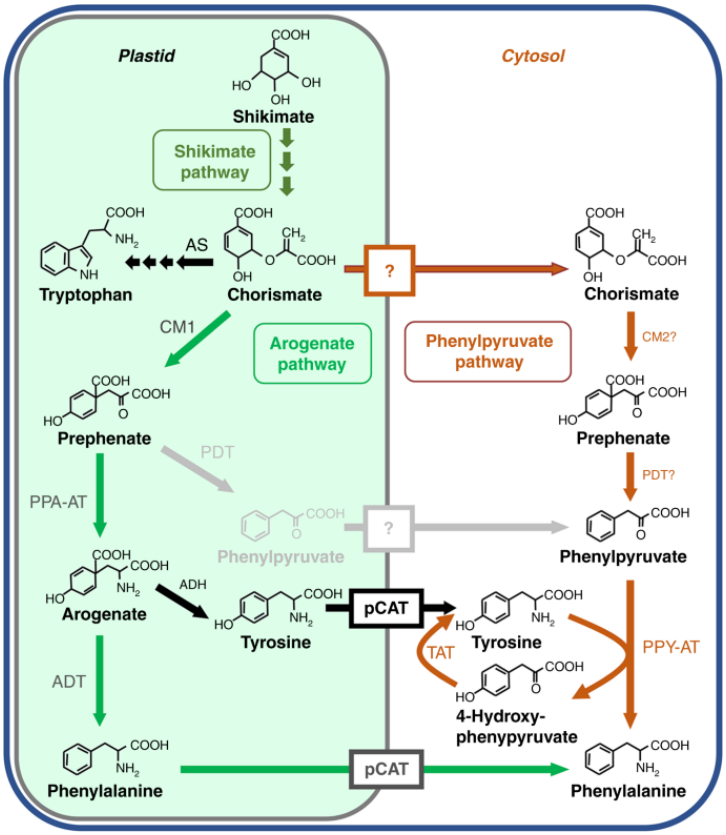


Figure 1.7. Proposed plant Phenylalanine biosynthetic pathways. Characterized enzymes and transporters are shown in solid colors. Uncharacterized enzymes and transporters (boxes) are shown in gray or with question marks. ADH arogenate dehydrogenase, ADT arogenate dehydratase, AS anthranilate synthase, pCAT plastidial cationic amino acid transporter, CM chorismate mutase, PDT prephenate dehydratase, PPA-AT prephenate aminotransferase, PPY-AT phenylpyruvate aminotransferase, TAT Tyrosine aminotransferase. (Qian et al., 2019)

1.15. Common first three steps in biosynthesis pathway of 1-phenethylisoquinoline scaffold and general phenylpropanoid pathway

Most phenolic compounds in plants originate from the phenylpropanoid and the phenylpropanoid-acetate pathways. Both the phenylpropanoid and phenylpropanoid-acetate pathways share the first three biochemical steps, the core phenylpropanoid pathway (Figure 1.8) which is also similar to the first three steps in our proposed PIA backbone biosynthesis pathway.

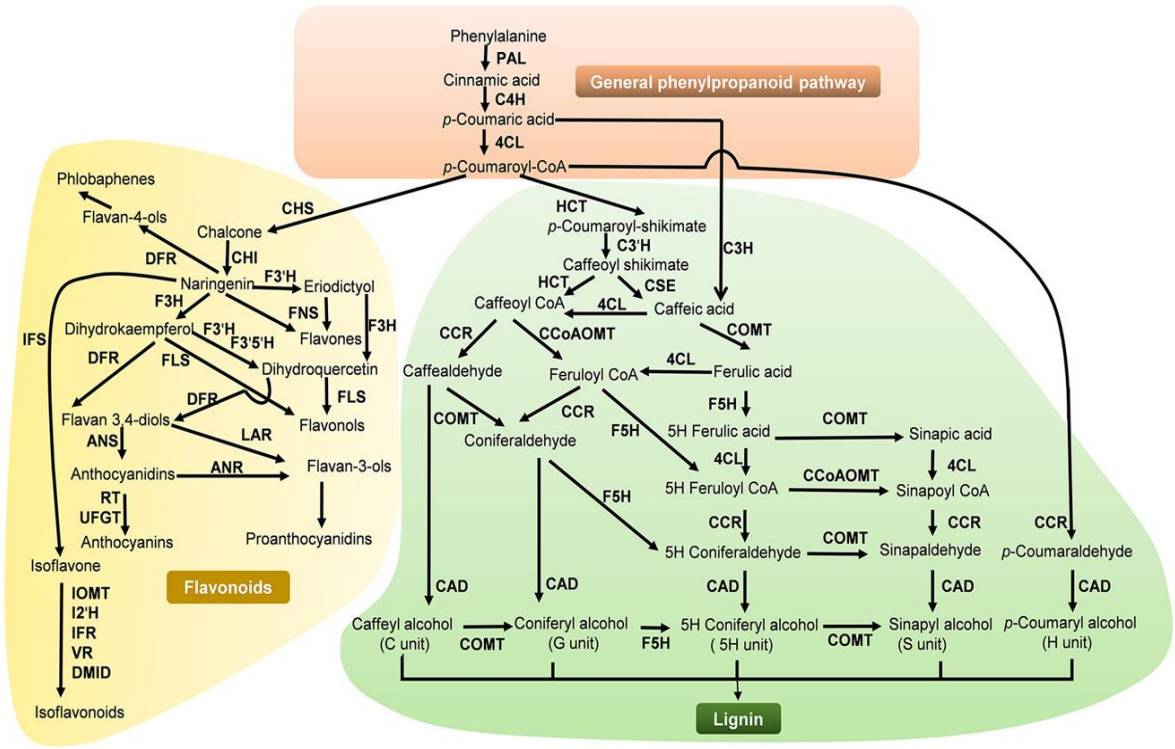


Figure 1.8. A scheme of phenylpropanoid metabolism in plants. Bold font indicates enzymes. 4CL, 4-Coumarate-CoA ligase; ANR, anthocyanidin reductase; ANS, anthocyanin synthase; C3H, Coumarate 3-hydroxylase; C3′H, p-Coumaroyl shikimate 3′ hydroxylase; C4H, Cinnamic acid 4-hydroxylase; CAD, Cinnamyl alcohol dehydrogenase; CCoAOMT, caffeoyl CoA 3-O-methyltransferase; CCR, Cinnamoyl-CoA reductase; CHI, chalcone isomerase; CHS, chalcone synthase; COMT, caffeate/5-hydroxyferulate 3-O-methyltransferase; CSE, caffeoyl shikimate esterase; DFR, fihydroflavonol 4-reductase; DMID, 7,2′-dihydroxy, 4′-methoxyisoflavanol dehydratase; F3′5′H, flavonoid 3′5′-hydroxylase; F3H, flavanone 3-hydroxylase; F3′H, flavonoid 3′-hydroxylase; F5H, ferulate 5-hydroxylase; FLS, flavonol synthase; FNS, flavone synthase; HCT, Hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase; I2′H, isoflavone 2′-hydroxylase; IFR, isoflavone reductase; IFS, isoflavone synthase; IOMT, isoflavone O-methyltransferase; LAR, leucoanthocyanidin reductase; PAL, Phenylalanine ammonialyase; RT, rhamnosyl transferase; UFGT, UDPG-flavonoid glucosyltransferase; VR, vestitone reductase.(Dong & Lin, 2021)

After diverging from an ancestral green alga some 500 million years ago, the first land plants encountered challenging abiotic factors such desiccation, UV radiation, and microbial pathogen attacks (J. Huang et al., 2010; Kenrick & Crane, 1997). Among other significant adaptations, the establishment of the phenylpropanoid pathway enables terrestrial plants to endure these significant

stresses (Ferrer et al., 2008; J. Huang et al., 2010). It is generally accepted that the development of the phenolic metabolism was necessary to support the moving of plants from the water to the land which occurred about 450 million years ago. This phenolic metabolism was probably a prerequisite for adapting to a more restricted environment and facing obstacles. In higher plants, this pathway results in the synthesis of antioxidants, UV-screens, defense compounds, and precursors of biopolymers like sporopollenin, lignin, and suberin, which are necessary for water conduction in vascular tissues and for defense against desiccation, UV exposure, and pests (Boerjan et al., 2003).

1.16. Phenylalanine ammonia lyase (PAL), the entry point enzyme

The first common enzyme in the proposed biosynthesis pathway of Phenethylisoquinoline scaffold (from the Phe branch) as well as the biosynthesis of plant phenolics, is Phenylalanine ammonia lyase (PAL), the entry point enzyme for the majority of plant phenolics which serves as a crucial transitional regulator between primary and secondary metabolism. This enzyme conducts the first stage of phenylpropanoids metabolism, which involves the deamination of L-Phe to generate trans-Cinnamate and Ammonia. To produce $NH4^+$, the cytosol protonates the ammonia produced by PAL. The nitrogen cycle which is actively involved in phenylpropanoid and phenylpropanoid-acetate metabolism in cells, eliminates the need for Nitrogen during phenolic synthesis and prevents ammonia accumulation which is toxic at high concentrations (Figure 1.9) (Lazar, 2005).

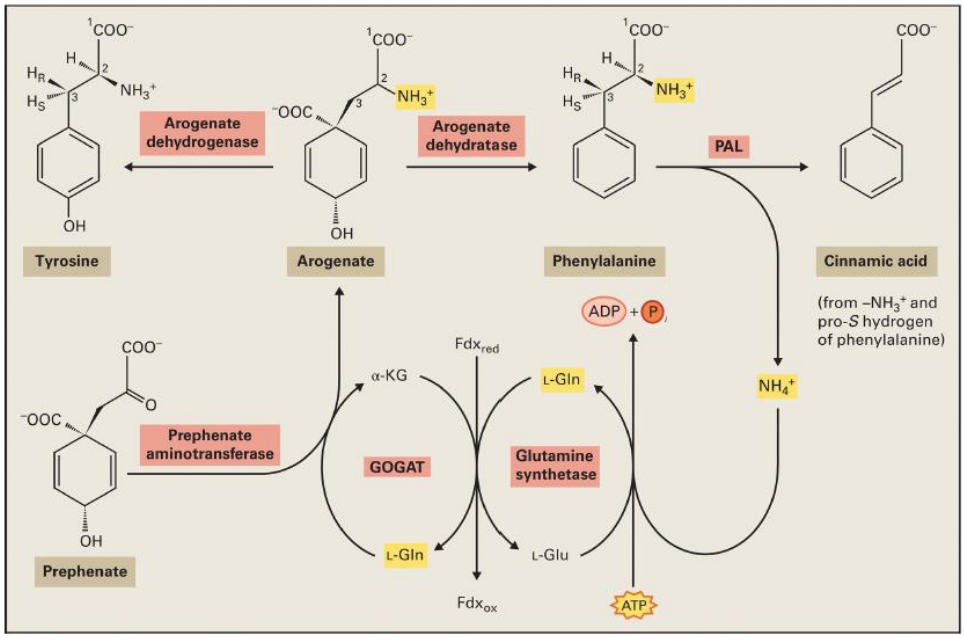


Figure 1.9. The nitrogen cycle present in cells. GOGAT, glutamine: α-ketoglutarate aminotransferase; L-Gln, glutamine; L-Glu, glutamate; α-KG, α-ketoglutarate; Fdxred, reduced ferredoxin; Fdxox, oxidized ferredoxin. (Lazar, 2005b)

Phenylalanine ammonia-lyase is one of the aromatic amino acid lyases (PTAL, EC 4.3.1.25), along with histidine ammonia-lyases (HAL, EC 4.3.1.3), Tyrosine ammonia-lyases (TAL, EC 4.3.1.23), and promiscuous Phenylalanine/Tyrosine ammonia-lyases. PTALs are primarily restricted to the family Poaceae but can also be found in other orders of dicotyledonous plants and ferns. TAL, on the other hand, obviously does not occur in plants (Barros & Dixon, 2020).

This Enzyme was discovered first by Koukol and Conn in *Hordeum vulgare* (barley) in 1961 (KOUKOL & CONN, 1961). PAL has also been detected in prokaryotic organisms and fungus (Hyun et al., 2011). It is believed that PAL in plants was acquired through horizontal gene transfer, first from bacteria to fungi and subsequently from arbuscular-mycorrhizal fungi to plants (Barros & Dixon, 2020; Emiliani et al., 2009). PALs are homotetrameric soluble enzymes that, in the majority of cases, create the so-called MIO (3,5-dihydro-5-methylidene-4H-imidazol4-one) group by autocatalyzing the reaction of three neighboring amino acids (alanine, serine, and glycine) (Ritter & Schulz, 2004).

The function and regulation of this protein (PAL) have been elucidated by the crystal structure analysis of the parsley (*Petroselinum crispum*) PAL(Lois et al., 1989). The cytosolic PAL protein is a member of histidine ammonia lyase (HAL) gene family which is widely distributed in plants, fungi, bacteria, and animals. The enzymes in this family consist mostly of α-helices, exist as tetramers in solution, and require no exogenous cofactors (Calabrese et al., 2004). L-Phe is the only substrate that the majority of PAL isoforms can use effectively, suggesting its position as the universal precursor for the phenylpropanoid and phenylpropanoid-acetate pathways. Most plants have multiple PAL genes, ranging from four in *Arabidopsis thaliana* to nearly two dozen in rare cases like cultivated tomato (*Solanum lycopersicum*)(Chang et al., 2008). These genes are members of a small gene family that exhibit variable expression throughout various stages of plant development, giving rise to distinct PAL isoform functions. All PAL isoforms nearly exclusively use L-Phe in most plants, such as Arabidopsis thaliana and poplar (Populus sp.), while L-Tyr is either a very poor substrate or not at all. Members of different PAL gene families or isoforms appear to serve various roles in the plant. For instance, in contrast to *A. thaliana*'s PAL3 (*At*PAL3), which has weak activity and whose precise function in the plant is unknown, *At*PAL1, *At*PAL2, and *At*PAL4 are thought to be involved in lignin biosynthesis, while only *At*PAL1 and *At*PAL2 are thought to provide precursors for biosynthesis of flavonoid and sinapate ester, which means *At*PAL4 plays no role in the production of flavonoids (Cochrane et al., 2004). Under normal circumstances,

AtPAL1 and AtPAL2 may perform redundant functions or different responsibilities, but when either is altered, they can take the place of the other. A double *At*PAL1/*At*PAL2 mutant causes severe development abnormalities and a substantial decrease in the synthesis of lignin, flavonoids, and sinapate ester, while a single knockout mutation in either has no discernible phenotype (J. Huang et al., 2010; Rohde et al., 2004).

The absence of the PAL in humans makes this enzyme a promising biopharmaceutical candidate for the treatment of a variety of diseases (Kawatra et al., 2020). Therefore, it has grown significantly in importance, in a number of clinical applications such as the effective treatment of phenylketonuria due to its inherent capacity to degrade L-Phenylalanine. As a result, pegylated versions of PAL have been developed for phenylketonuric patients (Shiva Hemmati, 2015).

1.17. Cinnamate-4-hydroxylase (C4H) the second common enzyme in the core phenylpropanoid and Phenethylisoquinoline scaffold pathway

The Cinnamate-4-hydroxylase (C4H) or CYP73 is one of the several membrane-anchored (ER membrane-bound) cytochrome P450 monooxygenases which generates 4-Coumarate (Teutsch et al., 1993). This enzyme is the second enzyme in the central phenylpropanoid pathway and is specialized for trans-Cinnamate. It requires molecular oxygen and a cytochrome P450 reductase to function. Cinnamate 4-hydroxylase (C4H) is the first plant-specific enzyme in the phenylpropanoid pathway because unlike PAL, the first enzyme in the pathway, which except plants is also present in other microorganisms, this enzyme is absent in green algae and first was discovered in bryophytes (Renault et al., 2017). A CYP73 duplication took place in the ancestor of seed plants and was maintained in Taxaceae and the majority of angiosperms.

In the case of secondary metabolism in plants, membrane-bound Proteins, particularly cytochrome P450 enzymes, were suggested to form metabolons for their acting as a catalyst (Jørgensen et al., 2005a; Ralston & Yu, 2006a). Typically, P450 enzymes are attached to the cytoplasmic surface ofthe endoplasmic reticulum (ER) by their N terminus, with the primary protein fold projecting from the membrane's surface (Bayburt & Sligar, 2002). Therefore, P450s can only move in two dimensions. They could be connected to membrane regions with a particular lipid composition or proteins that build membranes.

Numerous investigations have shown that the supramolecular structure of the soluble enzymes surrounding membrane-anchored oxygenases from the cytochrome P450 superfamily severely restricts the plant phenolic pathway (Achnine et al., 2004; Bassard et al., 2012; Chen et al., 2011). It is believed that this spatial configuration is essential for facilitating fast shifts into branch routes and maintaining the flux of precursors in particular pathways. The PAL-C4H couple was reported as the first instance of metabolic channeling (metabolon) in the phenylpropanoid pathway. In fact, a metabolon is a temporary structural-functional complex produced by the successive action of enzymes in a metabolic process. Metabolon is held together by non-covalent bonds as well as structural proteins found in the cell's membrane and cytoskeleton. The development of metabolons facilitates the direct transit (channeling) of an intermediate product from one enzyme into the active site of the following enzyme in the metabolic process (Czichi & Kindl, 1975, 1977). In the secondary metabolism of plants, metabolon production and metabolic channeling allow plants to efficiently produce particular natural compounds while avoiding metabolic interference. Channeling can take place within a metabolon, utilize compartmentalization within the same cell, or include various cell types. A channeling theory was reinforced by the discovery that this proportion converted Phe into p-Coumaric acid more effectively than exogenously provided Cinnamate. Additional research distinguished the PAL1 and PAL2 behavior in tobacco (*Nicotiana tabacum*), where it was discovered that PAL1, but not PAL2, localized to the ER (Rasmussen & Dixon, 1999). Exogenous radiolabeled cinnamate failed to equilibrate with the pool of Cinnamate generated from PAL directly in tobacco cell culturesand microsomal membranes. Following C4H overexpression, both green fluorescent protein (GFP)-tagged PALs were partially re-localized to the ER (Achnine et al., 2004), and further studies suggested that PAL1 had a greater affinity for ER binding sites. Although the data suggested a loose association, colocalisation with C4H was validated by double immunolabeling and fluorescence resonance energy transfer (FRET). *Nicotiana benthamiana* has more recently provided evidence that PAL1 is localized close to the ER membrane (Bassard et al., 2012). Possible effects of the association between PAL and C4H include a decrease in the cell's Cinnamate pool and a reduction in the feedback inhibition of PAL (Blount et al., 2000). Ro and Douglas, 2004 were unable to show metabolic channeling in yeast cells expressing poplar PAL and C4H, despite the fact that evidence exists to indicate its existence at the entrance point into the phenylpropanoid pathway (Ro & Douglas, 2004).

In tobacco (*Nicotiana tabacum*), alfalfa (*Medicago sativa*), and Arabidopsis (*Arabidopsis thaliana*), downregulating C4H leads to a proportional reduction in lignin concentration (Blount et al., 2000;

Sewalt et al., 1997; van Acker et al., 2013). Antisense suppression in tobacco reduces C4H activity, which affects the subunit composition of lignin as well as its content (Sewalt et al., 1997). Furthermore, it has been postulated that tobacco's C4H can open a metabolic channel on the ER through which Cinnamic acid can be transported from PAL to C4H without diffusing into the cytosol (Achnine et al., 2004). As a result, altering C4H may have a bigger effect on metabolic flux than would be predicted based purely on its metabolic function.

1.18. Coumarate CoA ligase (4CL)

The highly active molecule p-Coumaroyl-CoA is formed by the third enzyme in the core phenylpropanoid pathway, 4-Coumarate CoA ligase, using p-Coumarate, ATP, and coenzyme A. This molecule occupies a crucial location in the larger plant metabolic network. Both Chalcone synthase (CHS), the first committed step in the phenylpropanoid-acetate pathway, and Hydroxycinnamoyl transferase (HCT), the first committed step in the phenylpropanoid pathway to lignans, use p-Coumaroyl-CoA as their substrate (Figure 1.8) (Dong & Lin, 2021).

This enzyme is the primary branch point enzyme in the phenylpropanoid pathway. Being the final enzyme of the three shared common steps in general phenylpropanoid pathway, it helps to channelize precursors for various phenylpropanoids therefore plays a crucial function at the point where several important branch pathways diverge from general phenylpropanoid metabolism in the production of phenylpropanoids. Plants have developed a sophisticated gene-metabolic network throughout the history of their lengthy evolutionary history to adapt to various environmental changes. Lignin is a significant metabolite that is essential for protection against biotic and abiotic stresses (Q. Hu et al., 2018; Y. Hu et al., 2009) and 4-Coumarate-CoA ligases are used in the phenylpropanoid route to produce lignin. In fact, 4CL products are used as substrates by numerous oxygenases, reductases, and transferases to produce Lignin, Flavonoids, Anthocyanins, Aurones, Stilbenes, Coumarins, Suberin, Cutin, Sporopollenin, and other compounds (Lavhale et al., 2018) . Two different groupings, type I and type II, can be made up of the 4CL family genes in dicots (Ehlting et al., 1999). Unlike type II genes, which are mostly used in the production of phenylpropanoids other than lignin, type I genes are primarily used in the biosynthesis of lignin (Raes et al., 2003).

The first 4CL gene was cloned from *Petroselinum Hortense*(Ragg et al., 1981). The examination of 4CL genes from numerous plants supports the idea that various plants had varying numbers of 4CL gene families. There are four 4CL genes in *A. thaliana* that encode catalytically active proteins; *At*4CL3 is strongly expressed in mature leaves and flowers and primarily controls the synthesis of flavonoids (Ehlting et al., 1999; Zhong et al., 2022). In *P. tomentosa*, there are two distinct 4CL genes. *Pt*4CL1 plays a role in the production of lignin and is preferentially expressed in tissues involved in lignification. *Pt*4CL2 participates in the manufacture of flavonoids and other phenolic compounds and is selectively expressed in the epidermis of stems and leaves (W.-J. Hu et al., 1998; Zhong et al., 2022). The ligation of coenzyme A (CoA) with Cinnamic acid and its methoxy/hydroxy derivatives, such as Coumaric acid, Caffeine and Ferulic acid, is catalyzed by the 4CL (Figure 1.10)(Lavhale et al., 2018).

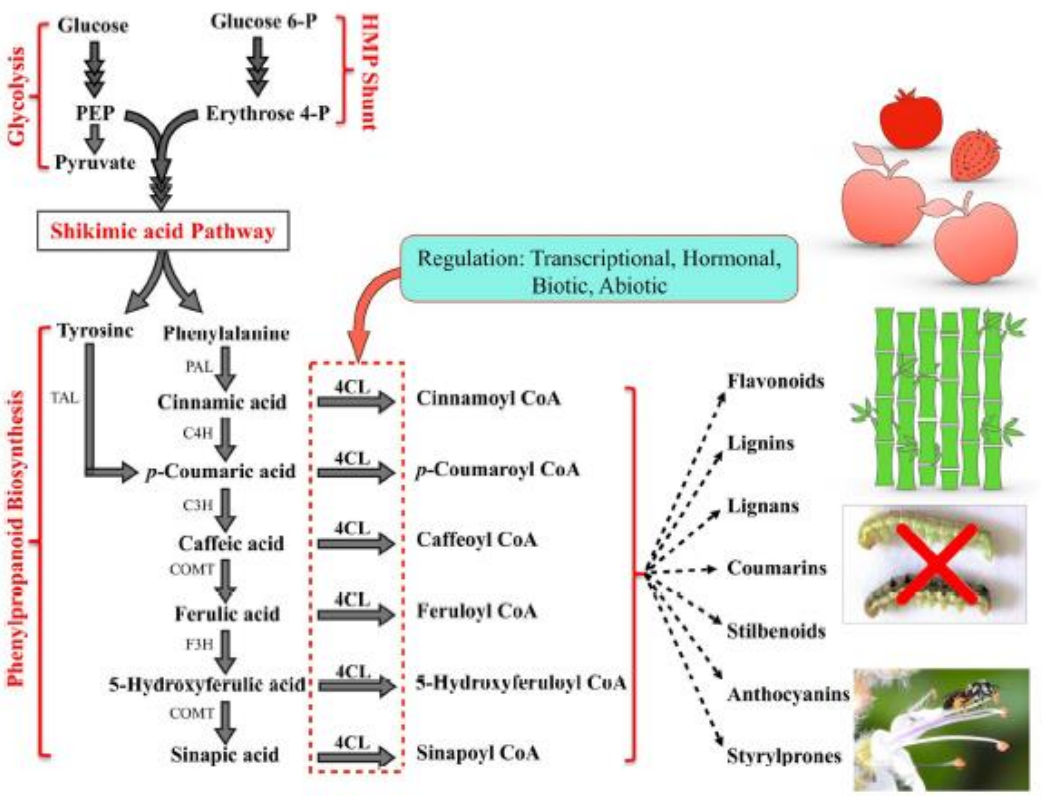


Figure. 1.10 General phenylpropanoid pathway and role of 4CL as a branch point enzyme. *HMP shunt* hexose monophosphate shunt, *4CL* 4-Coumarate-CoA ligase, *PEP* phosphoenolpyruvate, *PAL* Phenylalanine ammonia lyase, *TAL* Tyrosine ammonia lyase, *C4H* Cinnamate 4-hydroxylate, *C3H p*-Coumarate 3-hydroxylase, *COMT* caffeic acid *O*-methyltransferase, *F5H* ferulate 5-hydroxylase.

4CL as a member of adenylate-forming enzyme family, has two conserved peptide motifs, box I (SSGTTGLPKGV) and box II (GEICIRG) (Allina et al., 1998; Ehlting et al., 1999). All members of the adenylate-forming enzyme family share the same box I, which is the adenosine

monophosphate (AMP) binding domain. This AMP binding domain is found in abundance in the proteins of all living Organisms (Fulda et al., 1994; Lavhale et al., 2018). Box II doesn't directly participate in catalysis, and its purpose is unknown (Lavhale et al., 2018; Stuible et al., 2000). Furthermore, distinct 4CL isoforms have different substrate-binding domains (SBD). For two isoforms, *At*4CL1 and *At*4CL2, the domain responsible for determining substrate specificity in Arabidopsis thaliana was investigated using a domain-swapping technique. Both enzymes can use the substrate 4-Coumarate, however only At4CL1 can also use Ferulate (Ehlting et al., 2001). It has been discovered that the 4CL genes exist in small gene families that encode identical, substantially identical, or divergent proteins. These genes have been examined in a variety of plants. In silico genome study identified fourteen genes in *Arabidopsis thaliana* as candidate 4CL, although only eleven of them were investigated for their functional characterization. The remaining three, however, feature a peroxisome targeting sequence that is missing from genuine 4CLs (based on C-terminal signal peptide analysis) (Costa et al. 2005). Serine-lysine-leucine [SKL] or its variants make up peroxisomal targeting signal (PTS) type 1 (PTS1), a C-terminal un-cleaved tripeptide (Gould et al., 1989; Miura et al., 1992; Miyazawa et al., 1989), while type 2 (PTS2) is an N-terminal cleavable peptide with 11 to 36 amino acids (Swinkels et al., 1991; Tsukamoto et al., 1994). Both sequences serve as independent necessary and sufficient PTSs and are utilized by a variety of organisms across the evolutionary tree, from yeast to humans.

Similar to C4H, 4CL gene expression is triggered by environmental factors like pathogen infection, wounding, UV light exposure, and methyl jasmonate treatments. By controlling the expression of 4CL, phytohormones, like Abscisic acid (ABA), Jasmonic acid (JA) and Salicylic acid (SA) affect various metabolic pathways, including the phenylpropanoid pathway. For instance, when ABA, SA, and methyljasmonate (MeJA) were applied to a 2-month-old *Plagiochasma appendiculatum* thallus, Pa4CL1 was upregulated in response to the SA and MeJA treatments while it was downregulated in reaction to the ABA treatment (Gao et al., 2015). When exposed to MeJA and SA, Hc4CL from *Hibiscus cannabinus* displayed downregulation. Hc4Cl transcript levels, however, with the ABA therapy marginally reduced in 1 hour and then increased to their maximum in 24 hours (Choudhary et al. 2013).

1.19. Cinnamoyl-CoA reductase (CCR)

Cinnamoyl-CoA reductase catalyzes the conversion of the CoAs p-Coumaraldehyde, Coniferaldehyde, and Sinapoyl to their respective aldehydes. This enzyme catalyzes the conversion of p-Coumaroyl to CoA to p-Coumaroyl aldehyde, which is not only an important step in the biosynthesis of the *Cephalotaxus* alkaloids but also the first step in the phenylpropanoid pathway of lignin-specific branch, which leads to lignin production. CCR is regarded as a crucial enzyme in regulating the amount and quality of lignins (Davin et al., 2008; Gross, 1981; Vanholme et al., 2010). After cellulose, lignins are the second most prevalent substance on the planet. They contribute significantly to structural support and water/nutrient conduction, and they hold the majority of the wood's methoxyl content. In addition, lignification is thought to be a plant's defensive mechanism against environmental stressors like wounding and pathogen invasion (Barros et al., 2015; Zhao & Dixon, 2011). Wengenmayer et al. (1976) initially described CCR activity in soybean (Glycine max) suspension cultures; Lacombe et al. (1997). first described the cloning of a CCR cDNA from *Eucalyptus gunnii* (Lacombe et al., 1997; WENGENMAYER et al., 1976).

Numerous plant species have been shown to have homologs of the CCR gene families, including 11 genes in *Arabidopsis thaliana*, 9 in *Populus trichocarpa*, 26 in rice (*Oryza sativa*), and 10 in *Eucalyptus grandis* (Carocha et al., 2015; Costa et al., 2003; Kawasaki et al., 2006; Shi et al., 2010). According to expression research, several CCRs may have various roles in plants. For instance, in *Arabidopsis thaliana* AtCCR1 engages in developmental lignification, while AtCCR2 takes involvement in stress and elicitor responses (Lauvergeat et al., 2001). *A. thaliana*'s lignin contents have been seen to drop by up to 50% when AtCCR1 is down-regulated (Goujon et al., 2003). Although CCR-like genes in rice have been linked to defense-related functions in a number of studies (Kawasaki et al., 2006), the biochemical and physiological functions of rice CCR and CCR-like genes are mostly unknown. Different CCR paralogs in barrel clover (*Medicago truncatula*) appear to have differential substrate selectivity, with CCR1 showing a preference for feruloyl-CoA and CCR2 for p-Coumaroyl-CoA and caffeoyl-CoA (Zhou et al., 2010). Using mutants and transgenics, the function of CCR in lignification has been studied. Thinner walls, brown vascular tissue, less lignin, an increase in the S-G lignine subunits´ ratio, and fewer growth and tissue abnormalities could be seen in the vascular tissue of transgenic tobacco (*Nicotiana tabacum*), where CCR was down-regulated (Piquemal et al., 2002). Although the CCR gene family in *Cephalotaxus*

hainanensis is encoded by 10 extremely identical genes, only one gene member was shown to have the ability to catalyze the reduction of p-Coumaroyl to CoA to p-Coumaroyl aldehyde in vitro (Zhao et al. 2022).

1.20. NADP-dependent alkenal double-bond reductase (DBR)

NADPH dependent double bond reductases (DBR) mediated the 7,8-double bond reduction of Phenylpropanal substrates, p-Coumaroyl aldehyde, to their corresponding dihydro molecules, Phenylpropyl aldehyde, in the hypothesized *Cephalotaxus* alkaloids biosynthesis pathway (Zhao et al. 2022). The most recent discovery of the biosynthesis pathway for colchicine identified 4-hydroxydihydrocinnamaldehyde (4-HDCA) as an essential intermediate produced from L-Phenylalanine. However, up until a recent study by Xiong et al. 2022, the existence of a functional gene that catalyzes the synthesis of 4-HDCA remained debatable. This study involved the cloning and characterization of the *Colchicum autumnale´*s alkenal double-bond reductase (*Ca*DBR). According to in vitro enzyme studies in this research, CaDBR1 could catalyze the formation of 4-HDCA from 4-Coumaraldehyde but not from Cinnamaldehyde. The involvement and contribution of this enzyme in the biosynthesis pathway of 4-HDCA and colchicine alkaloids in *C. autumnale* have been verified and confirmed by heterologous production of *Ca*DBR1 in *Escherichia coli* and stable transformation of this gene in tobacco BY-2 cells (Xiong et al., 2022). This enzyme is an essential component of the process that separates lignin biosynthesis from the formation of 4-hydroxydihydrocinnamaldehyde (4-HDCA).

Five candidates were found when the transcriptome of *C. hainanensis* was searched for double-bond reductases (DBR). In fact, *Ch*DBR2 and ChDBR3 were able to convert p-Coumaroyl aldehyde to 4-hydroxydihydrocinnamaldehyde (4-HDCA) in vitro while none of the five *Ch*DBRs were able to convert Cinnamaldehyde to Phenylpropyl aldehyde (Zhao et al. 2022).

Regarding the biosynthesis of the Phenethylisoquinoline scaffold, the above-mentioned enzymes are involved in the PIA's intermediate step, which results in the formation of 4-hydroxydihydrocinnamaldehyde (4-HDCA) from Phenylalanine. Two more enzymes are required to create Dopamine from Tyrosine in order to have the Pictet-Spengler condensation, which is mediated by *Ch*PSS.

1.21. Biosynthetic pathway of Dopamine

Dopamine is the final product of Tyrosine branch contributing in production of 1-Phenethylisoquinoline scaffold in cooperation with 4-HDCA. In plants, Dopamine is produced through two distinct pathways, and Tyrosine serves as the precursor to both. The Tyrosine decarboxylase (TYDC) first converts Tyrosine into Tyramine, which is then hydroxylated by the monophenol hydroxylase (MH) to make Dopamine. Tyrosine is first hydroxylated by the enzyme Tyrosine hydroxylase (TH) to create levodopa (L-Dopa), which is then decarboxylated by the enzyme Dopa decarboxylase (DD) to create Dopamine (Kong et al., 1998; Liu et al., 2020; Steiner et al., 1996). The biosynthetic processes used by various plants to make Dopamine differ, though.

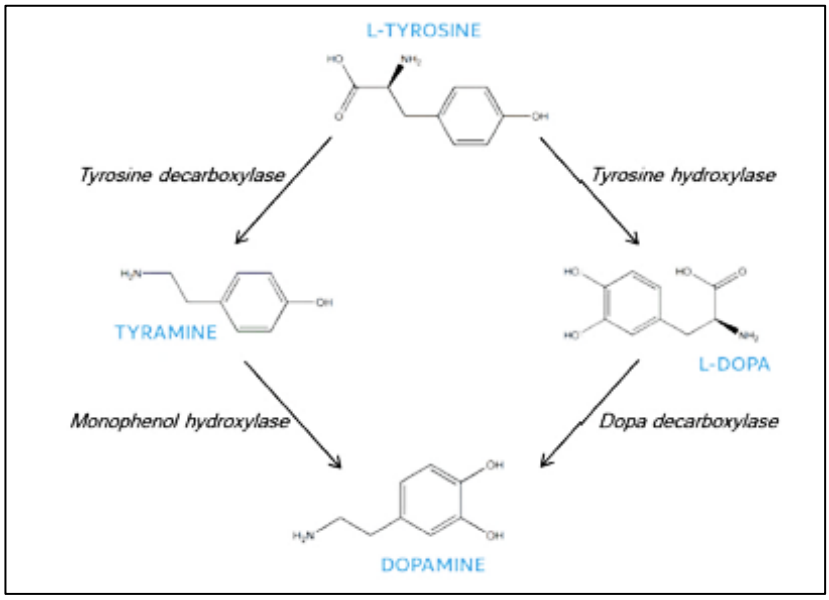


Figure 1.11. Dopamine biosynthetic pathways in plants (Liu et al., 2020)

Tyrosine, for instance, is first hydroxylated to create L-Dopa in *Musa sapientum* before being decarboxylated to create Dopamine, whereas in cacti (*Opuntia stricta*) and purslane (*Portulaca oleracea L.*), Tyrosine is first decarboxylated to begin the production of Dopamine (Liu et al., 2020; Lundström & Agurell, 1971). In addition, Phenylalanine is converted to Tyrosine, which is then converted to L-Dopa or decarboxylated to Tyramine in the peyote cactus *Lophophora williamsii* (Liu et al., 2020; Paul, 1973). Environmental factors also have an impact on the synthesis of Dopamine in plants, in addition to the variability in those pathways. For instance, abscisic acid (ABA) treatment of potato plants can increase the activity of TYDC, DD, and TH; high salt treatment also increased the activity of TYDC; UV exposure also increased the activity of DD; drought also increased the activity of TH and DD; low temperatures can reduce the activity of DD;

and dark treatment and red-light treatment also inhibited the activity of TYDC, TH (Liang et al., 2017).

Both Tyrosine and DOPA can be decarboxylated by *Ch*TyDC1 in the *C. hainanensis* (Zhao et al. 2022). These results are in line with the evidence from other plants, where it was demonstrated that TyDC absorb both Tyrosine (which produces Tyramine) and L-DOPA (which produces Dopamine) (Torrens-Spence et al., 2020). Since a member of the CYP71AD family with a hydroxylase function that is present in red beet (Polturak & Aharoni, 2018), qualified as a candidate for the ensuing hydroxylation, no CYP71AD family member was found when searching the transcriptome of *C. hainanensis* (Zhao et al. 2022). However, it appears that other plants that develop BIAs lack this particular hydroxylase as well.

Search for alternatives in *C. hainanensis* transcriptome, Zhao et al. 2022 considered polyphenol oxidase (PPO) as in one hand it was noted that silencing this gene in *Juglans regia* caused a buildup of Tyramine, the key precursor to Dopamine (Jumper et al., 2021) and on the other hand PPOs often catalyze the oxidation of aromatic rings. Searching for candidates in the *C. hainanensis* transcriptome, and *Ch*PPO1 was able to produce L-DOPA once Tyrosine was added to it (Zhao et al. 2022). Therefore, *Ch*PPO1 can be considered the elusive aromatic hydroxylase in *C. hainanensis*. Additionally, after feeding Tyrosine to *Ch*TyDC1 and *Ch*PPO1, Dopamine biosynthesis was accomplished thanks to their recombinant co-expression under the same bacterial promoter. Therefore, Dopamine synthesis process of PIA production, PPO might functionally substitute for cytochrome P450.

1.22. Metabolic engineering for the production of natural products

Many of the biosynthetic processes involved in the synthesis of plants´ chemically complicated and pharmaceutically valuable chemicals have been clarified thanks to developments in sequencing and recombinant DNA technologies. Metabolic engineering is a more effective way to increase these natural product titers and produce new chemicals thanks to its ever-expanding toolbox of biosynthetic components. Advances in expression systems and regulation have made it possible to fine-tune pathways for increased efficiency, and the characterization of individual pathway components has facilitated the construction of hybrid pathways for the production of novel

compounds. Heterologous production platforms have made it possible to access pathways from difficult to culture strains. These developments in the many facets of metabolic engineering have led to intriguing scientific breakthroughs as well as making it a more and more practical strategy for the enhancement of natural product biosynthesis.

The first crucial choice to be made in the production of pharmaceutically useful chemicals from plants is whether to enhance the native generating strain via plant cell fermentation or to transfer the necessary pathway into a heterologous host for enhancement. If the native producer is genetically recalcitrant, grows slowly, or does not grow well under industrial fermentation conditions, a heterologous host might be preferable. It was a common notion in the 1980s to use plant cell cultures to manufacture these important chemicals. This strategy, meanwhile, has been challenged by the inability to standardize such cultures and to obtain specific chemicals in culture. This is the reason that, starting in the 1990s, most groups stopped using plant cell fermentation, while it has nonetheless proved profitable in a few instances.

Therefore, a heterologous host might be preferable. The source of the pathway to be transmitted and the kind of metabolite produced are frequently factors in the selection of a heterologous host. In light of this, it may be desirable to use a heterologous host. The source of the route to be transferred and the type of metabolite produced are typically taken into account when choosing a heterologous host. Three feasible groups of heterologous hosts for the production of plant-valued natural chemicals are plant cells, fungi like *Saccharomyces cerevisiae* and *Aspergillus nidulans*, and bacteria like *Streptomyces* or the well-known *Escherichia coli*.

The best results in metabolic engineering of cells for the production of heterologous compounds may be obtained, as shown in picture 1.12, by increasing precursor supply, boosting the effectiveness and expression of biosynthetic pathway enzymes, modulating gene expression at the individual gene level using constitutive promoters, using particular regulators, and downregulating or deleting specific genes to eliminate competing pathways. We first require fundamental knowledge of the biosynthetic pathway of 1-phenethylisoquinoline scaffold in order to reach the point of heterologous production of this molecule in a suitable manner and to perform all these manipulations to boost the productivity of the production. As was already mentioned, the overall phenylpropanoid pathway and the synthesis of 1-phenethylisoquinoline scaffold share a number of initial phases. The phenylpropanoid metabolic process is controlled by a number of complex regulatory mechanisms, particularly those that affect the first three enzymes in this pathway, according to numerous studies

that have been conducted over the past several decades on the regulatory evaluations of the phenylpropanoid pathway. Although understanding these regulatory mechanisms helps us understand the fundamentals of plant-specific metabolism, we still don't have a complete understanding of the sub-cellular analysis of the enzymes involved in this complex, extremely dynamic biosynthetic machinery. Our ability to logically design plant metabolic pathways for future applications or the metabolic engineering of 1-phenethylisoquinoline scaffold will undoubtedly be improved by increasing our knowledge of these sub-cellular processes, as we believe that compartmentation and distribution are key factors in these processes.

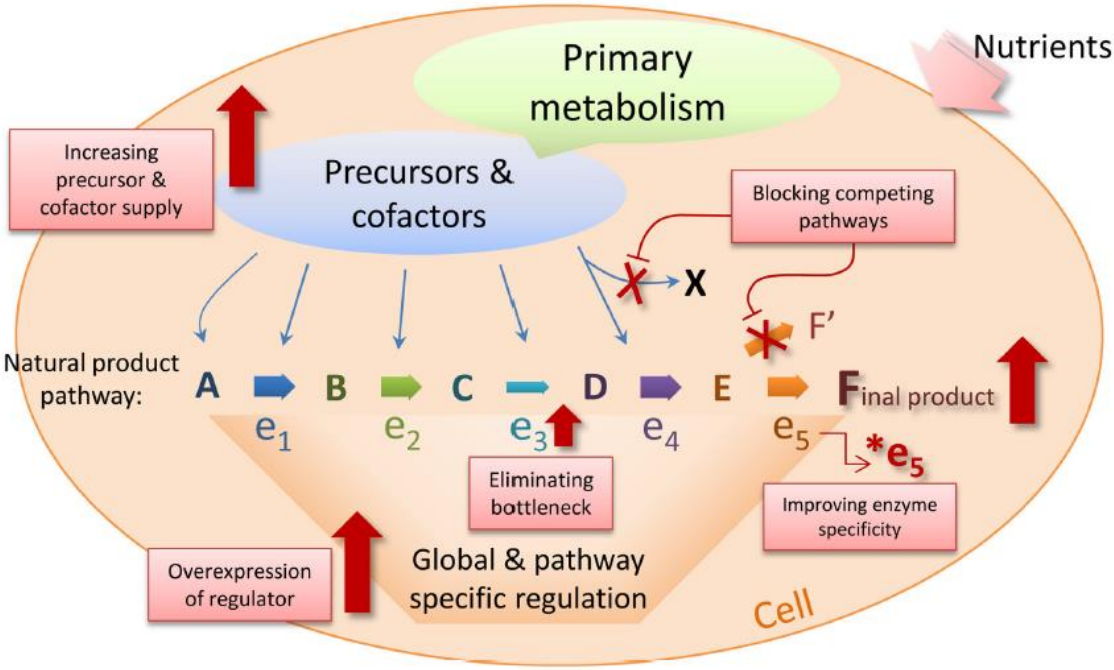


Figure 1.12. General metabolic engineering strategies for improvement of product

1.23. the aims of this project…

We believe that a deeper understanding of sub-cellular mechanisms will undoubtedly increase our capacity to logically create plant metabolic pathways for potential applications or the metabolic engineering of 1-phenethylisoquinoline scaffold. Therefore, we focused our research on the sub-cellular localisation of enzymes involved in the synthesis of 1-Phenylethylisoquinoline scaffold in order to improve our understanding and ability to logically design the metabolic engineering of Cephalotaxine alkaloids for future uses. The metabolic synthesis of entry enzyme (PAL), the most

important enzyme controlling the metabolic flux into the phenylpropanoid pathway and ultimately towards the formation of the scaffold for 1-phenylethylisoquinoline, will also be studied in vivo as part of this study. After delivering the necessary precursors, High-performance chromatography (HPLC) will be employed to assess the efficient metabolite production of PAL over expressed cell lines. The potential for phenylpropanoid pathway dispersion inside the cell, an intriguing new field of inquiry in plant science that explored the cross-membrane dynamics and intracellular communication of various phenylpropanoid biosynthesis intermediates, will then be assessed. An additional objective of this project would be to use relative gene expression analysis to investigate the effects of a precursor supplement on the expression of endogenous and exogenous PAL in transgenic cell lines.

2. Materials and Methods

2.1. Materials

2.1.1. Machines and devices

Machine	Type	Provider
Balance	Analytical plus	Ohaus Corporation, USA
TissueLyser	TissueLyzer	QIAGEN GmbH, Hilden
Vortexer	VortexGenie	Bender and Hobein AG
Schweiz Eppi Incubator Göttingen	TSC THermo Shaker	Biometra GmbH,
Shaker Incubator	GFL 3033	GFL Gesellschaft fu¨r Labortechnik GmbH
Incubator	Memmert	Memmert GmbH + Co. KG
Centrifuge	Heraeus Pico 17	Thermo scientific, Waltham, MA, USA
PCR-tube centrifuge	Mini-Zentrifuge Rotilabo	Carl Roth GmbH + Co. KG, Karlsruhe
96-well plate centrifuge	VWR® PCR Plate Spinners	VWR International GmbH, Darmstadt
Spectrophotometer	NanoDrop ND-1000	Peqlab Biotechnologie GmbH, Erlangen
Thermocycler	FlexCycler CFX96 Touch™ Real–Time	Analytik Jena AG, Jena Bio - Rad Laboratories GmbH, München
Gel scanner	Safe Imager	Invitrogen GmbH, Karlsruhe
Stereo microscope	Leica Stereolupe 420, Leica DFC 500	Leica Microsystems GmbH
(Wetzlar) Camera	Leica DFC 500	Leica Microsystems GmbH
(Wetzlar) Keyence	VHX digital microscope	Keyence Deutschland GmbH
Scanner	Ecosys M4125idn	Kyocera
Digital camera	Cyber-shot DSC-WX350	Sony
Cleanbench	Sicherheitswerkbank HB2448	Heraeus Instruments GmbH, Hanau

Table 2.1. Machines used in this study

2.1.2. Enzymes

Purpose	Enzyme	Provider
PCR	Taq DNA polymerase	New England Biolabs, Frankfurt am Main
A-tailing	Taq DNA polymerase	New England Biolabs, Frankfurt am Main
Cloning	T4 Ligase	New England Biolabs, Frankfurt am Main
cDNA synthesis	M-MuLV Reverse Transkriptase	New England Biolabs, Frankfurt am Main
RNA extraction	RNase freie DNase I	QIAGEN GmbH, Hilden
qPCR	GoTaq polymerase	Promega GmbH, Mannheim

Table 2.2. Enzymes used in this study

2.1.3. Chemicals

Purpose	Chemical	Provider
Dilutions	Nuclease-free H2O	Biozym Scientific GmbH
PCR	dNTPs	New England Biolabs, Frankfurt am Main
cDNA synthesis	dNTPs	New England Biolabs, Frankfurt am Main
Agarose Gel Electrophoresis	SYBRsafe	New England Biolabs, Frankfurt am Main
Agarose Gel Electrophoresis	Midori green Xtra	Nippon Genetics Europe GmbH
Agarose Gel Electrophoresis	Agarose NEEO Ultra-quality	Carl Roth GmbH + Co. KG, Karlsruhe
Agarose Gel Electrophoresis	100bp ladder	New England Biolabs, Frankfurt am Main
Agarose Gel Electrophoresis	1kb ladder	New England Biolabs, Frankfurt am Main
cDNA synthesis	Oligo-dT Primer	New England Biolabs, Frankfurt am Main
cDNA synthesis	RNase Inhibitor	New England Biolabs, Frankfurt am Main
qPCR	SYBRgreen	Thermo Fisher Scientific Inc., Waltham, MA, USA
qPCR	MgCl2	USB Corporation, Cleveland, OH, USA

Table 2.3. Chemicals use in this study

2.1.4. Buffers

Purpose	Buffer	Recipe / Provider
Agarose Gel Electrophoresis	Loading buffer	50% Glycerine / H2O 0,05% Bromphenolblue 0,05% Xylencyanol
Agarose Gel Electrophoresis	TAE-Buffer	2 M Tris 0,05 M EDTA 1 M acetic acid
Agarose Gel Electrophoresis	TB Buffer	500 mM KCl 100 mM Tris-Cl 15 mM MgCl2 [113]
cDNA synthesis	RT-Buffer	New England Biolabs, Frankfurt am Main
qPCR	GoTaq Buffer	Promega GmbH, Mannheim

Table 2.4. Buffers used in this study

2.1.5. Kits

Purpose	kit	Provider
PCR purification	MSB Spin PCRapace	Stratec molecular, Birkenfeld
MiniPrep	Roti-Prep Plasmid Mini	Carl Roth GmbH + Co. KG, Karlsruhe
RNA extraction	Spectrum Plant Total RNA Kit	Sigma-Aldrich, Darmstadt
Gateway cloning	Gateway™ BP Clonase™ II	Thermo Fisher Scientific Inc., Waltham, MA, USA
Gateway cloning	Gateway™ LR Clonase™ II	Thermo Fisher Scientific Inc., Waltham, MA, USA

Table 2.5. Kits used in this study

2.1.6. Agents for cloning

Purpose	Object	Provider / Recipe
Cultivation plates	LB-medium	Carl Roth GmbH, Mannheim
Resistance	ampicillin (100 g/ml)	Carl Roth GmbH, Mannheim
Amplification of plasmids	*E.coli* strain DH5α	Inoue method
Transformation	pGEM-T Easy Vector System	Promega GmbH, Mannheim
Blue-white screening	IPTG	Carl Roth GmbH + Co. KG, Karlsruhe
Blue-white screening	X-gal	Biomol GmbH, Hamburg

Table 2.6. Agents used in this study

2.1.7. Plasmids

Plasmid name	Bacterial selection	Plant selectable marker	Purpose	Provider
pGEM-T Easy Vector	Ampicillin	_	TA-cloning	Promega
PDONR Vector	Zeocin	_	Gateway cloning	Thermo Fisher Scientific
pK7FWG2 (GWV12)	Spectinomycin	Kanamycin	Plant expression vector C-terminal fusion to GFP- Gateway cloning	VIB-UGent Center for Plant Systems Biology
pK7WGF2 (GWV13)	Spectinomycin	Kanamycin	Plant expression vector N-Terminal fusion to GFP- Gateway cloning	VIB-UGent Center for Plant Systems Biology
pH7RWG2 (GWV10)	Spectinomycin	Hygromycin	Plant expression vector C-terminal fusion to RFP- Gateway cloning	VIB-UGent Center for Plant Systems Biology

Table 2.7. Plasmids used in this study

2.1.8. Software and Databases

Purpose	Software/ Database
Sequence quality evaluation	FinchTV version 1.4.0
Multiple sequence alignment	MEGA7 version 7.0.14
Evaluation of qPCR data	Bio-Rad CFX Manager version 3.1
OLIGO Primer Analysis Software	OLIGO version 7.5
Gene analyzing	SnapGene Viewer 4.2.11
BLASTing and sequence data	NCBI (National Center for Biotechnology Information)
Reference genome *N. tabacum*	https://www.genome.jp/kegg-bin/show_organism?org=nta
Subcellular localisation Prediction -TargetP - 2.0	https://services.healthtech.dtu.dk/service.php?TargetP-2.0

Table 2.8. Software and Databases used in this study

2.1.9. Genes

Gene	Accession (NCBI)	Abbreviation	known name in nick-lab
Cephalotaxus hainanensis isolate 1 Phenylalanine ammonia-lyase (PAL1)	MN398159.1	PAL1	PAL2344
Cephalotaxus hainanensis isolate 2 Phenylalanine ammonia-lyase (PAL2)	MN398160.1	PAL2	PAL2638
Cephalotaxus hainanensis isolate 3 Phenylalanine ammonia-lyase (PAL3)	MN398161.1	PAL3	PAL2680
Cephalotaxus hainanensis isolate 4 Phenylalanine ammonia-lyase (PAL4)	MN398162.1	PAL4	PAL2898
Cephalotaxus hainanensis isolate 1 Cinnamate 4-hydroxylase (C4H1)	MW691947.1	C4H1	C4H656
Cephalotaxus hainanensis isolate 2 Cinnamate 4-hydroxylase (C4H2)	MW691948.1	C4H2	C4H11
Cephalotaxus hainanensis isolate 3 Cinnamate 4-hydroxylase (C4H3)	MW691949.1	C4H3	C4H12
Cephalotaxus hainanensis isolate 4 Cinnamate 4-hydroxylase (C4H4)	MW691950.1	C4H4	C4H82
Cephalotaxus hainanensis 4-Coumarate: CoA ligase (4CL1)	MZ833459.1	4CL1	4CL1
Cephalotaxus hainanensis 4-Coumarate: CoA ligase (4CL2)	MZ833460.1	4CL2	4CL2
Cephalotaxus hainanensis 2-Alkenal reductase (DBR1)	MZ833471.1	DBR1	DBR12
Cephalotaxus hainanensis 2-Alkenal reductase (DBR2)	MZ833473.1	DBR2	DBR1466
Cephalotaxus hainanensis 2-Alkenal reductase (DBR3)	MZ833474.1	DBR3	DBR1450
Cephalotaxus hainanensis Cinnamoyl-CoA reductase (CCR1)	MZ833461.1	CCR1	CCR1565
Cephalotaxus hainanensis Cinnamoyl-CoA reductase (CCR2)	MZ833462.1	CCR2	CCR1323
Cephalotaxus hainanensis Cinnamoyl-CoA reductase (CCR3)	MZ833464.1	CCR3	CCR12
Cephalotaxus hainanensis Cinnamoyl-CoA reductase (CCR4)	MZ833466.1	CCR4	CCR1363
Cephalotaxus hainanensis Cinnamoyl-CoA reductase (CCR5)	MZ833468.1	CCR5	CCR11512

Cephalotaxus hainanensis Polyphenol oxidase (PPO1)	MZ833479.1	PPO1	PPO1
Cephalotaxus hainanensis Tyrosine decarboxylase (TyDC1)	MZ833477.1	TyDC1	TyDC1780
Cephalotaxus hainanensis Phenethylisoquinoline scaffold synthase (PSS)	MZ833481.1	PSS	PR10-084

Table 2.9. Genes amplified in this study

2.1.10. Primers

Gene	Forward Or Reverse	Primer sequence	Size of transcript [bp]	Annealing Tm [°C]
PAL1	Sense	5‘ATGTTTTCAATGGCGGAGGTT 3’	2235 bp	60
	Antisense	5‘TCAAATGATAGAGAGAAGCCAGTAG 3’		
PAL2	Sense	5‘ATGGAGACTTCTCTGAATCAGAAC 3’	2271 bp	60
	Antisense	5‘TCAGAGCACTTGGTGGGAGG 3’		
PAL3	Sense	5‘ATGGTAGCGGGTGTGGAT 3’	2172 bp	59
	Antisense	5’TCAACCTATGGGGAAAGGAC 3’		
PAL4	Sense	5‘ATGGCAGCAACACAAGACTTTAC 3’	2139 bp	60
	Antisense	5‘TTAAATTGGGATTGGGGCG 3’		
C4H1	Sense	5‘ATGGAGATGGATTTGTTTAACGG 3’	1506 bp	59
	Antisense	5‘TCATGCTCTGGGTTTAGCC3’		
C4H2	Sense	5‘ATGGGGAAAAAGATGGCGG 3’	1551 bp	59
	Antisense	5‘CTAGTAAACTCTGGGCTTGG 3’		
C4H3	Sense	5‘ATGGCTTCCTTGCTGAAC 3’	1566 bp	59
	Antisense	5‘TCACAGAAGGCGGAAGCGG 3’		
C4H4	Sense	5‘ATGGATTTGTTTAATGGCCTG3’	1500 bp	59
	Antisense	5‘TCATGCTCTGGGTTTGGC3’		
4CL1	Sense	5‘ATGGCGGGATGCGAAGG 3’	1629	59
	Antisense	5‘TTAAGCAACTACTACAAGTCTGGCTC3’		
4CL2	Sense	5‘ATGATTGAGGTGGACGAGCTG3’	1671	60
	Antisense	5‘TCATTTGGGCGTTACTTGGGC3’		
	Sense	5‘ATGGAAACGAAAGAAGTGATCAA3’		

DBR1	Antisense	5‘CTATGGATGATGGAGGCGAAC3’	1047	58
DBR3	Sense	5‘ATGGCAAGCAATGAAGTGAGTAA3’	1035	58
	Antisense	5‘TCATTCGTCACCTATGCAAACAA3’		
DBR4	Sense	5‘ATGGGAAAAGAGGTTGTGAATAA3’	1032	58
	Antisense	5‘TCATCCATCAGAAATACAAATTACCTGC3’		
CCR1	Sense	5‘ATGACTATGGACAATCAGACAGG3’	975	59
	Antisense	5’TCACTTGGAAATATGGCCCTTCT3’		
CCR2	Sense	5‘ATGGAGGAAGTTTGTGTGAC3’	966	59
	Antisense	5‘CTACAGCAGCCCCTTCTCTTG3’		
CCR4	Sense	5‘ATGCCTGTATTTTCTCACTTGGATG 3’	903	59
	Antisense	5‘CTACTGCTTTGTGTATCGACCCATG3’		
CCR6	Sense	5‘ATGACCGGGAAGCTAGTTTGT3’	981	59
	Antisense	5‘TCACATATTACTCTCCAACTGAAGC3’		
CCR8	Sense	5‘ATGGCGATGGTGAAGTCTG3’	1044	58
	Antisense	5‘TCTTCGGCAGCAGGAATTG3’		
TyDC1	Sense	5‘ATGGAAGGATGGAAGAGAGTGG3’	1506	59
	Antisense	5‘CGAGTAACTGAAAGCAACATGA3’		
TH1	Sense	5‘ATGGAGGGAGGAGGTTCG3’	810	59
	Antisense	5‘TTATCTTCTATCTTTCCCTGATGAC3’		
PPO1	Sense	5‘ATGGCTCCTACTTATCCCGA3’	1674	60
	Antisense	5‘TTAATGATATGGGAACACACACA3’		
PR10-084	Sense	5‘ATGGTGGCGAAAAGTATAACAGTAG3’	483	60
	Antisense	5‘ACAATATACGGTAGGGTTAGAGACGA3’		

Table 2.10. Primers for gene amplifications

2.2. Methods

2.2.1. Plant material, cell culture and elicitation

In this study we used *Nicotiana tabacum* cv. "Bright Yellow 2" plants (voucher KIT 8579) as a viable platform housing the shikimate pathway for synthesizing and determining the concentration of our desired alkaloids. *Nicotiana tabacum* L. cv. Bright Yellow 2 (Cell Strain BY-2) (Nagata et al., 1992) was grown in a liquid medium with 4.3 g/L Murashige and Skoog basal salts, 30 g/L

sucrose, 200 mg/L KH2PO4, 100 mg/L Inositol, 1 mg/L Thiamine, and 0.2 mg/L (0.9 µM) 2,4-D, pH 5.8 as shown in table 2.11. Weekly subcultures of the cells were performed by inoculating 1.5 ml of stationary cells into 30 ml of new media in 100 ml Erlenmeyer flasks. The cells were incubated at 26 °C while being continuously shaken at 150 rpm on an IKA Labortechnik KS260 basic orbital shaker (http://www.ika.de). Stock calli were sub-cultured on media containing 0.8% (w/v) agar (Roth, http://www.carlroth.com) every three weeks. The transgenic strains' suspension cultures and calli were grown in the same media as non-transformed wild-type cultures (BY-2 WT), but with the addition of 50 mg/L kanamycin.

Component	Concentration
Murashige and Skoog basal salts	4.3 g/L
Sucrose	30 g/L
KH2PO4	200 mg/L
Inositol	100 mg/L
Thiamine	1 mg/L
(0.9 µM) 2,4-D	0.2 mg/L

Table.2.11. MS medium components

2.2.2. PCR (mastermix / setting / gel electrophoresis / PCR product cleaning)

The target gene was amplified using a 25 µl reaction volume that contained 16 µl of nuclease-free H2O (Lonza, Biozym), fivefold Q5 Reaction Buffer (New England Biolabs), 200 µM dNTPs (New England Biolabs), 0.5 µM of forward and reverse primer (see primer list, table 2.10), 100-150 ng of cDNA template, and 0.75 units of Q5 High-Fidelity. The thermal cycler settings for the amplification included a denaturation step at 98°C for 30 seconds, followed by 35 cycles at 98°C for 30 seconds, 59–60°C for 30 seconds, and 72°C for 20–30 seconds/kb, with a final extension at 72°C for 2 minutes. NEEO ultra-quality agarose from Carl Roth, Karlsruhe, Germany, was used to conduct an agarose gel electrophoresis to assess the PCR. SYBRsafe (Invitrogen, Thermo Fisher Scientific, Germany) or Midori green Xtra (Nippon Genetics Europe GmbH) and blue light excitation were used to visualize DNA. The 100 bp or 1 kb size standard (New England Biolabs) was used to determine the fragment size. For further cloning (TA cloning) procedures, amplified DNA was purified using the MSB Spin PCRapace kit (Stratec), as directed by the manufacturer.

Component	Volume [µl]
5X Q5 Reaction Buffer	5
dNTPs (10 mM)	0.5
Forward primer (10 µM)	1.25
Reverse primer (10 µM)	1.25
Template DNA or cDNA	0.7
Q5 High-Fidelity DNA Polymerase	0.3
Nuclease-Free Water	16
Final volume	**25**

Table 2.12. (RT)-PCR single reaction

Temperature	Duration	Repeats
98 °C	30 s	1x
98 °C	10 s	35x
59 °C – 60°C	30 s	
72 °C	20–30 seconds/kb	
72 °C	2 min	1x
4 °C	∞	

Table 2.13. (RT)-PCR cycling conditions

2.2.3. Cloning (A-tail, ligation, transformation, blue-white screening, miniprep)

The cleaned PCR product was utilized in a total volume of 10 µl for the A-tailing process together with 1,7 µl of nuclease-free H2O (Lonza, Biozym), 1 µl of 10x Thermopol Buffer (New England Biolabs), 0,2 µl of 200 mM dNTPs (New England Biolabs), and 1 unit of Taq polymerase (New England Biolabs). The following phase of ligation employed this mixture directly after it had been incubated at 68°C for 60 min. 2 µl of A-tailing product were added to a mixture of 5 µl of ligation buffer, 1 µl of the pGEM-T Easy Vector System, 1 µl of T4 ligase, and 1 µl of nuclease-free water to complete this step overnight at 4 degrees Celsius.

For transforming the ligation product to DH5α first the ligation product was combined with 50 µl of chemo competent DH5α *E. coli* bacterial cells and incubated for 30 minutes on ice. Transformation then occurred for 50 seconds in a 42 °C water bath. The samples were immediately put back on ice for two minutes, and then, after adding 950 mL of LB medium, the transformed cells were incubated

for two hours at 37 °C with constant shaking at 180 rpm. Following two hours, the cells were centrifuged for two minutes at 3000 rpm, 850 µl of medium were then withdrawn, the pelleted cells were then resuspended in the medium, and they were then transferred to an LB-agar plate containing ampicillin (0,1%), IPTG (0,2%), and X-gal (0,8%). Plates were incubated at 37°C for an entire night before being kept at 4°C until the blue-white screening.

White (positive) colonies were selected sterile, moved to a single eppi with 1 ml liquid LB media, and 0,1% ampicillin. After two hours, the subsequent overnight incubation at 37°C and 180 rpm was temporarily interrupted to take 1 µl of media (containing the transformed E. coli), which was utilized as a template for a colony PCR. Because Taq DNA polymerase was employed in this amplification program (instead of Q5 High-Fidelity DNA Polymerase in the preceding amplification), the PCR settings, master mixes, and thermal cycler parameters were altered and were proceeded as the following.

For the colony PCR, the target genes were amplified using a 25 µl reaction volume containing 15.8 µl of nuclease-free H2O (Lonza, Biozym), one unit of tenfold Thermopol Buffer (New England Biolabs), 200 µM dNTPs (New England Biolabs), 0.2 µM of forward and reverse primer (see primer list, table 10), 100-150 ng DNA template (colony PCR culture) and 0.6 unit of Taq polymerase (New England Biolabs). For the thermal cycler utilized for the amplification, the following conditions were used: initial denaturation at 95°C for 3 min; 30 cycles at 95°C for 30 s, 58-60°C for 30 s, and 68°C for 1 minute/kb based on the length of the fragments; followed by an extension at 68°C for 5 min.

Component	volume [µl]
10X Standard Taq Reaction Buffer	2.5
dNTPs (10 mM)	1
forward primer (10 µM)	1
reverse primer (10 µM)	1
template DNA (Colony culture)	5
Taq DNA polymerase	0.2
Nuclease free H2O	up to 25 µl
Final volume	**25**

Table 2.14. Colony-PCR single reaction

Temperature	Duration	Repeats
95 °C	3 min	1x
95 °C	30 s	30x
58 °C – 60°C	30 s	
68 °C	1 minute/kb	
68 °C	5 min	1x
4 °C	∞	

Table 2.15. (RT)-PCR cycling conditions

PCR condition and master mixes plus thermal cycler conditions were the same as for the preceding program, and the specific primer pair for each gene was utilized. Colonies with the correct sized insert were cultivated further while false positive colonies were discarded.

Finally, using the Roth-Prep Plasmid Mini kit in accordance with the manufacturer's instructions, the plasmids of the positive colonies that had been cultured positively overnight were isolated by mini-prep. The volume and concentration of samples provided were accordance with the needs of the Eurofins company to whom the sequencing was outsourced. The universal T7 and SP6 primers were utilized for this sequencing.

2.2.4. Evaluation of sequencing results

Utilizing FinchTV version 1.4.0 software (J Patterson et al., 2004), the sequences' quality was evaluated. The T7 forward and SP6 reverse primers were used for two-direction sequencing and for sequences over 1200 pb a middle primer was also used to obtain the whole sequence of *Cephalotaxus hainanensis* gene´s open reading frame (ORF). The overlapping sections of the complete fragment were aligned using MEGA version 7.0.14 (Sudhir Kumar et al., 2016). As stated, the whole cDNA sequence for each gene were acquired. To ensure that our ORF has no mutation and is prepared to be expressed correctly in the future, MEGA7 was used to perform a multiple sequence alignment of the obtained cDNA sequences with the sequences of the corresponding gene that were supplied to us from our Chinese partner.

2.2.5. C-terminal fusion of confirmed ORFs to GFP-tagged expression vectors

Aiming expression and studying the subcellular localisation and regulatory patterns of the candidate genes, fluorescent labeling was performed using gateway recombination cloning technology (Gateway® Technology with Clonase® II by Invitrogen™ Life Technologies) through following

the supplier's user guide. By using the Gateway-cloning method (Invitrogen), the constructs for being expressed in tobacco BY2 cells were created in two steps: the BP reaction and the LR reaction to create the entrance clone and the expression clone, respectively. Using the attB1 and attB2 primer pairs, genes were initially amplified for this purpose with PCR. AttB1 and attB2 primers were the identical primers listed in table 2.10, after adding a specific sequence known as attB1 sequence to the beginning of forward primers and attB2 sequence to the end of reverse primers after the stop codon was removed.

attB1 sequence 5'GGGGACAAGTTTGTACAAAAAAGCAGGCTTC3' + Forward primers

attB2 sequence 5'GGGGACCACTTTGTACAAGAAAGCTGGGTC + Reverse primers without Stop codon

The PCR single reaction and cycling conditioned were followed as written in tables 2.12 and 2.13 with annealing temperature of 63°C.

2.2.6. BP reaction

To get an entry clone, the PCR products were cloned using a BP reaction into the pDONR/Zeo vector from Invitrogen. The elements of the reaction are listed below.

Component	volume [µl]
PCR product (50-150 ng)	1
pDONR/Zeo (150ng/µL)	1
BP clonase II enzyme	1
TE buffer	To 5

Table 2.16. Components for BP reaction

18 hours were spent incubating the BP reaction at 25 °. Then 1 µl of proteinase K was added to stop the reaction. This was followed by 10 minutes incubation at 37 °C. After that, the mixture underwent heatshock transformation protocol as described in 2.5 to be send to *E. coli* DH5 competent cell. The transformed cells were placed on an LB plate containing 50 µg/mL zeocin, and they were then incubated at 37°C overnight. In the next day a single colony was chosen and cultured in 5 mL liquid LB medium with 50 µg/mL zeocin at 37°C overnight with constant shaking at 180 rpm. The

following day, using the Roth®-Prep Plasmid MINI kit (Roth, Germany), plasmids were extracted, and the insertion was subsequently confirmed through sequencing (GATC Biotech, Cologne, Germany). In the LR reaction, these entry clones were utilized.

2.2.7. LR reaction

A LR reaction was utilized to clone the expression vectors fused with GFP, as demonstrated below:

Component	volume [µl]
Entry clone (50-150 ng)	1
pK7FWG2 (150ng/µL)	1
LR clonase II enzyme	1
TE buffer	To 5

Table 2.17. Components for LR reaction

Apart from the selection antibiotic, which was 100 µg/mL Spectinomycin, the remaining steps were the same as those outlined for the BP reaction. These expression vectors were employed to transform it into *Agrobacterium tumefaciens* after being confirmed by sequencing.

2.2.8. Transformation of *Agrobacterium tumefaciens* with GFP-labelled expression constructs

After being defrosted on ice, 100 µl of chemo-competent cells of *Agrobacterium tumefaciens* (strain LBA 4404; Invitrogen Corporation, Paisley, UK) was incubated for 20 min with 100 ng binary expression vectors encoding the desired genes. The transformation was done based on the freeze-thaw method described by (Weigel & Glazebrook, 2006). The bacteria were then spread out onto solid LB (Lennox Broth, Roth, https://www.carlroth.com) agar media with antibiotics (100 g μml^{-1} rifampicin, 300 µg ml^{-1} streptomycin, and 100 µg ml^{-1} spectinomycin) and incubated for 3 days at 28 °C in the dark. A number of the single colonies were separately introduced into 1 ml of LB liquid medium that contained the same selective antibiotics and was then cultured at 28 °C with vigorous shaking for a whole night in order to be tested by colony PCR (described in 2.5) with their specific primers. One of the positive colonies confirmed by PCR was selected and used for establishment of stable transgenic tobacco BY-2 cell lines.

2.2.9. Transformation and establishment of stable transgenic tobacco BY-2 cell lines

A method described by (Buschmann et al., 2011) was used to create stable cell lines for overexpressing *Cephalotaxus hainanensis´* genes in the upstream of PIA´s backbone production pathway, with a number of modifications for improved performance. There are two separate preparations in this procedure. One of them involves making BY-2 WT cells, while the other includes preparing agrobacterium for transformation.

2.2.9.1. Preparing BY-2 WT cells for establishment of stable transgenic tobacco BY-2 cell lines

1.5 ml of 7-day-old BY-2 WT cells were sub-cultured and maintained for 3 days in the same manner as normal suspension cell cultures being incubated at 26 °C with continuous shaking at 150 rpm. Then, three flasks containing 31.5 ml of suspension cell culture from each flask were combined and washed twice with 450 ml of Paul medium (4.3 g l-1 Murashige and Skoog salts; (Duchefa Biochemie, https://www.duchefa-biochemie.com/), 10 g l-1 sucrose, pH 5.8). Using a thermo fisher Nalgene Reusable Bottle Top Filter, 500mL, 45mm Collar, Pk 1 and a nylon mesh with 70 µm pores in diameter, the washing procedures were carried out. A 5- to 6-fold concentrated cell suspension resulted after resuspending the washed cells in 15–18 ml of MS media.

2.2.9.2. Preparing *Agrobacterium tumefacien*s cells for establishment of stable transgenic tobacco BY-2 cell lines

Agrobacterium tumefaciens transformants cell suspension that had been prepared as described in 2.8 was combined with 6 ml of this concentrated cell suspension. A 150 µl overnight culture of the chosen colony, which was verified already by colony PCR, was added to 5 ml of LB liquid medium with the same selective antibiotics, and the mixture was vigorously shacked overnight at 28 °C. To achieve an OD600 of 0.8, the next day 1.5 ml of the overnight culture was injected into 5 ml of fresh LB-medium (without antibiotics) to reach the OD of 0.8 after roughly 5 hours of cultivation. The Agrobacterium cells were then centrifuged in 50 ml falcon tube at 8000 g for 10 minutes at 4 °C using a Heraeus Pico 17 Centrifuge (Thermo Fischer Scientific). After discarding LB supernatant, 180 µl of MS media was used to resuspended the bacteria and by utilizing a bench-top vortexer (Bender & Hobein Zurich, Switzerland) mix the suspension until it is homogenous.

In order to completely combine the prepared BY-2 WT cells with the 180 µl homogenized agrobacteria solution, the mixture was shaken on an orbital shaker for five minutes at a speed of 100

rpm and at a 30° angle above the horizontal. This mixture was then deposited onto petri plates that had previously been covered in a single layer of sterile filter paper and were filled with Paul medium that had been solidified with 0.5% (w/v) PhytagelTM (Sigma-Aldrich) without any antibiotics. Instead of incubating the plates at the suggested 27 °C as stated in the original article, they were parafilm-sealed and kept at 22 °C in the dark. After three to four days, the filter paper and cell plaques were transferred to MS agar plates that had been pretreated with 300 mg/ml cefotaxime and 100 mg/ml kanamycin. The plates were then incubated at 26 °C in the dark throughout this time. After around 3 weeks of incubation, the calli that had developed were moved to new MS agar plates (along with the required antibiotics and cefotaxime) for additional development. Once enough calli had reached the proper sizes, a suspension culture was then created from the calli.

2.2.10. Transient transformation for visualization of subcellular compartments

For further analysis and a comparison of cell features, transient transformation by particle bombardment was used to visualize the subcellular compartments of Tobacco BY2 cells. According to the procedure outlined by (Finer et al., 1992), cells that were three days old were employed for the transformation. The appropriate plasmid DNA was coated on gold particles (1.5-3.0 μm; Sigma-Aldrich, http://www.sigmaaldrich.com) in accordance with the standard Bio-Rad manual with the following modifications. The macro carriers were filled with 1 μg of plasmid-DNA that was coated on 12.5 μl (1.5 mg) of gold suspension particles (Bio-Rad, http://www.bio-rad.com/). In a 2 ml Eppendorf reaction tube, 800 μl of 3-day-old transformed and untransformed tobacco BY-2 cells were allowed to settle for 5 minutes. 300 μl of the supernatant were then taken out. The remaining 500 μl of cells were resuspended and then equally distributed onto the PetriSlidesTM (Millipore, Billerica, USA) containing 1.5 ml of the aforementioned solid medium. According to (Finer et al., 1992), the loaded PetriSlidesTM and macrocarriers were both put into the chamber of the specially manufactured particle gun and bombarded three times at a pressure of 1.5 bar in a vacuum chamber at -0.8 bar. The cells were first cultured for 16–24 hours at 26 °C in complete darkness following the bombardment before being examined under a fluorescence microscope.

2.2.11. Precursor feeding, combination and co-cultivation experiments

The BY-2 cell suspension cultures were fed a variety of precursors, including Phenylalanine, Cinnamic acid, Coumaric acid (Sigma-Aldrich). The 3-day-old transgenic suspension cells were treated with a final concentration of 0.5 mM of the precursor. Cell cultures were incubated at 26 °C with continual shaking at 150 rpm. The identical dose of each precursor in control experiment sets were applied to wild type BY-2 cells. Finally, the cells and supernatant of seventh-day cells were used for HPLC analysis. For this purpose, a Büchner funnel and quick vacuum filtration was utilized to collect the dry cell samples and supernatant for further HPLC analysis. After being dissolved in 10 ml of methanol, 0.2 gr dry cells were homogenized in an ultrasonic mixer for 1 hour before being kept apart from the gathered supernatant at 4 °C (Figure 2.1).

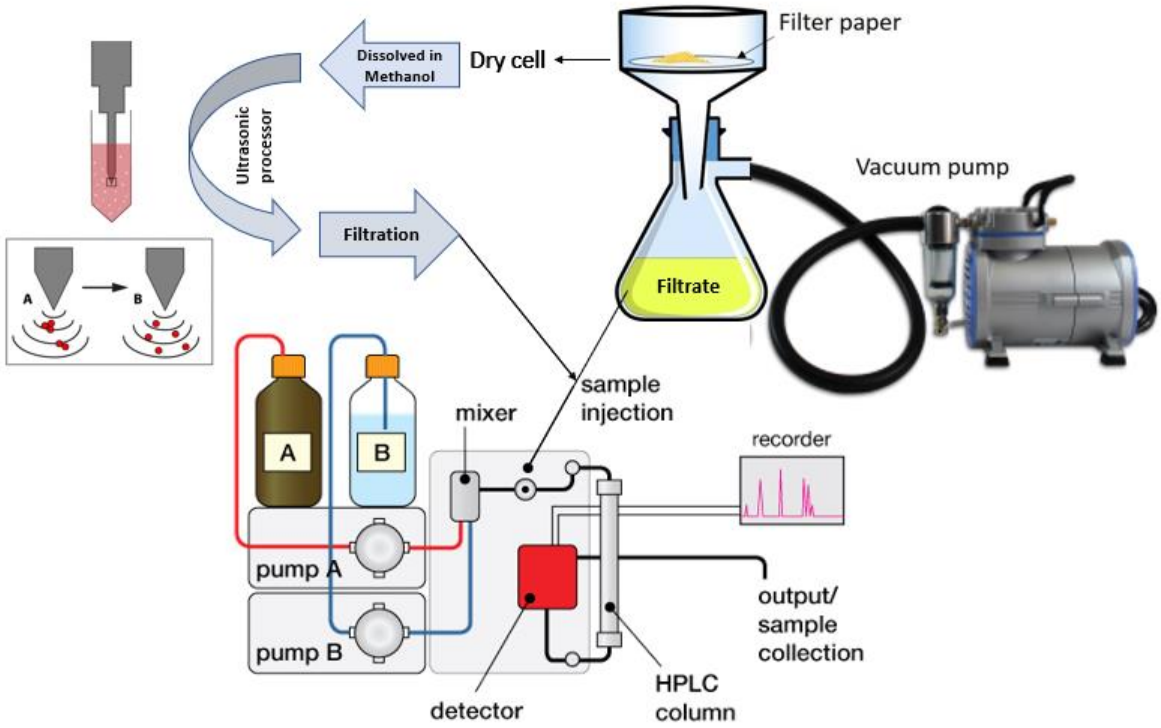


Figure 2.1. The process for preparing samples for HPLC analysis

2.2.12. Separation of alkaloids by High-Performance Liquid Chromatography

The Phenomenex Gemini-NX 5 C18 110A 150 mm x 4.6 mm column (Phenomenex, http://www.phenomenex.com/) was utilized with an Agilent-1200-Series HPLC system equipped

with a diode array detector (G1315D), Agilent ChemStation software, and a column temperature of 30 °C with a flow rate of 0.5 ml min-1. For both cell and medium extracts, the injection volume was 10 µl, and the peaks were measured at 260 and 310 nm. From 200 nm to 700 nm in wavelength, UV spectra were gathered.

	Amount	Solvents
B	20.0%	Methanol
C	0.1%	Formic acid
D	79.9%	Water

Table 2.18. HPLC Solvents and amounts

	Time	%B	%C	%D	Flow(mL/min)	Max press. (bar)
1	0.00	20.0	0.1	79.9	0.500	400
2	5.00	55.0	0.1	44.9	0.500	400
3	14.00	90.0	0.1	9.9	0.500	400
4	20.00	90.0	0.1	9.9	0.500	400
5	25.00	20.0	0.1	79.9	0.500	400
6	30.00	20.0	0.1	79.9	0.500	400

Table 2.19. HPLC conditions

2.2.13. RNA extraction and cDNA synthesis

A Büchner funnel was used to filter off the liquid medium from 2 milliliters of cycling BY-2 WT and transgenic cell lines (seven days after subcultivation; 100 mg of cells). The cells were transferred into a 2 ml reaction tube with a spatula, immediately placed into liquid nitrogen, and then ground in a Tissuelyser (Qiagen, (Http://Www.Qiagen.Com.)) with a 5 mm steel bead. Using an enhanced RNA extraction technique created by Universal RNA Purification Kit, total RNA was extracted following the manufacturer's instructions. (Roboklon, (Http://Www. Roboklon.Com,)). Electrophoresis on a 1% (w/v) agarose gel was used to verify the RNA preparation's integrity and purity. The concentration of RNAs were measured by the means of Nanodrop (peqlab Biotechnologie GmbH, ND-1000 Spectrophotometer, Erlangen, Germany). Using 100 ng of template RNA and the SuperScript® II Reverse Transcriptase (Https://Www.Lifetechnologies.Com), RNA was converted into cDNA. Reverse transcription was carried out on 1 µg of RNA using the reverse transcriptase M-MuLV to synthesize cDNA. To achieve the desired final RNA concentration of 1 µg for this process, a primarily mastermix (table

2.20) containing 0,4 µl of 10 mM dNTPs and 1 µl of oligo dT (100 µM) was mixed with the corresponding amount of RNA. With RNase-free water, the reaction was adjusted to a total volume of 16 µl depending on the concentration and amount of RNA. The 4 µl second mastermix (table2.20), consisting of 2 µl of reverse transcriptase buffer, 0.25 µl reverse transcriptase (MuLV), 0.50 µl RNase inhibitor, and 1.25 µl RNase-free H2O, was added after the first master mix had been incubated at 65°C for 5 min. The reaction was continued at 42°C for 60 min, then at 90°C for 10 min, and finally at 12°C for the final storage phase (table 2.20).

mastermix I	
dNTPs (10 mM)	1 µl
oligodT (100 µM)	0.4 µl
cycler step I	
70 °C	5 min
mastermix II	
MulV RT buffer	2 µl
RNase inhibitor	0.5 µl
MulV RT	0.25 µl
nuclease free water	1.25 µl
cycler step II	
42 °C	60 min
90 °C	10 min
12 °C	∞

Table 2.20. Protocol and cycler program for cDNA synthesis

2.2.14. Gene regulation analysis with quantitative PCR (qPCR)

In order to look into any potential post-transcriptional gene regulation in tobacco, the transcript levels of the first two genes of the phenylpropanoid pathway (PAL and C4H) were measured by quantitative real-time PCR (qPCR) using the CFX96 TouchTM Real-Time PCR Detection System from Bio-Rad Laboratories GmbH (Munich) and the qPCR primers listed in table 2.21.

Gene	Forward Or Reverse	Primer sequence	Size of transcript [bp]	Annealing Tm [°C]
PAL1	Sense	5‘ATGCGAAGTCCTGAAAACTAACAA 3’	104 bp	60
	Antisense	5‘TGATAGAGAGAAGCCAGTAGCCC 3’		
PAL2	Sense	5‘GAGAAGATGTTTGGAGAAGTGC 3’	100 bp	60
	Antisense	5‘AGGTGACTGTGGAGGATGTAGA 3’		
PAL3	Sense	5‘GTCTCCACAGTCGCCAAAAAAA 3’	161 bp	59
	Antisense	5’ AGGTTCTGCATTAAGGGGTAGT 3’		
PAL4	Sense	5‘CGTGAGACATGAACTTGAAACC 3’	140 bp	60
	Antisense	5‘CATTCCAACCTTCCACACACTT 3’		
C4H1	Sense	5‘ATCCTGAATGGTGGGACAAA 3’	162 bp	60
	Antisense	5‘CAATGGACAAGGCCAAAATG 3’		
C4H2	Sense	5‘TCCCATATTAGCACTCTCCAT 3’	145 bp	60
	Antisense	5‘GGGCTTGGCAATGATGAGA 3’		
C4H3	Sense	5‘GGCGGTAGGAGGCAAAGT 3’	167 bp	60
	Antisense	5‘CACTTACATCGATCTTCTGCTCA 3’		
C4H4	Sense	5‘TCCTGAATGGTGGGACAGAC 3’	168 bp	60
	Antisense	5‘AGCCTTCCAATGGCCAAG 3’		

Table 2.21. qPCR primer sequences for external *Cephalotaxus hainanensis´* overexpressed genes in transgenic cell lines

Gene	Forward Or Reverse	Primer sequence	Size of transcript [bp]	Annealing Tm [°C]
Spec.WT.PAL	Sense	5‘TCTGAAGGGGCAAGAGCTG3’	141 bp	60
	Antisense	5‘GTGACATGACTCTGTTCCATTG3’		
Spec.WT.C4H	Sense	5‘CCGATTCCTGTGCCAGTTTT3’	136 bp	60
	Antisense	5‘TTCTTTGGCTAGTTCAGGGG3’		

Table 2.22. qPCR primer sequences for internal *Nicotiana tabacum´* genes in transgenic cell lines

The qPCR analysis was performed in 20 µl reaction tubes with final concentrations of 200 nM of each primer and dNTP, 1x GoTaq Buffer, 2.5 mM MgCl2, 0.5 units of GoTaq polymerase, 1x SYBR green, and 1 l of 1:20 cDNA dilutions (table 2.23).

Component	volume [µl]
5x GoTaq Puffer buffer	4
5 mM each dNTP Mix	0.4
10 µM primer forward	0.4
10 µM primer forward	0.4
0.5 U/µl GoTaq Polymerase (Promega)	0.1
SybrGreen	0.95
cDNA template (1:10)	1
MgCl2 (50mM)	1
dd H2O	11.75
Final volume	**20**

Table 2.23. qPCR reaction components

For the qPCR experiment, the thermal cycler's operating parameters were as follows: a 3 minutes initial denaturation at 95 °C, followed by 39 cycles of 15 s at 95 °C, 40 s at 60 °C, and a plate read step, finalizing with a 10 sec denaturation step at 95°C for the melting curve. The melting curve was performed using 60 total repeats of 5 second steps that were initiated at 65°C and increased by 0.5°C every step.

Temperature	Duration	Repeats
95 °C	3 min	1x
95 °C	15 s	39x
60°C	40 s	
Plate read		

Table 2.24. qPCR thermal cycler's condition

Housekeeping genes Ribosomal protein L25 and Elongation factor (EF) which demonstrated the highest expression stability have been chosen as an internal standard, and three technical replicates of the identical procedure have been carried out for each of the three biological replicates.

By averaging the normalized ct values (G), which serve as an internal standard for the corresponding sample, and subtracting this value from the housekeeping gene reference ct value (H), the relative expression of genes was determined.

Formula: $\Delta ct = ct(G) - ct(H)$

The relative gene expression value was calculated using the resulting difference and the $2^{-\Delta ct}$ algorithm. This number has been used to show the expression patterns of the genes in presence or absence of treatments under tight regulation of a constitutive 35S promoter. Any differences would be as a result of post transcriptional regulation of the genes in phenylpropanoid pathway. The standard deviation for each triple pair of biological replicates has been determined to test the significance.

3. Results

Paul Srere (1967) analyzed enzyme concentrations in tissues in his famous Science article and came to the conclusion that the local concentrations reported in vivo are hundreds of times greater than those employed in vitro (Srere, 1967). This was nothing more than reclaiming space as a category that is relevant to molecular biology as well. As a result, if one wants to comprehend and control biosynthetic pathways, the structure of molecules and the compartmentalization of metabolic activity become extremely important (Nick, 2016). Therefore, the current research started with the sub-cellular localisation of enzymes involved in the production pathway of 1-Phenylethylisoquinoline scaffold. To achieve this, 21 genes (four PALs, four C4Hs, two 4CLs, five CCRs, three DBRs, one TyDC, one PPO, and one PR10) listed in table 3.1 were isolated from *Cephalotaxus hainanensis* cDNA and then were inserted in to the expression vector pK7FWG2.0 (Karimi et al., 2002) with C-terminal fusion of green fluorescent protein (GFP) as a reporter under the control of the CaMV 35S promoter which were later transformed in to *Nicotiana tabacum* (tobacco BY-2 cells) using stable transformation techniques via *Agrobacterium tumefaciens*. The expression vector's (pK7FWG2.0) map, used in this investigation is shown in Figure 3.1.

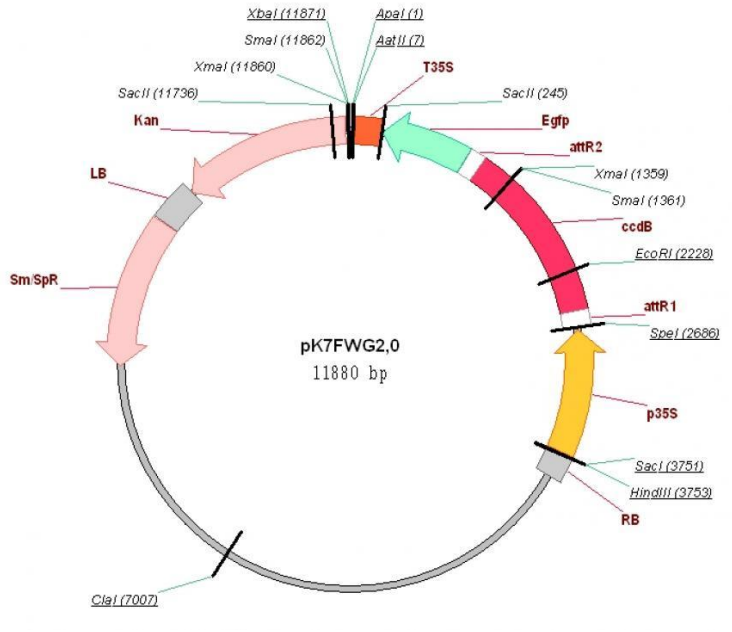


Figure 3.1. pK7FWG2.0 expression vector's map, used in this investigation

The sub-cellular localisation studies then were followed by in vivo studies regarding metabolic production of entry enzyme (PAL) in this pathway as the most important enzyme regulating the metabolic flux into the phenylpropanoid pathway and consequently, towards the production of 1-Phenylethylisoquinoline scaffold. In order to do this, the production of PAL over expressed cell lines was performed by high-performance chromatography (HPLC) after being supplied with proper precursors. The effect of precursor supplement on the expression of endogenous and exogenous PAL and C4H in transgenic cell lines was then investigated by relative gene expression analysis.

Gene name	Abbreviation	Full-length cDNA sequence obtained	Names
Phenylalanine ammonia-lyase	PAL	4	▪ PAL-1 (2235bp) ▪ PAL-2 (2271bp) ▪ PAL-3 (2172bp) ▪ PAL-4 (2139bp)
Cinnamate 4-hydroxylase	C4H	4	▪ C4H-1 (1506bp) ▪ C4H-2 (1551bp) ▪ C4H-3 (1566bp) ▪ C4H-4 (1500bp)
4-Coumarate CoA ligase	4CL	2	▪ 4CL-1 (1629bp) ▪ 4CL-2 (1671bp)
Cinnamoyl CoA reductase	CCR	5	▪ CCR-1 (903bp) ▪ CCR-2 (966bp) ▪ CCR-3 (981bp) ▪ CCR-4 (975bp) ▪ CCR-5 (1044bp)
NADPH-dependent double bond reductases	DBR	3	▪ DBR-1 (1047bp) ▪ DBR-2 (1032bp) ▪ DBR-3 (1035bp)
Tyrosine / Dopa decarboxylase	TyDC	1	▪ TyDC-1 (1506bp)
Polyphenol Oxidase	PPO	1	▪ PPO-1 (1674bp)
Norcoclaurine synthase-like	NCS-like (PR10-0843)	1	▪ PR10-0843 (483bp)

Table 3.1. Isolated *Cephalotaxus hainanensis* enzymes and number of their isoforms used in this study

As described in table 1.3 eight different enzymes are involved in converting Phenylalanine and Tyrosine as primary precursors to the 1-Phenylethylisoquinoline scaffold. The lab of Dr. Fei Qiao in Hainan has carried out a high throughput RNA-seq analysis and reconstructed the transcriptome de novo to provide molecular access to the biosynthesis process of *Cephalotaxus* ester alkaloids. 4.3 Gbp of raw readings were assembled de novo into over 40 000 unigenes with a mean length of 1 kb,

for an overall assembly size of 45.8 Mbp (deposited in the GenBank under accession SRR1509462). A framework for further analysis of gene functions was created by annotating the reconstructed unigenes with regard to their possible function (Qiao et al., 2014c). In their research, they discovered distinct isoforms for some genes in the pathway of 1-Phenylethylisoquinoline scaffold production, and in their most recent study, they explored the function of all these isoforms by heterologous expression of the extracted genes from *Cephalotaxus hainanensis* in *Saccharomyces cerevisiae* after feeding them with the appropriate precursor (Zhao et al. 2022). The number of different isoforms for each gene which were extracted from *Cephalotaxus hainanensis* cDNA and then were labeled with GFP in this study were shown in table 3.1.

3.1. Sub-cellular localisation of enzymes involved in 1-Phenylethylisoquinoline scaffold production pathway

In figure 3.2 the upstream pathway of Phenylalanine and Tyrosine production in plants were shown. As depicted in this picture these two amino acids are synthesized through the Enbden-Meyerhof-Parnas (EMP), Hexose monophosphate (HMP), and shikimate pathways (Y.-C. Li, 2014). In fact, phosphoenolpyruvate from the glycolysis pathway and erythrose 4-phosphate from the pentose phosphate pathway both enter the shikimate pathway in plastids, where they are converted to Phenylalanine and Tyrosine.

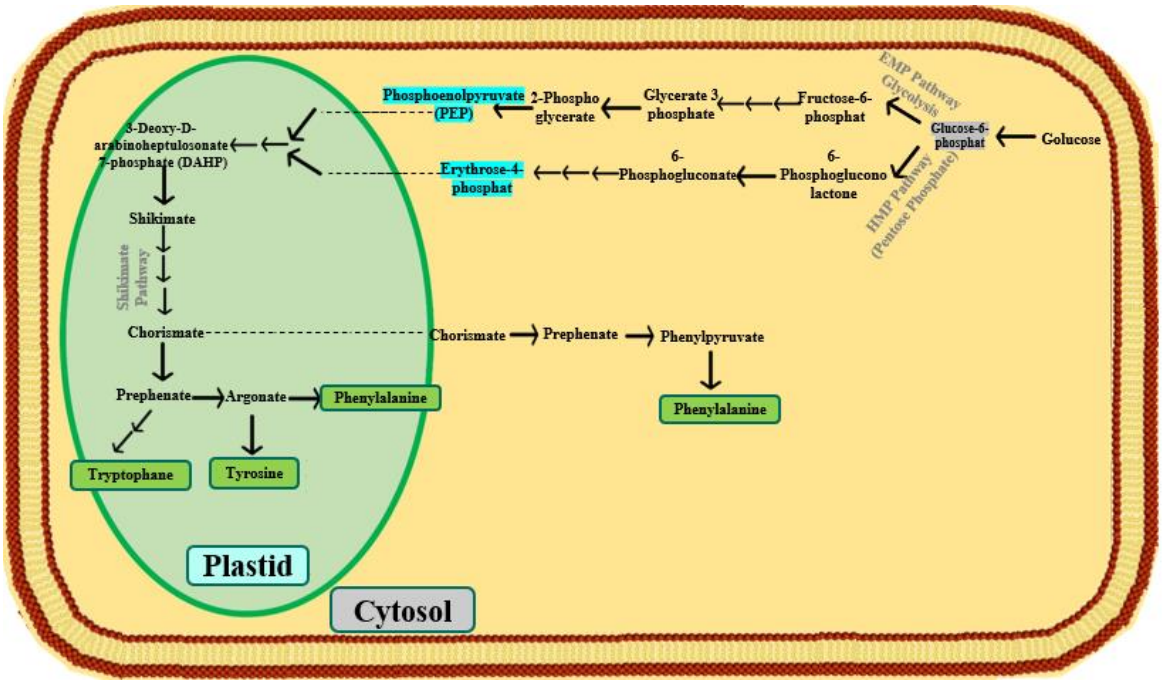


Figure 3.2. Upstream biosynthesis pathways for aromatic amino acids production

3.1.1. Sub-cellular localisation of *Ch.* Phenylalanine ammonia-lyase (*Ch*.PAL) using C-terminal GFP fusions and stable expression in *Nicotiana tabacum* followed by spinning disc confocal microscopy

The first enzyme in the 1-Phenylethylisoquinoline scaffold production pathway is PAL which is responsible for converting Phenylalanine to Cinnamic acid. The Phenylalanine could either be produced in cytosol from phenylpyruvate or in plastids from Arogenate (Y.-C. Li, 2014). The four stable *Ch*. PAL overexpressed *Nicotiana tabacum* cell lines were observed by the confocal microscope for the sub-cellular localisation analysis. Among the four isoform of PAL three of them (PAL2, PAL3 and PAL4) were shown to have the cytosolic localisation (Figure 3.3) while PAL1 showed a plastidic localisation (Figure 3.4).

3.1.1.1. Cytosolic PAL for converting Phe resulted from phenylpyruvate in cytosol

As shown in microscopy images in Fig. 3.3, bright green fluorescence was clearly detected, which was the signal of the *Ch*.PAL-GFP fusion. In the same condition, the wild-type tobacco BY-2 cells displayed no green fluorescence. In this picture a strong green fluorescence signal was distributed around the vacuole, indicating that the *Ch*.PAL-2, *Ch*.PAL-3 and *Ch*.PAL-4 proteins were localized in the cytoplasm.

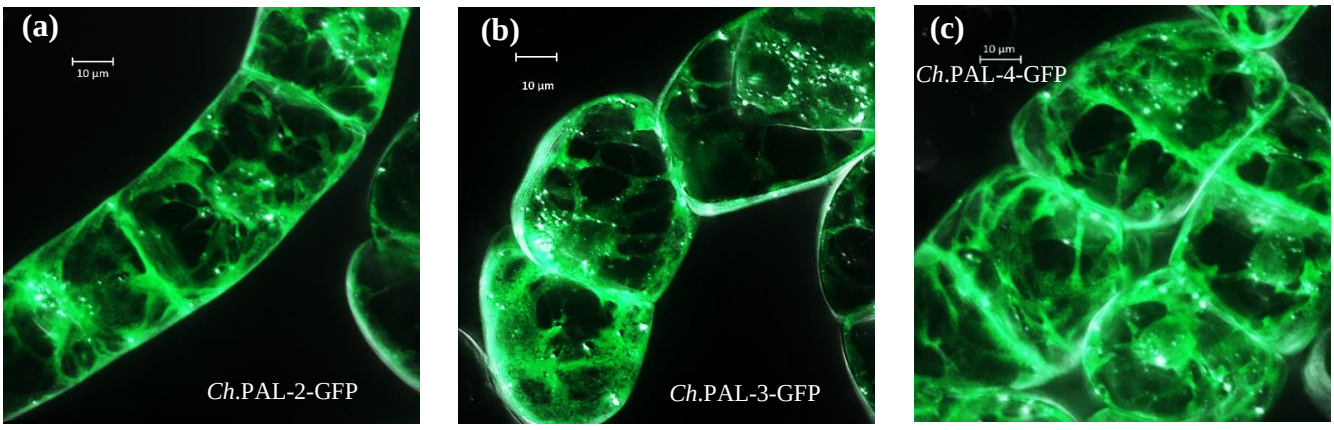


Figure 3.3. Sub-cellular localisation results of *Ch*.PAL-2-GFP (a), *Ch*.PAL-3-GFP (b), and *Ch*.PAL-4-GFP (c) in BY-2 cells overexpressing corresponding enzymes using Confocal Microscopy (Bars = 10 µm)

3.1.1.2. Plastidic PAL for converting Phe resulted from Arogenate in plastids

The *Ch*.PAL-1-GFP on the other hand showed a small round shape pattern localisation which were different with the three other PALs. The tpFNRmEos (Schattat et al., 2012) (a generous gift from Professor Dr. Jaideep Mathur, Guelph University, Canada), was transiently transformed into the *Ch*-PAL-1-GFP overexpressors using the *Agrobacterium* based protocol (Buschmann et al., 2011). This was done to test for a potential co-localisation of the fusion protein with plastids. The tight colocalisation of the tpFNRmEos and the GFP signal (Figure 3.4) provides evidence for a plastidic localisation of the overexpressed *Ch*. PAL-1-GFP.

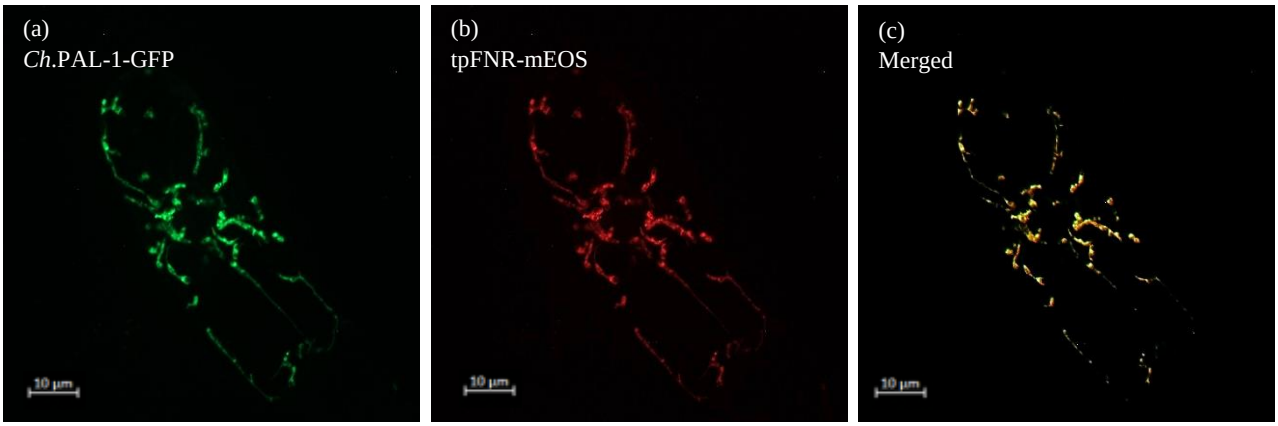


Figure 3.4. Colocalisation of *Ch*.PAL-1-GFP and plastid marker tpFNR-mEOS in BY-2 cells overexpressing *Ch*.PAL-1-GFP showed plastid localisation of *Ch*.PAL-1-GFP. (a) *Ch*.PAL-1-GFP signal, (b) plastid marker tpFNR-mEOS, (c) merged signal of both channels showing the tight colocalisation of the PAL-1 with plastids. (Bars = 10 µm)

3.1.2. Sub-cellular localisation of *Ch.* Cinnamate 4-hydroxylase (*Ch*.C4H) using C-terminal GFP fusions and stable expression in *Nicotiana tabacum* followed by spinning disc confocal microscopy

C4H, the second enzyme in the production pathway of the 1-Phenylethylisoquinoline scaffold, is in charge of converting Cinnamic acid into Coumaric acid. For the sub-cellular localisation study the confocal microscope was used to observe the four stably transformed *Ch*. C4H gene which were overexpressed in *Nicotiana tabacum*. Bright green fluorescence, which was the signal of the Ch.C4H-GFP fusion, was plainly identified, as shown in microscopy photos in Fig. 3.5 while the wild-type tobacco BY-2 cells did not exhibit green fluorescence under the same circumstances.

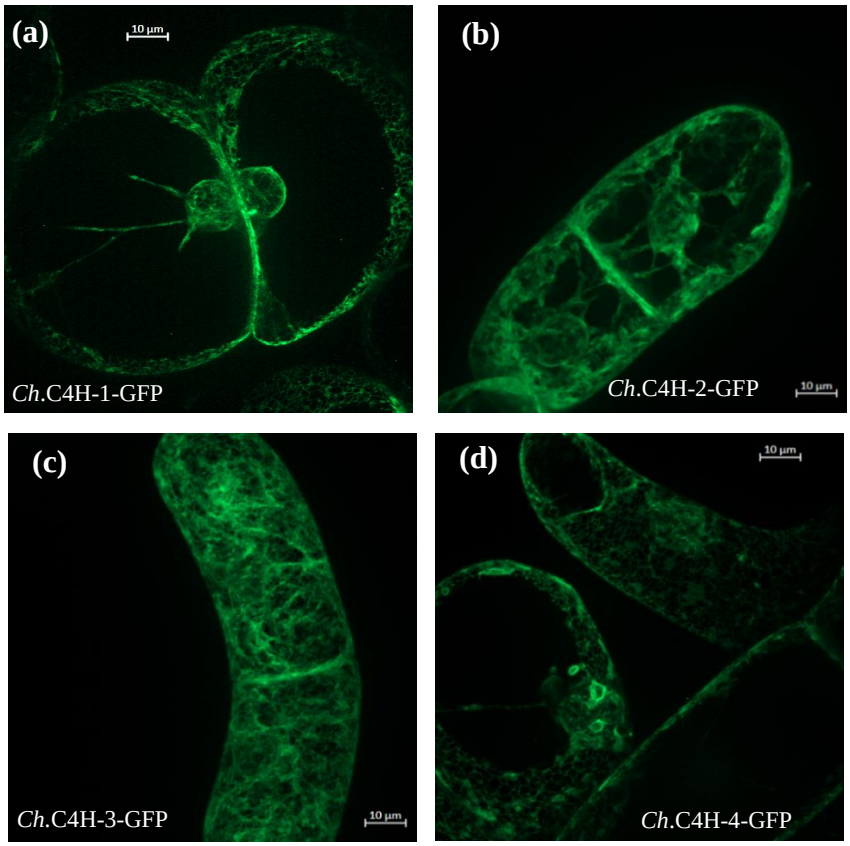


Figure 3.5. Sub-cellular localisation results of *Ch*.C4H-1-GFP (a), *Ch*.C4H-2-GFP (b), *Ch*.C4H-3-GFP (c) and *Ch*.C4H-4-GFP (d) in BY-2 cells overexpressing corresponding enzymes using Confocal Microscopy (Bars = 10 µm)

The reticulated GFP pattern of all four *Ch*.C4H-GFP isoforms in *Nicotiana tabacum* displayed a possible Endoplasmic reticulum localisation. A rhodamine-labeled variant of the ER tracker was utilized to confirm whether this pattern was generated by a localisation in the endoplasmic reticulum. A close colocalisation of both signals supported the ER localisation of C4Hs (Figure 3.6).

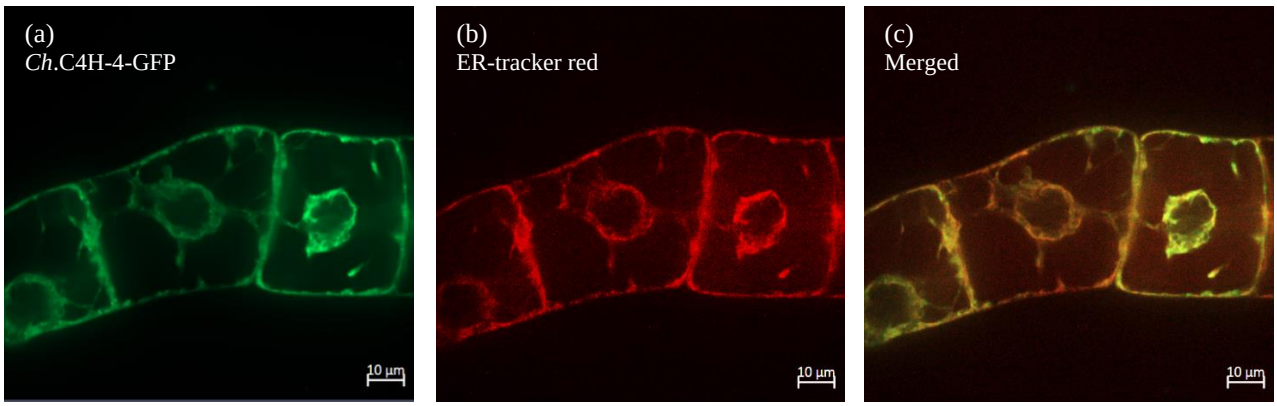


Figure 3.6**.** Colocalisation of *Ch*.C4H-GFP and rhodamine-conjugated ER-tracker in BY-2 cells overexpressing *Ch*.C4H-GFP showed endoplasmic reticulum localisation of this protein. (a) *Ch*.C4H-GFP signal, (b) rhodamine-conjugated ER-tracker, (c) merged signal of both channels showing the tight colocalisation of the Cinnamate 4-hydroxylase with the endoplasmic reticulum. (Bars = 10 µm)

3.1.3. Sub-cellular localisation of *Ch.* 4-Coumarate-CoA ligase (*Ch*.4CL) using C-terminal GFP fusions and stable expression in *Nicotiana tabacum* followed by spinning disc confocal microscopy

P-Coumaric acid is transformed into P-Coumaroyl-CoA by 4CL, the third enzyme in the route that produces the 1-Phenylethylisoquinoline scaffold. For the sub-cellular localisation research, two suspension culture of stable transformed *Ch*.4CL overexpressed cell lines were observed under the confocal microscope. The bright green fluorescence signal of *Ch*.4CL-GFP was clearly visible in the microscopy images in Fig. 3.7. In contrast, the wild-type tobacco BY-2 cells did not exhibit green fluorescence under the same conditions.

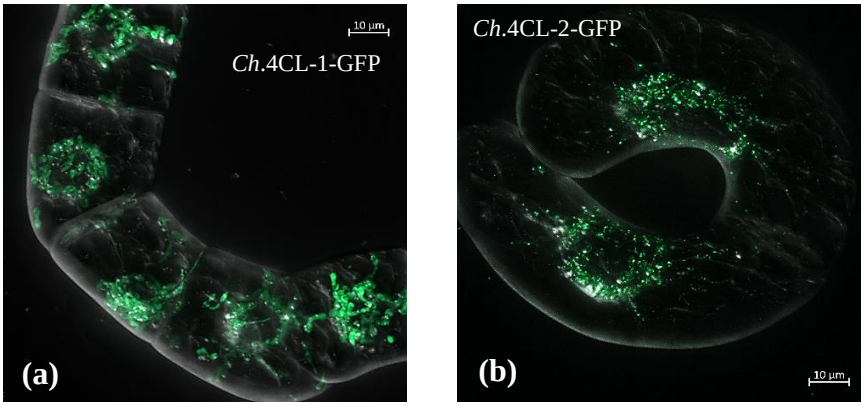


Figure 3.7. Sub-cellular localisation results of *Ch*.4CL-1-GFP (a) and *Ch*.4CL-2-GFP (b) in BY-2 cells overexpressing corresponding enzymes using Confocal Microscopy (Bars = 10 µm)

In contrast to the *Ch*.4CL-2 over expressor line localisation, the *Ch*.4CL-1-GFP displayed a circular shape pattern localisation around the nucleus. The circular shape localisation pattern surrounding the nucleus appears to be plastidic. This hypothesis was verified by transiently transforming the plastid marker tpFNRmEos (Schattat et al., 2012) into *Ch*-4CL-1-GFP overexpressors utilizing an Agrobacterium-based technique (Buschmann et al., 2011). This was done to see if the fusion protein could co-localize with plastids. The close colocalisation of the tpFNRmEos and the GFP signal (Figure 3.8) suggests that the overexpressed *Ch*. 4CL-1-GFP is plastidic.

The *Ch*. 4CL-2-GFP over expressed cell lines exhibited punctuated signals suggesting that the GFP fusion of this enzyme was located in small organelles. To clarify the sub-cellular localisation of *Ch*.4CLs, a peroxisome-targeted mCherry marker (PTS1-Cherry) (Ching et al., 2012) was transiently introduced into the background of *Ch*. 4CL-2-GFP over expressed cell line using the Agrobacterium mediated protocol (Buschmann et al., 2011). The tight colocalisation of the mCherry

and the GFP signal (Figure 3.9) provides evidence for a correct localisation of the overexpressed *Ch*. 4CL-2-GFP in peroxisomes.

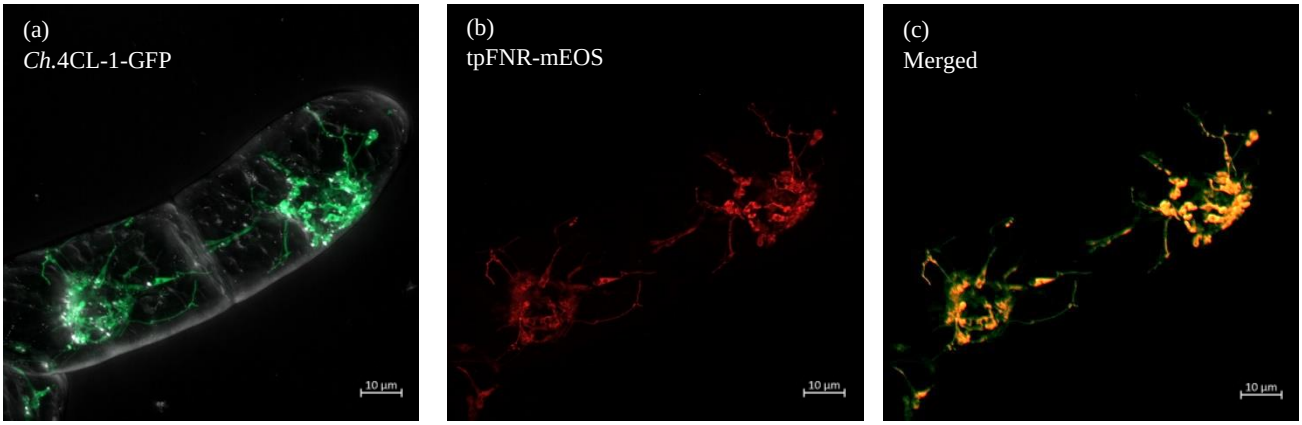


Figure 3.8. Colocalisation of *Ch*.4CL-1-GFP and plastid marker tpFNR-mEOS in BY-2 cells overexpressing *Ch*.4CL-1-GFP showed plastid localisation of this protein. (a) *Ch*.4CL-1-GFP signal, (b) plastid marker tpFNR-mEOS, (c) merged signal of both channels showing the tight colocalisation of the *Ch*.4CL-1 with plastids. (Bars = 10 µm)

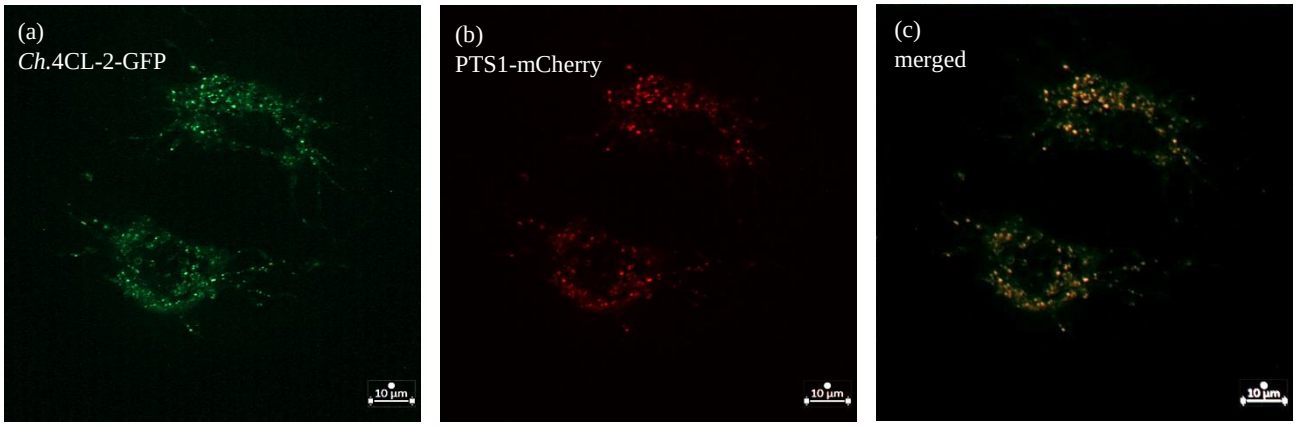


Figure 3.9. Colocalisation of *Ch*.4CL-2-GFP and peroxisomal marker PTS1-mCherry in BY-2 cells overexpressing *Ch*.4CL-2-GFP showed Peroxisomal localisation of *Ch*.4CL-GFP. (a) *Ch*.4CL-2-GFP signal, (b) peroxisomal marker PTS1-mCherry, (c) merged signal of both channels showing the tight colocalisation of the 4-Coumarate-CoA ligase with peroxisomes. (Bars = 10 µm)

3.1.4. Sub-cellular localisation of *Ch.* Cinnamoyl-CoA reductase (*Ch*.CCR) using C-terminal GFP fusions and stable expression in *Nicotiana tabacum* followed by spinning disc confocal microscopy

Cinnamoyl-CoA reductase (CCR), the subsequent enzyme in the 1-Phenylethylisoquinoline scaffold synthesis pathway, catalyzes the conversion of a substituted p-Coumaroyl-CoA arising from 4CL to its corresponding Cinnamaldehyde while using NADPH and H+ and releasing free CoA and NADP+. Five *Nicotiana tabacum* suspension cultures of stable transformed *Ch*.CCR overexpressed cell lines were examined using a confocal microscope for the sub-cellular localisation research. In the microscope pictures shown in Fig. 3.10, the *Ch*.CCR-GFP fluorescence signal, which is bright green, is readily discernible. Contrarily, under the identical circumstances, the wild-type tobacco BY-2 cells did not show any green fluorescence. The *Ch*.CCR proteins were localized in the cytoplasm, according to the strong green fluorescence signal that was present in all samples and was dispersed around the vacuole.

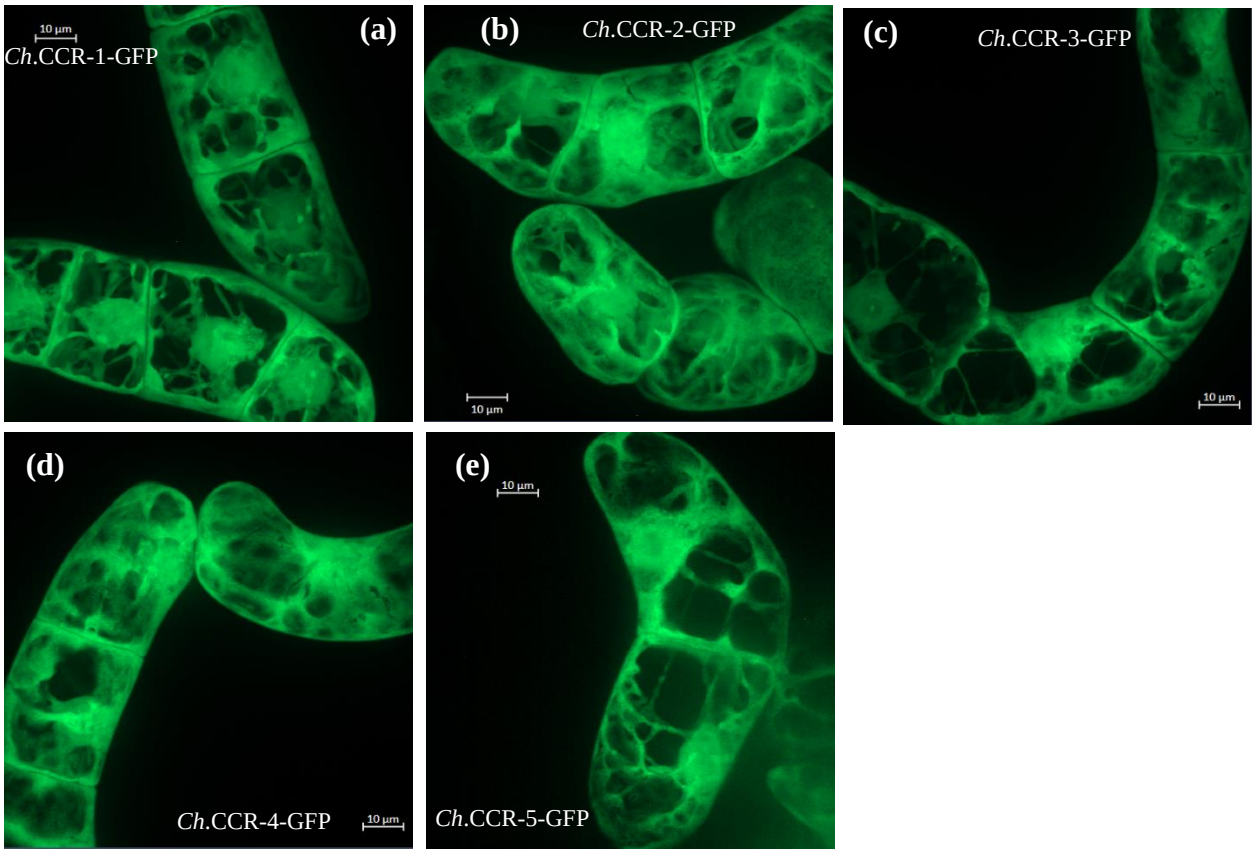


Figure 3.10. Sub-cellular localisation results of *Ch*.CCR-1-GFP (a), *Ch*.CCR-2-GFP (b), *Ch*.CCR-3-GFP (c) *Ch*.CCR-4-GFP (d) and *Ch*.CCR-5-GFP (e) in BY-2 cells overexpressing corresponding enzymes using Confocal Microscopy (Bars = 10 µm)

3.1.5. Sub-cellular localisation of *Ch.* Alkenal double bond reductase (*Ch*.DBR) using C-terminal GFP fusions and stable expression in *Nicotiana tabacum* followed by spinning disc confocal microscopy

The final enzyme in the Phenylalanine branch of 1-Phenylethylisoquinoline scaffold production pathway is alkenal double bond reductase (DBR), which catalyzes the NADPH-dependent reduction of the unsaturated double bond of numerous secondary metabolites. NADPH dependent double bond reductases (DBRs) mediated the 7,8-double bond reduction of Phenylpropanal substrates, p-Coumaroyl aldehyde, to their corresponding dihydro molecules, Phenylpropyl aldehyde, in the hypothesized *Cephalotaxus* alkaloids biosynthesis pathway. In this study, three alkenal double bond reductase genes that had been isolated from *Cephalotaxus hainanensis* cDNA and were stably transformed and overexpressed in *Nicotiana tabacum*'s cells after being labelled with GFP. For the sub-cellular localisation study, the stable transgenic cell lines were studied under a confocal microscope. The vivid green *Ch*.DBR-GFP fluorescence signal is clearly visible in the microscope images displayed in Fig. 3.11. The wild-type tobacco BY-2 cells, on the other hand, did not exhibit any green fluorescence under the same conditions. The strong green fluorescence signal, which was present in all samples and was spread around the vacuole, indicated that the *Ch*.DBR proteins were localized in the cytoplasm.

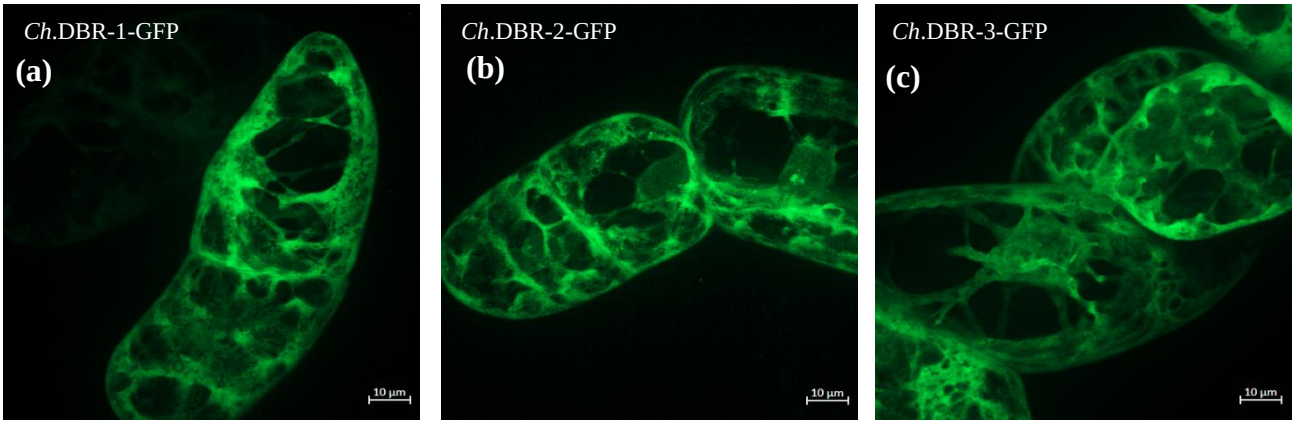


Figure 3.11. Sub-cellular localisation results of *Ch*.DBR-1-GFP (a), *Ch*.DBR-2-GFP (b) and *Ch*.DBR-3-GFP (c) in BY-2 cells overexpressing corresponding enzymes using Confocal Microscopy (Bars = 10 µm)

3.1.6. Sub-cellular localisation of *Ch.* Tyrosine decarboxylase (*Ch*.TyDC) using C-terminal GFP fusions and stable expression in *Nicotiana tabacum* followed by spinning disc confocal microscopy

Tyrosine decarboxylase is the initial enzyme in a different branch of the 1-Phenylethylisoquinoline scaffold biosynthesis pathway that begins with Tyrosine which is converting this product into Tyramine. Only one Tyrosine decarboxylase gene was isolated from *Cephalotaxus hainanensis* cDNA for this investigation, and it was then overexpressed and stably transformed in *Nicotiana tabacum* cells after being labeled with GFP. Using a confocal microscope, the stable transgenic cell lines were investigated for the sub-cellular localisation investigation. The microscopy picture shown in Fig. 3.12 clearly indicates the bright green *Ch*.TyDC-GFP fluorescence signal but under the same conditions, no green fluorescence was observed in the wild-type tobacco BY-2 cells. The *Ch*.TyDC-1-GFP protein was observed to be localized in the cytoplasm, based on its strong green fluorescence signal, which was detected around the vacuole.

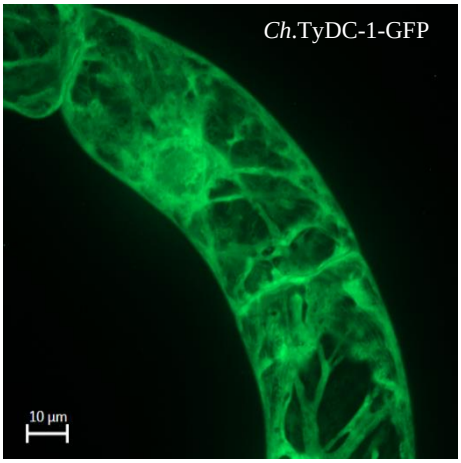


Figure 3.12. Sub-cellular localisation results of *Ch*.TyDC-1-GFP in BY-2 cells overexpressing corresponding enzymes using Confocal Microscopy (Bars = 10 µm)

3.1.7. Sub-cellular localisation of *Ch.* Polyphenol oxidase (*Ch*.PPO) using C-terminal GFP fusions and stable expression in *Nicotiana tabacum* followed by spinning disc confocal microscopy

The final enzyme in the 1-Phenylethylisoquinoline scaffold production route, *Ch*.PPO-1-GFP, which converts Tyramine to Dopamine. Using a confocal microscope, the stable transgenic *Ch*.PPO-1-GFP cell lines were examined for the sub-cellular localisation investigation. The microscope images shown in Fig. 3.13 clearly show the bright green *Ch*.PPO-1-GFP fluorescence signal while under the same conditions, no green fluorescence was seen in the wild-type tobacco BY-2 cells.

A possible plastidic localisation was predicted by the strong green fluorescence signal that was observed in small round-shaped organelles. To test this, the Agrobacterium-based technique (Buschmann et al., 2011) was used to transiently transform the plastid marker tpFNRmEos (Schattat et al., 2012) into *Ch*-PPO-1-GFP overexpressors background. The strong colocalisation of the tpFNRmEos and the GFP signal (Figure 3.13), supports the localisation of the overexpressed *Ch*. PPO-1-GFP in plastids.

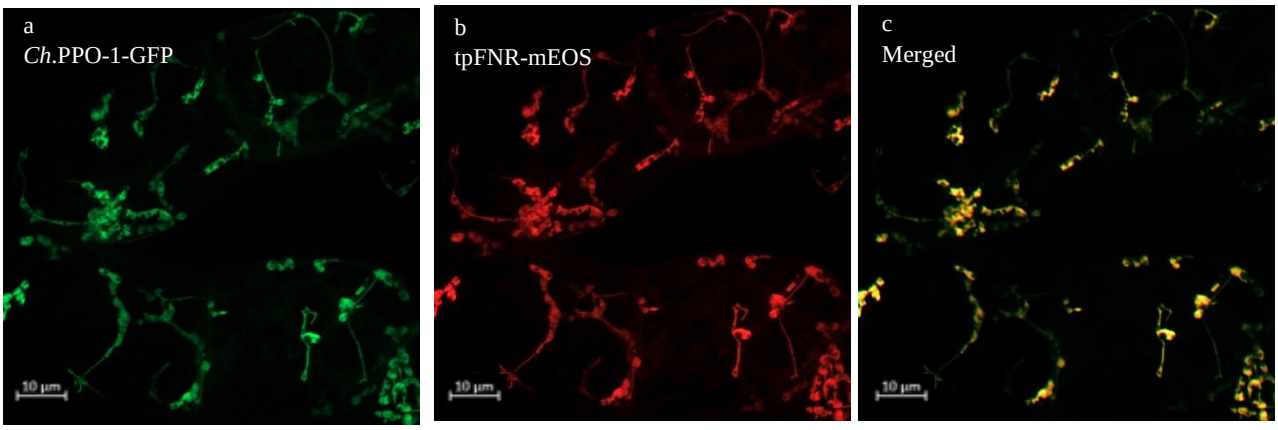


Figure 3.13. Colocalisation of *Ch*.PPO-1-GFP and plastid marker tpFNR-mEOS in BY-2 cells overexpressing *Ch*.PPO-1-GFP showed plastid localisation of *Ch*.PPO-1-GFP. (a) *Ch*.PPO-1-GFP signal, (b) plastid marker tpFNR-mEOS, (c) merged signal of both channels showing the tight colocalisation of the *Ch*.PPO-1 with plastids. (Bars = 10 µm)

3.1.8. Sub-cellular localisation of *Ch.* phenethylisoquinoline scaffold synthase (*Ch*.PSS) using C-terminal GFP fusions and stable expression in *Nicotiana tabacum* followed by high-end spinning disc confocal microscopy

Base on the confocal microscopy analysis of this study, *Ch*.PSS, the enzyme that catalyzes the Picket-Spengler condensation of dopamine and 4-HDCA, the products of the phenylalanine and tyrosine branches, is also located in the cytosol.

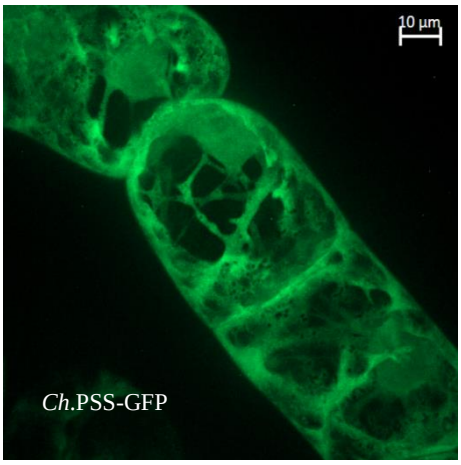


Figure 3.14. Sub-cellular localisation results of *Ch*.PSS-GFP in BY-2 cells overexpressing corresponding enzymes using Confocal Microscopy (Bars = 10 µm)

3.1.9. Summary of sub-cellular localisation of the functional genes in the 1-Phenylethylisoquinoline scaffold production pathway

A summary of the sub-cellular localisation analysis of the functional genes in the 1-Phenylethylisoquinoline scaffold production pathway, which were identified based on the confocal microscopy analysis in this study were shown in figure 3.15. These results reveals that various cellular compartments are involved in this pathway. The first three enzymes, which are the same enzymes as the three shared common enzymes in the general phenylpropanoid pathway —PAL, C4H, and 4CL — have the highest congestion in this route. In this study these enzymes were identified to be localized in the cytosol, endoplasmic reticulum, and peroxisomes, three separate cellular compartments. The results of colocalisation also revealed plastidic placement for one of the *Ch*.PAL isoforms (*Ch*.PAL-1) and one of the *Ch*.4CL isoforms (*Ch*.4CL-1) indicating a putative plastidic route for the synthesis of p-Coumaroyl-CoA. The remaining of the manufacturing route for 4-HDCA, the ultimate product of the Phenylalanine branch in creation of 1-Phenylethylisoquinoline scaffold, occurs in the cytosol. A part of the Tyrosine branch that leads to the final deliverable of Dopamine occurs in the cytosol, whereas another portion occurs in the plastids. These results suggest a possible presence of another isoform of Tyrosine decarboxylase (TyDC) in plastids which due to difficulties extracting all of the reported isoforms of TyDC from *Cephalotaxus hainanensis* cDNA could not be confirmed in this study.

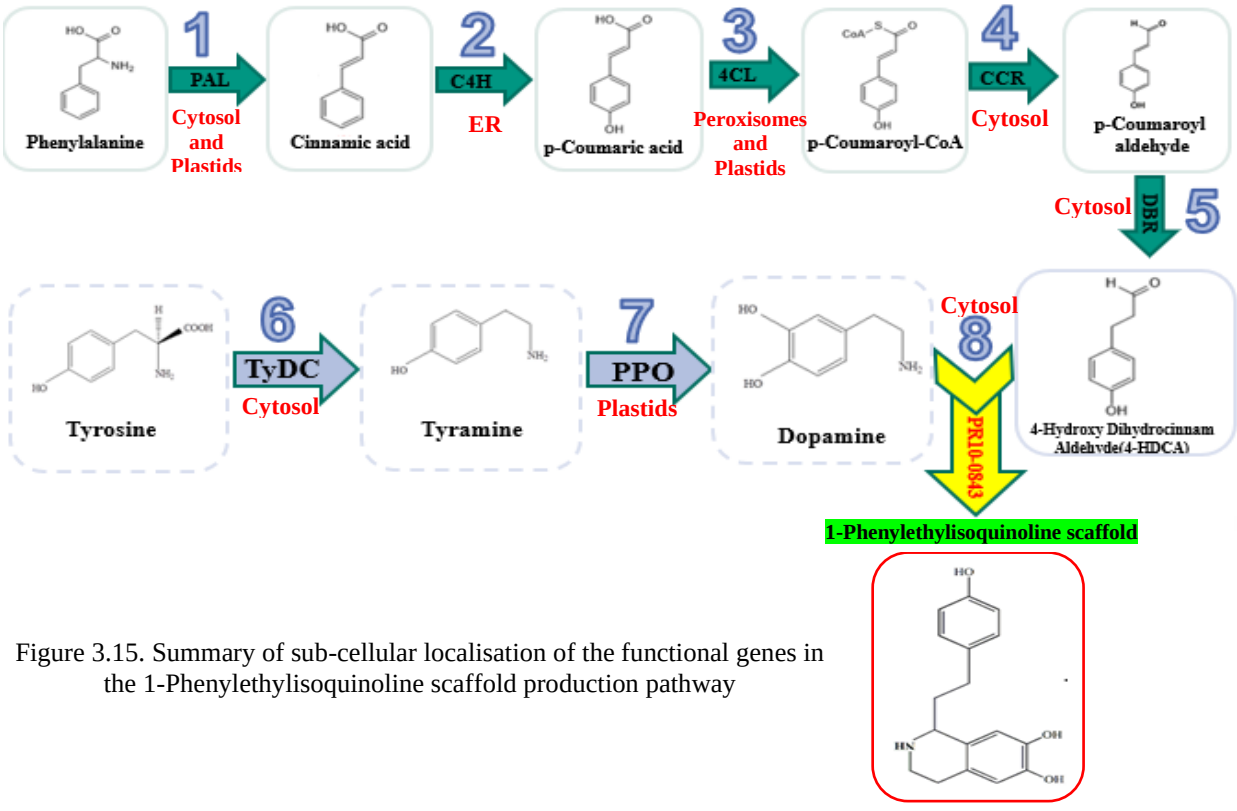


Figure 3.15. Summary of sub-cellular localisation of the functional genes in the 1-Phenylethylisoquinoline scaffold production pathway

3.2 Feeding studies in PAL overexpressor lines

High performance chromatography (HPLC) was used to analyze the metabolic activity of the PAL overexpressed cell lines in vivo. All of the isoforms of *Ch*. PAL were fed with the proper precursors before being harvested seven and fourteen days after inoculation. Prior to doing the HPLC analysis, the HPLC condition and flow through were adjusted to achieve the optimum results for metabolite separation. Based on the standards' HPLC analysis results (Figure 3.16), the peak for Caffeine acid, by collecting the UV spectra in 310 nm wavelength, was detected around 9.08 minutes, followed by those for Coumaric acid, Ferulic acid, and Cinnamic acid in 11.14, 11.42, and 14.82 minutes, respectively. The peak for Cinnamic acid was measured in 280 nm wavelength.

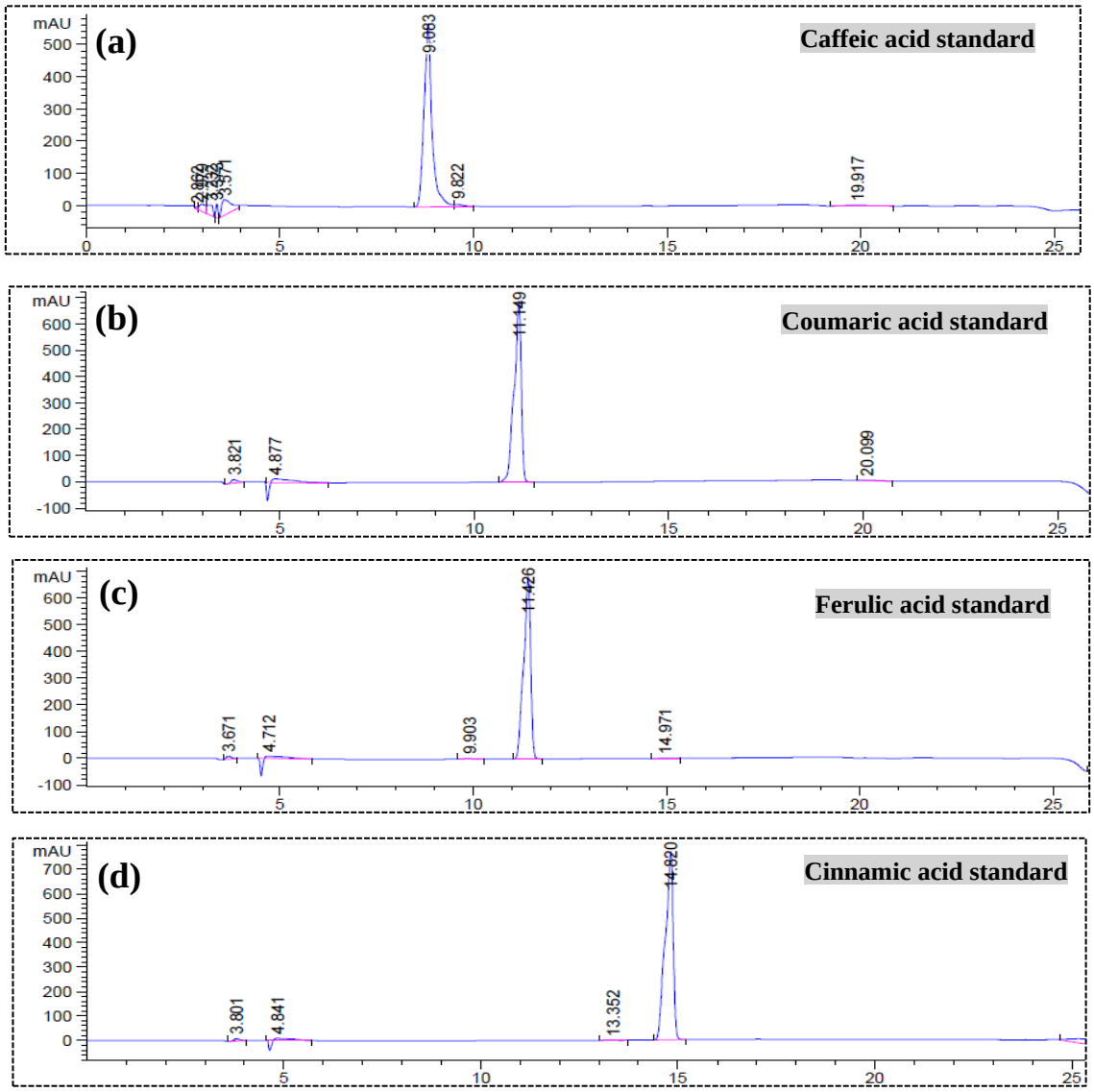
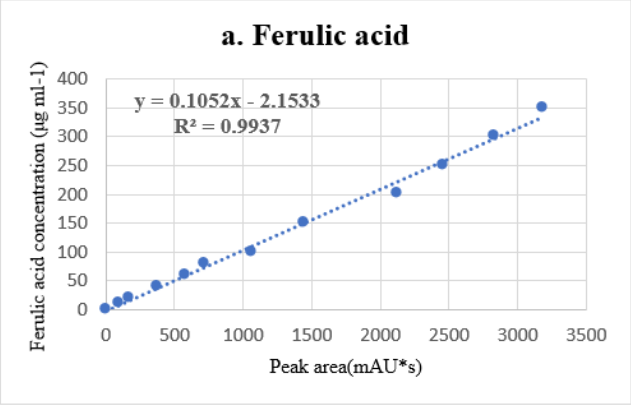
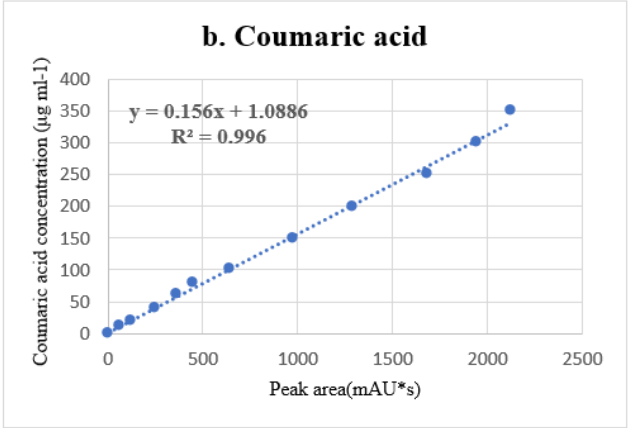


Figure 3.16. Chromatogram of pure standard solutions of Caffeic acid **(a)** Cinnamic acid **(b)**, Ferulic acid **(c)**, Diode array detection (HPLC-DAD; 310 nm) and Cinnamic acid **(d)** Diode array detection (HPLC-DAD; 280 nm)

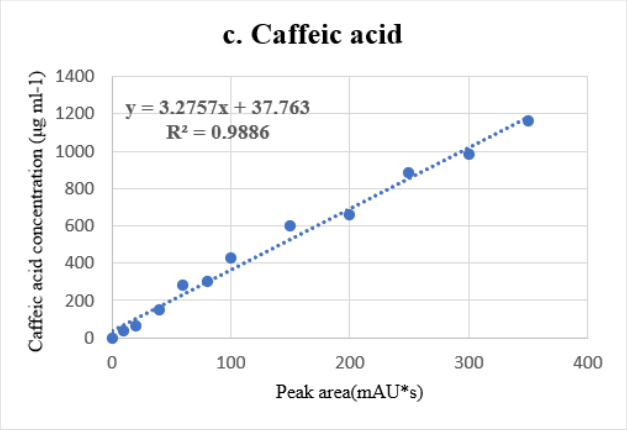
The amount of metabolite production in different isoforms were compared with each other and also with the wild type. To measure the metabolite synthesis of overexpressors, calibration curves were created by injecting 10 µl of several standard solutions (Caffeic acid, Coumaric acid, and Ferulic acid) with known concentrations, resulting in varying quantity and response data (peak area). To join the acquired curve's points, a line of best fit (regression line) was employed. The final metabolite concentration can be calculated using the regression analysis result from the calibration curve, which describes the line of best fit through the data points. According to the "correlation coefficient," which is the square root of the regression coefficient, all data points may be fitted to a straight line with reasonable accuracy. The intercept of the regression equation did not have a significant positive or negative value, which shows that the sample preparation and analysis were precise and accurate.



Ferulic acid (µg ml-1)	Peak area(mAU*s)
0	0
10	96.856429
20	179.6325
40	370.0893
60	586.45899
80	727.7867
100	1061.3347
150	1445.2689
200	2128.5698
250	2458.6598
300	2830.2819
350	3184.06714



Coumaric acid (µg ml-1)	Peak area(mAU*s)
0	0
10	64.849
20	124.698
40	248.4634
60	358.8569
80	446.792
100	644.895
150	973.5841
200	1290.5692
250	1688.12578
300	1947.16827
350	2129.7153



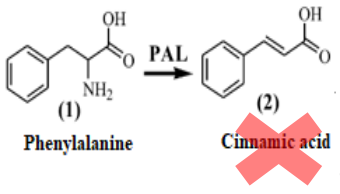
Caffeic acid (µg ml-1)	Peak area(mAU*s)
0	0
10	38.25698
20	66.26104
40	152.4005
60	284.657
80	304.80256
100	427.516
150	598.9562
200	657.3041
250	887.5691
300	985.9562
350	1159.5692

Figure 3.17. Calibration curves for quantitative analysis of Ferulic acid (a), Coumaric acid (b) and Caffeic acid (c) production in *Ch*.PAL-3 overexpressor lines.

3.2.1. Metabolic output of PAL over expressed cell lines

The same quantity of fresh weight (2 gr) of the seven-day-old tobacco BY2 cell cultures overexpressing *Ch*.PAL-GFP were inoculated in fresh medium and the Phenylalanine precursor was added to the final concentration of 0.5 mM in order to study the metabolic output of four different PAL overexpressor lines by high performance chromatography (HPLC) and compare the production of different isoforms. The injection volume was set at 10 µl. All of the peaks were measured at 310 nm. UV spectra were collected in the 200-700 nm wavelength range. Quantitative determination for any possible unknown peak was performed by an internal standard method with the use of 1-hydroxypyrene solutions from human urine which was added manually to each sample to the final concentration of 30 µg ml^{-1}.

3.2.1.1. Coumaric acid and Ferulic acid main products of PAL overexpressed cell lines after seven days incubation



After being fed with Phenylalanine, the HPLC results of four different *Ch*.PAL overexpressor lines revealed no Cinnamic acid production, but Ferulic acid and Coumaric acid production after seven seven days of incubation, while caffeic acid was the main production of the wild type tobacco BY2 cell line following Phenylalanine feeding and seven days of incubation.*Ch*.PAL-3 and *Ch*.PAL-4 overexpressed cell lines produced more product than the other two cell lines, according to the HPLC data in figure 3.19. The amount of metabolite production was calculated based on the standard curves´ equations on the figure 3.17.

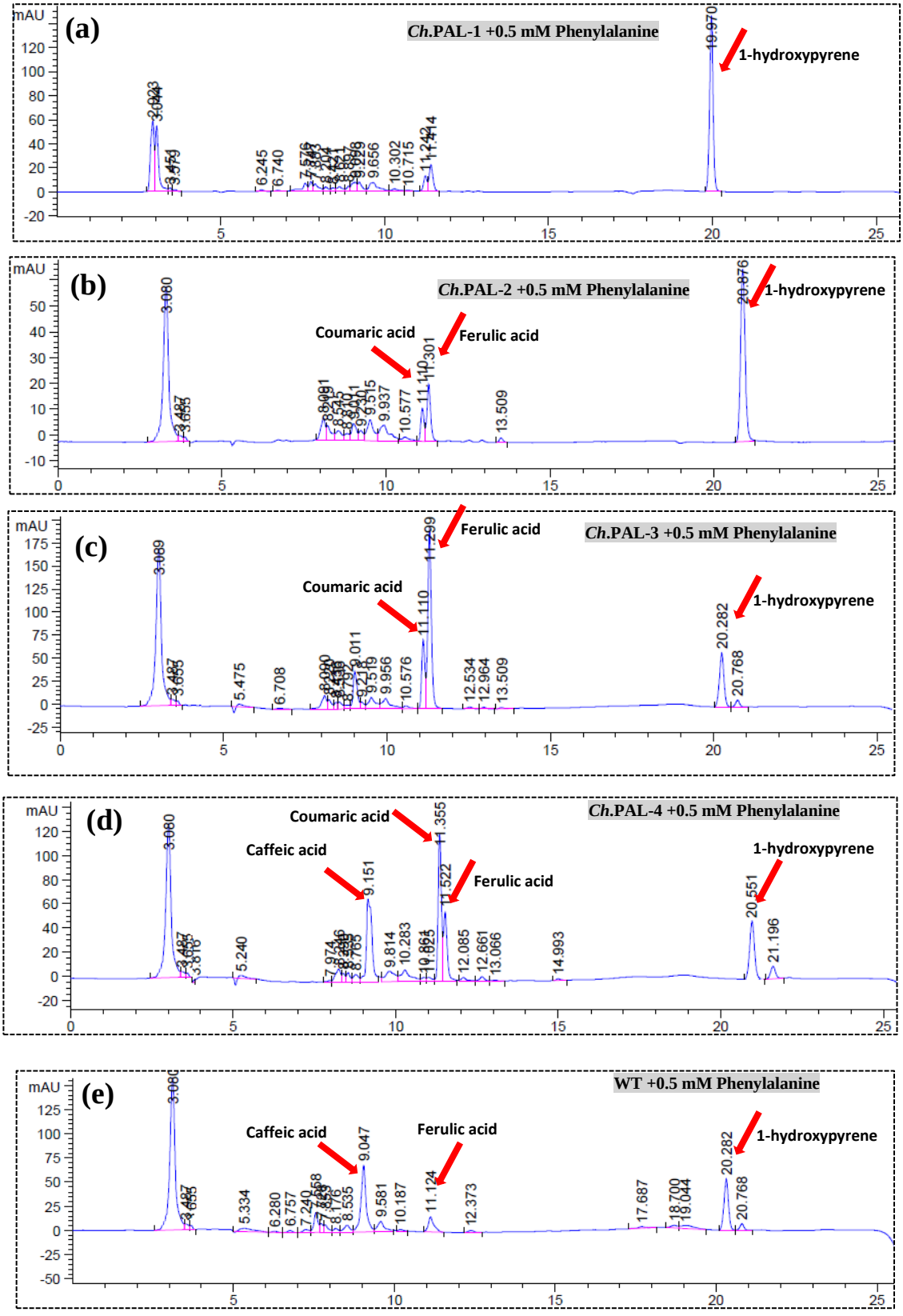


Figure 3.18. HPLC Chromatogram results of *Ch*.PAL-1 (a), *Ch*.PAL-2 (b), *Ch*.PAL-3(c), *Ch*.PAL-4 (d) overexpressors and wild type (e) after treatment with Phenylalanine and seven days incubation, Diode array detection (HPLC-DAD; 310 nm) and injection volume:10 µl

The increase in the production in transgenic tobacco cell lines carrying *Ch*.PALs in comparison to wild type is as a result of over expressing *Ch*.PALs under the control of the constitutive cauliflower mosaic virus (CaMV) 35S promoter which is constitutively on. Despite of this constitutive promoter which is present in all transgenic cell lines, the difference between different metabolite production in 4 different *Ch*.PAL overexpressors could be as result of distinct PAL isoform functions or their different affinities of PALs to Phenylalanine as substrate.

According to calculations based on the standard curve formulae on Figure 3.17, the quantity of Ferulic acid produced by *Ch*.PAL-3 overexpressor line after seven days of incubation was 120µg ml^{-1}, and the amount of Coumaric acid production was 40 µg ml^{-1}. When compared to *Ch*.PAL-3, Coumaric acid was produced in greater quantities in the overexpressor line *Ch*.PAL-4 in contrary to Ferulic acid which was produced less in this cell line.

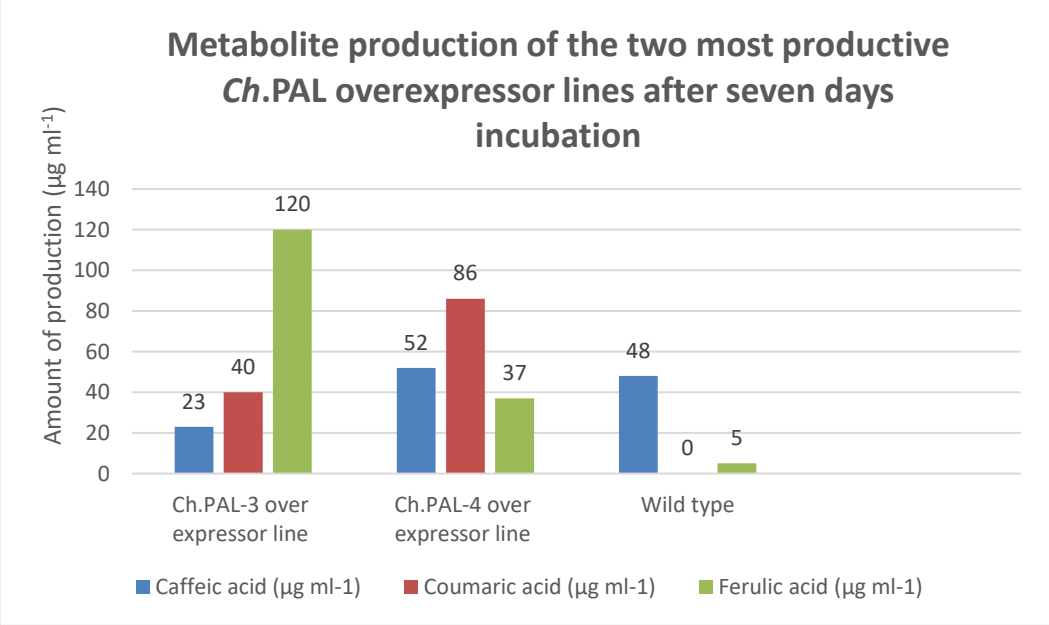


Figure 3.19. Quantitative comparison of metabolite production between the two most prolific overexpressor lines (*Ch*.PAL-3, *Ch*.PAL-4) and wild type after seven days incubation

3.2.1.2. Ferulic acid, the main product of PAL overexpressed cell lines after fourteen days incubation

After fourteen days incubation four different *Ch*.PAL overexpressor lines that have been fed with Phenylalanine showed significant amount of Ferulic acid production instead of Cinnamic acid

production, while the wild type tobacco BY2 cell line produced Caffeic acid and a bit more Ferulic acid in comparison to seven days old samples after Phenylalanine feeding.

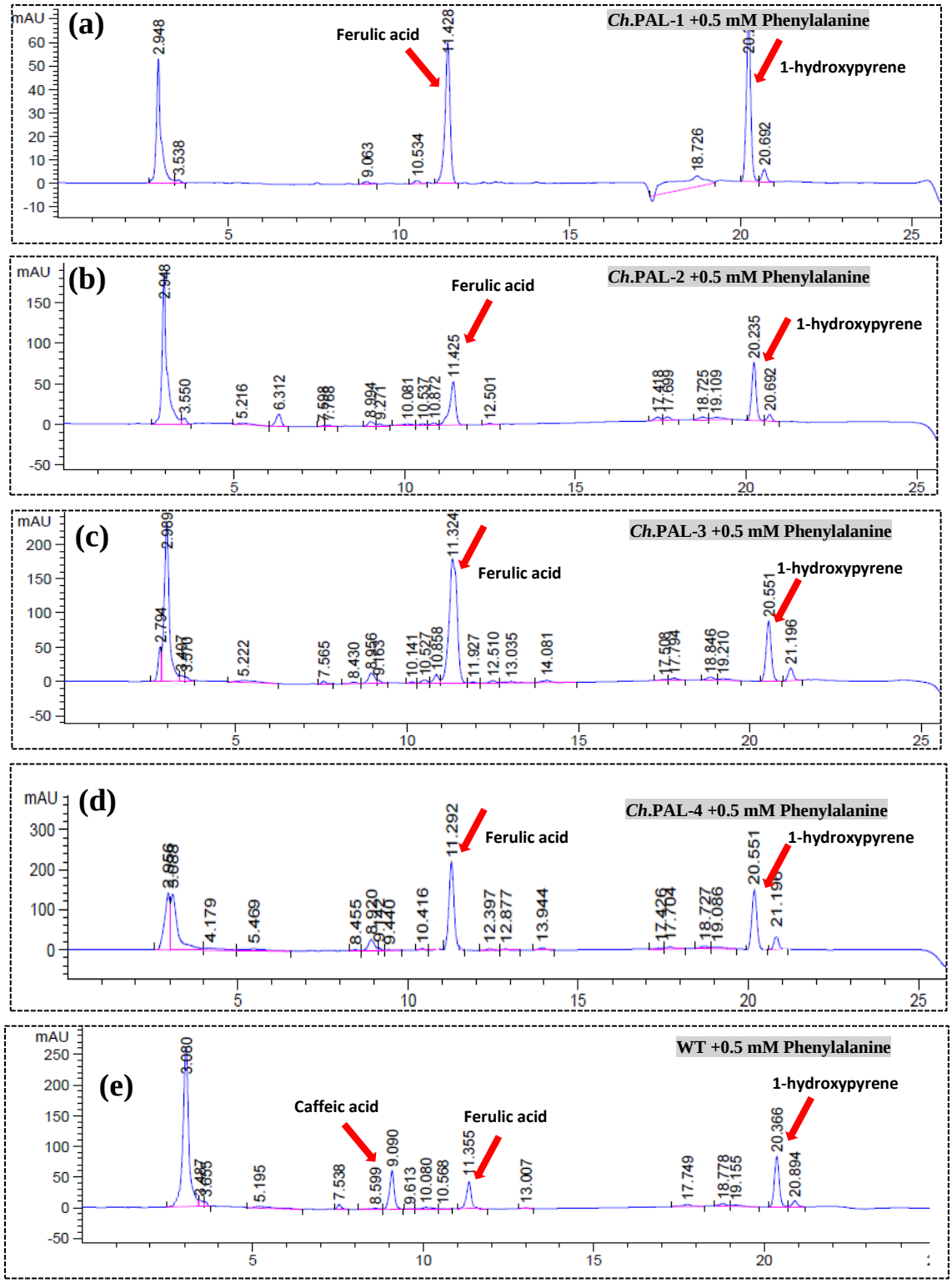


Figure 3.20. HPLC Chromatogram results of *Ch*.PAL-1 (a), *Ch*.PAL-2 (b), *Ch*.PAL-3(c), *Ch*.PAL-4 (d) overexpressors and wild type (e) after treatment with Phenylalanine and fourteen days incubation, Diode array detection (HPLC-DAD; 310 nm) and injection volume:10 µl

According to the HPLC data in figure 3.20, like seven days old samples, the two overexpressed cell lines *Ch*.PAL-3 and *Ch*.PAL-4, produced more metabolites than the other two cell lines. Based on the standard curve calculations in figure 3.17, the amount of Ferulic acid synthesis was determined in *Ch*.PAL-3 and *Ch*.PAL-4 overexpressor lines.

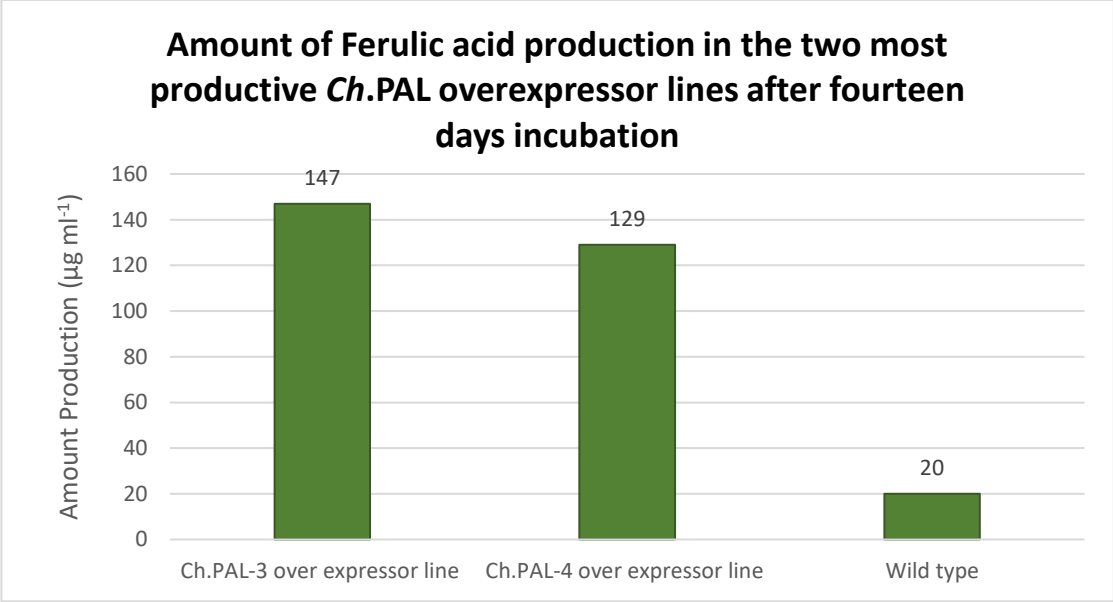


Figure 3.21. Quantitative comparison of Ferulic acid production between the two most productive overexpressor lines (*Ch*.PAL-3, *Ch*.PAL-4) and wild type after fourteen days incubation

Figure 3.21 depicts an overview of PAL transgenic cell lines' metabolite production. By increasing the Phenylalanine concentration in the cell cultures, HPLC analysis revealed no Cinnamic acid but Coumaric acid and Ferulic acid after seven days and a substantial amount of Ferulic acid after fourteen days, whereas wild type cells produced very little metabolite after being fed Phenylalanine. Endogenous Tobacco BY2 genes were certainly engaged in converting the predicted Cinnamic acid, which was the principal product of PAL enzyme, to Coumaric acid and Ferulic acid, as seen in this figure. Phenylalanine is converted to Ferulic acid after fourteen days by engaging three internal BY2 genes. C4H, 4-Coumarate 3-hydroxylase (C3H) which converts the Coumaric acid to Caffeic acid and Caffeic acid 3-O-methyltransferase (COMT) which is responsible for the production of Ferulic acid from Caffeic acid. This study's suggested possible explanation for the synthesis of Ferulic acid by PAL overexpression cell lines is that PAL overexpression raises the concentration of Cinnamic acid in cells, which may be toxic to cells at high concentrations. The reported negative feedback loop of Cinnamic acid by the expression of PAL at the entry point into the phenylpropanoid pathway through the activity of C4H (Blount et al., 2000) could also be due to the toxicity of Cinnamic acid

for cells. Because of this, the cells may choose to consume this substance by directing it toward the formation of Ferulic acid by raising the activity of nearby genes such C4H, which could account for the production of Coumaric acid in samples that have been collected after seven days. By raising the Coumaric acid concentration, other genes' activity, including C3H and COMT, would likely be stimulated, which will result in a significant amount of Ferulic acid being produced in fourteen days old samples.

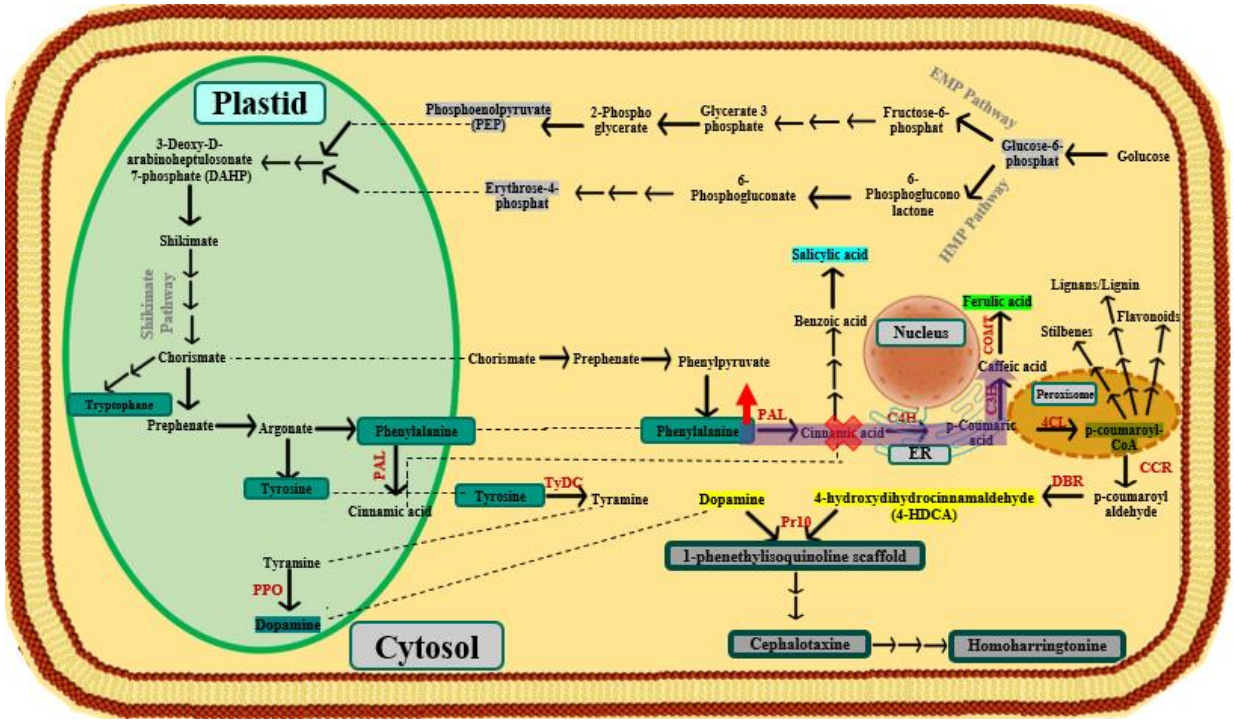


Figure 3.22. Schematic illustration of the enzymatic activity of *Ch*.PAL-3 overexpressor line for Ferulic acid production after fourteen days as indicated in purple color

3.2.2. Ferulic acid the main product of cocultivation of PAL and C4H over expressed cell lines

Co-cultures of PAL and C4H overexpressed cell lines were performed to determine whether Cinnamic acid, which is produced and converted to Ferulic acid by PAL overexpression, could be consumed by subsequent C4H overexpression lines. The goal of this experiment was to see if it was possible to redirect the pathway away from unwanted Ferulic acid production toward more Coumaric acid production. To accomplish this, the same amount of *Ch*.PAL-3 overexpressing cell lines was co-cultured with all the isoforms of *Ch*.C4H overexpressing cell lines before being fed 0.5 mM Phenylalanine and incubated for fourteen days. Contrary to expectations, HPLC data revealed

that Ferulic acid predominated in all co-cultivations, even at higher concentrations (184 µg ml^{-1}) than when PAL overexpressors were cultivated singly. (Figure 3.24).

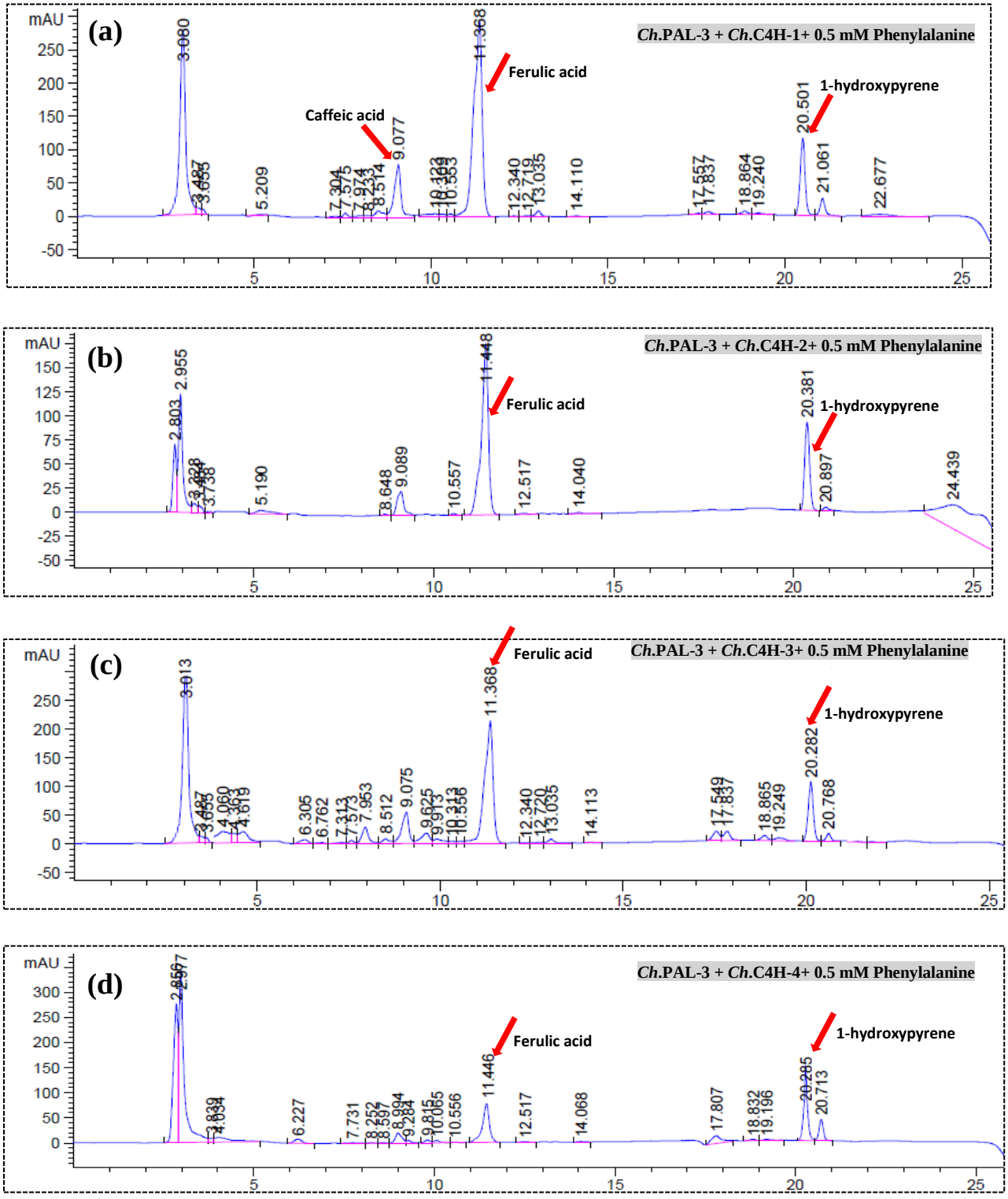


Figure 3.23. HPLC Chromatogram results of four different cocultivation of *Ch*.PAL-3 with *Ch*.C4H-1 (a), *Ch*.C4H-2 (b), *Ch*.C4H-3(c), *Ch*.C4H-4 (d) overexpressors after treatment with Phenylalanine and fourteen days incubation, Diode array detection (HPLC-DAD; 310 nm) and injection volume:10 µl.

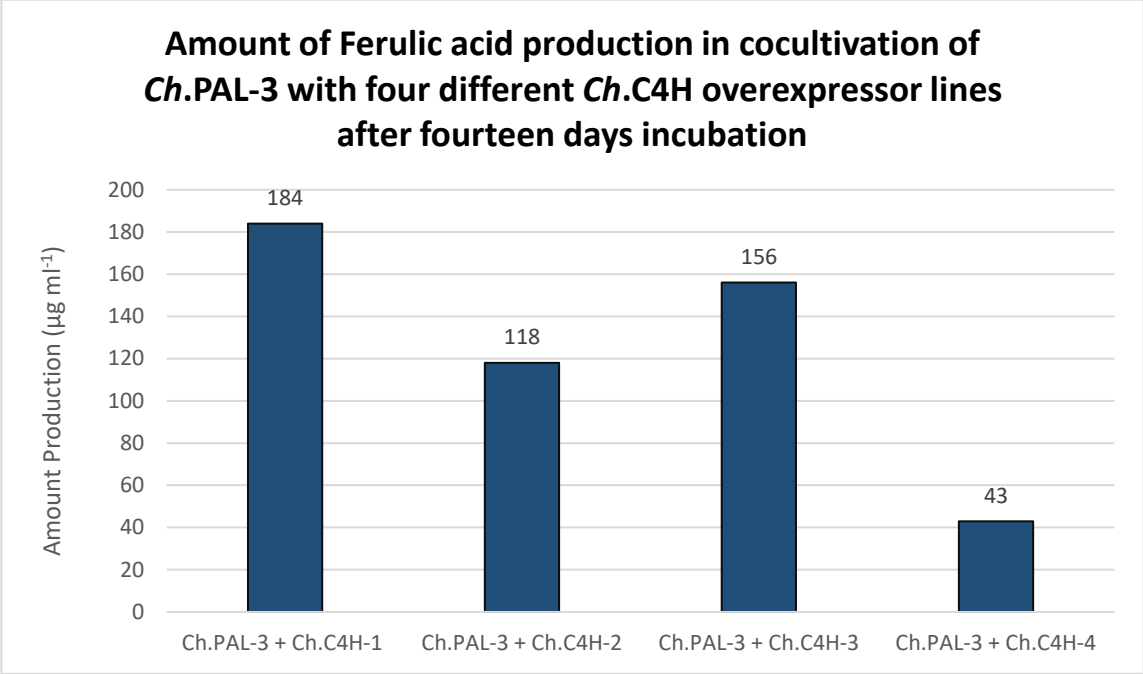


Figure 3.24. Quantitative comparison of Ferulic acid production between four different co-cultivation of *Ch*.PAL-3 with *Ch*.C4H-1, *Ch*.C4H-2, *Ch*.C4H-3, *Ch*.C4H-4 overexpressors after treatment with Phenylalanine and fourteen days incubation

These findings demonstrate that when *Ch*.PAL-3 and *Ch*.C4Hs are grown together, Ferulic acid synthesis is far more significant. However, the most fruitful co-cultivation was between *Ch*.PAL-3 and *Ch*.C4H-1, which is noteworthy because *Ch*.C4H-1 was the least productive *Ch*.C4H over expressor when grown alone after being supplied with 0.5 mM Cinnamic acid as a precursor and cultured for seven days. (Figure 3.25). On the contrary the cocultivation of *Ch*.PAL-3 with the *Ch*.C4H-4, which was the most prolific *Ch*. C4H over expressor of all, produced the lowest yield of Ferulic acid. Therefore, based on these findings, we can infer that the C4H cell lines have an inhibitory effect on the *Ch*. PAL overexpressor lines to a greater or lesser extent depending on how active they are. This means that the *Ch*.C4H over expressed cell lines' ability to suppress the *Ch*. PAL overexpressor lines increases with their activity. Figure 3.25 shows the high-performance chromatography (HPLC) analysis of the metabolic output of four distinct C4H overexpressor lines. To compare the production of various isoforms, the same fresh weight (2 gr) of tobacco BY2 cell cultures overexpressing *Ch*.C4H-GFP were inoculated in new media and the Cinnamic acid precursor was added to the final concentration of 0.5 mM and then cells were incubated for seven days. The injection volume for HPLC analysis was set to 10 µl. At 310 nm, all of the peaks were measured. The wavelength range of 200–700 nm was used to capture UV spectra.

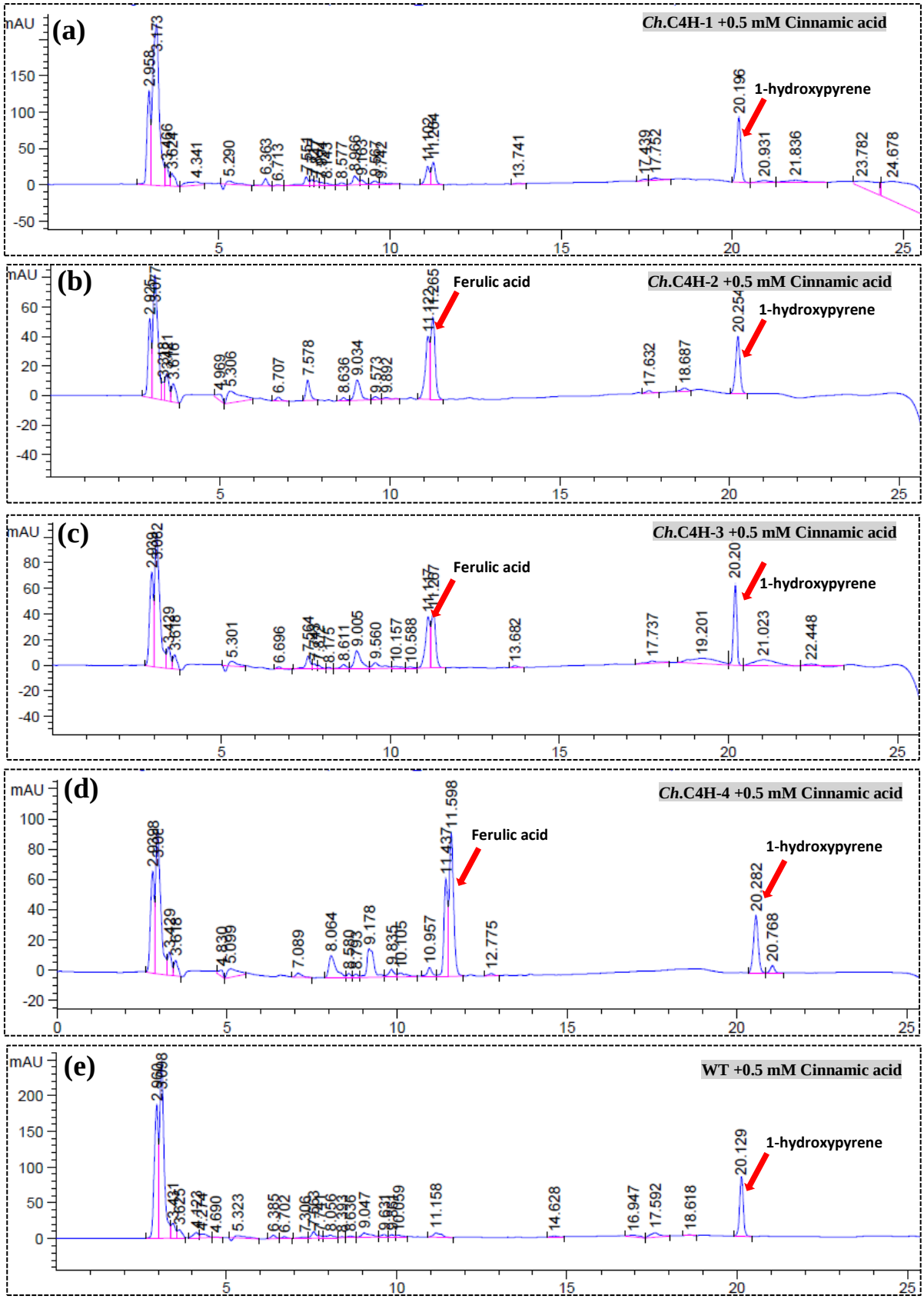


Figure 3.25.HPLC Chromatogram results of *Ch*.C4H-1 (a), *Ch*.C4H -2 (b), *Ch*.C4H -3(c), *Ch*.C4H -4 (d) overexpressors and wild type (e) after treatment with Cinnamic acid and fourteen days incubation, Diode array detection (HPLC-DAD; 310 nm) and injection volume:10 µl

3.2.3. More production of Ferulic acid by Sodium fluoride (NaF) treatment of PAL overexpressed cell lines after fourteen days incubation

The presence or absence of the precursor Phenylalanine had no discernible effect on the HPLC results of *Ch.* PAL over expressor lines. For example, as shown in figure 3.26, Ferulic acid production by *Ch.*PAL-3 overexpressed cell lines was almost the same in transgenics treated with Phenylalanine and in cell lines with no treatment after fourteen days incubation.

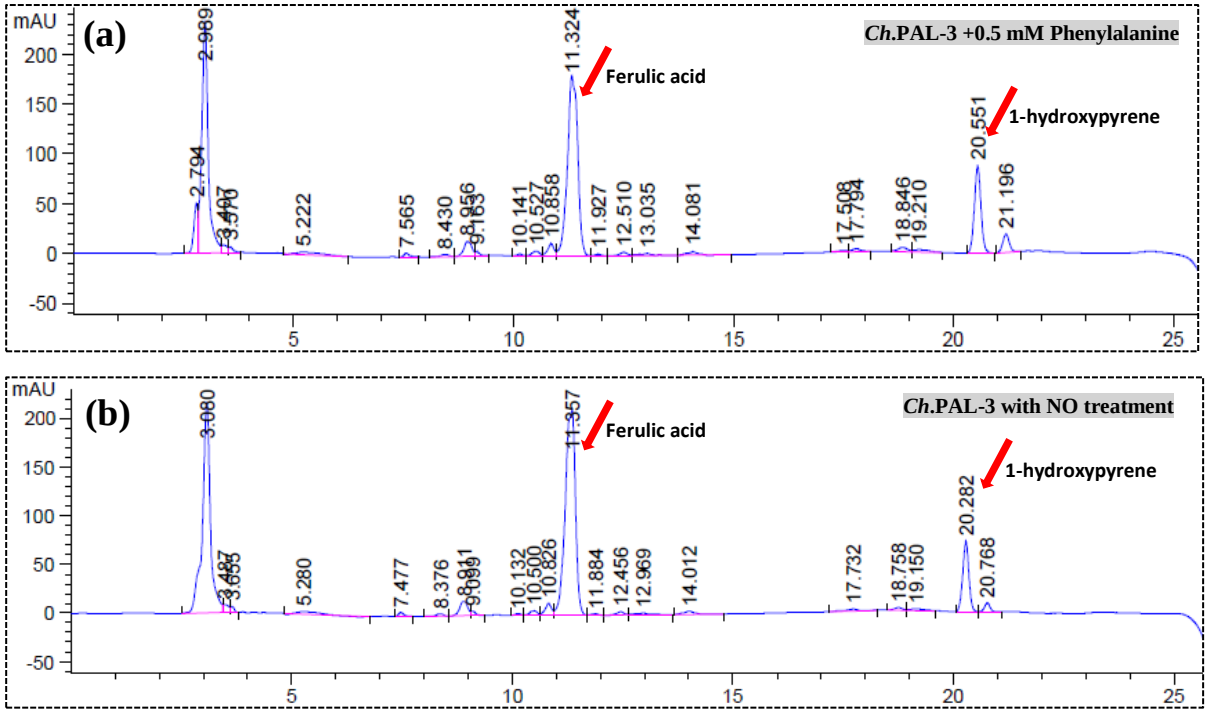


Figure 3.26. HPLC Chromatogram results of *Ch.*PAL-3 overexpressor after treatment with Phenylalnine (a) and without any treatment (b) and fourteen days incubation, Diode array detection (HPLC-DAD; 310 nm) and injection volume:10 µl

As previously stated in 3.2.1.2 figure 3.21, *Ch.*PAL-3 overexpressed cell lines produced 147 µg ml^{-1} Ferulic acid after fourteen days incubation in the presence of 0.5 mM phenylalanine, whereas the quantity produced with no treatment by this cell line was almost the same,136 µg ml^{-1}. That was the reason for investigating into the idea of inducing cells to make more phenylalanine on their own rather than adding it to the medium manually. There is evidence that sodium fluoride can serve as a metabolic inhibitor of the EMP pathway by inhibiting the activities of Aldolase and pyruvate kinase, directing the route away from the production of pyruvate towards the formation of phenylalanine (Y.-C. Li, 2014) as it was shown in figure 3.27 by red arrows.

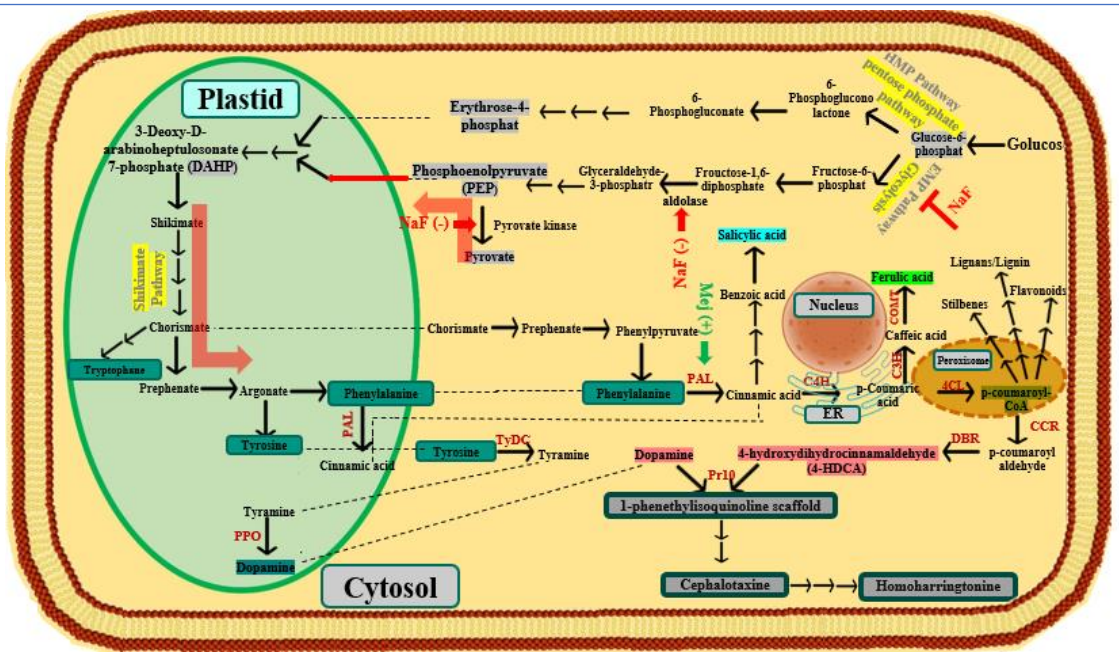


Figure 3.27. Schematic illustration of the enzymatic regulations in *Ch*.PAL-3 overexpressor line after treatment with Sodium Fluoride (NaF) and Methyl jasmonate (MeJA) as indicated in red and green colors

As a result, 0.2 mM sodium fluoride was applied to the wild type as well as *Ch*.PAL-3 over expressors. After fourteen days, HPLC results analysis revealed a considerable increase in Ferulic acid synthesis either in wild type and also in *Ch*.PAL-3 transgenic cell lines reaching up to 358 µg ml^{-1} (Figure 3.28). These findings showed that directing carbon flow towards the shikimate pathway, and subsequently toward increased production of the amino acid phenylalanine, might assist increase the production of the pathway, but the function of two endogenous BY2 C3H and COMT genes leads the pathway towards more production of Ferulic acid instead of P-Coumaroyl-CoA (the product of 4CL gene).

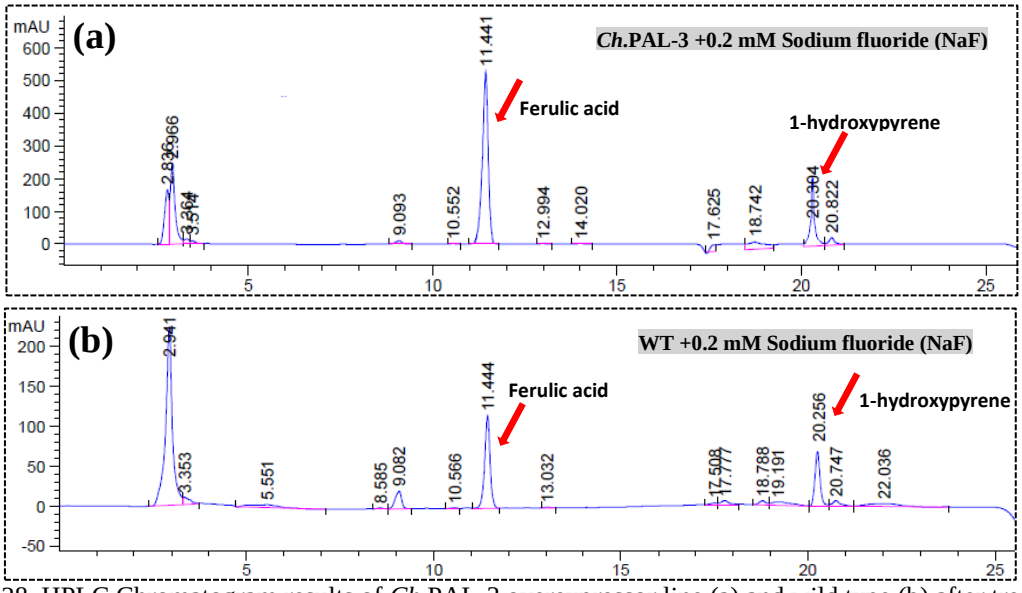


Figure 3.28. HPLC Chromatogram results of *Ch*.PAL-3 overexpressor line (a) and wild type (b) after treatment with Sodium Fluoride (NaF) and fourteen days incubation, Diode array detection (HPLC-DAD; 310 nm) and injection volume:10 µl

3.2.4. Increasing the production of Ferulic acid by Methyl jasmonate (MeJA) treatment of *Ch*. PAL overexpressed cell lines after fourteen days incubation

Methyl jasmonate (MeJA) treatment has been shown to be effective as an elicitor, activating the gene expression of related enzymes involved in the synthesis of plant secondary metabolites as PAL, which enhances the production of Cephalotaxin in suspension cultures of *Cephalotaxus mannii* (Y.-C. Li, 2014). To test the MeJA effect on the production of *Ch*.PAL-3 over expressors, this cell line was treated with 0.1 mM MeJA and then incubated for fourteen days. Figure 3.29 indicates an increase in the Ferulic acid production up to 238 µg ml^{-1}.

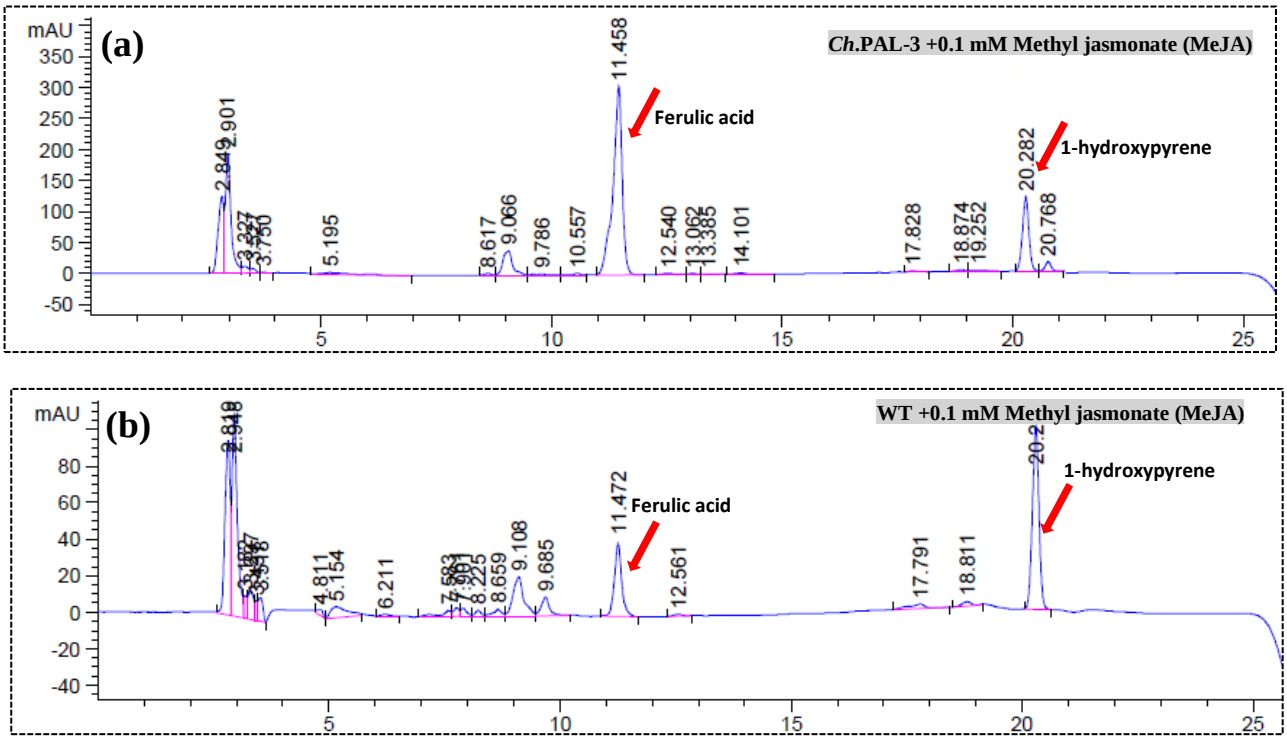


Figure 3.29. HPLC Chromatogram results of *Ch*.PAL-3 overexpressor line (a) and wild type (b) after treatment with Methyl jasmonate (MeJA) and fourteen days incubation, Diode array detection (HPLC-DAD; 310 nm) and injection volume:10 µl

These findings demonstrate the robust function of internal (BY2) C3H and COMT, which direct metabolic flow away from desired Coumaroyl-CoA towards Ferulic acid. Therefore, our knowledge of the sub-cellular localisation of enzymes in this route thus could be useful for redirecting the process towards the formation of P-Coumaroyl-CoA instead of Ferulic acid production.

3.2.5. Caffeic acid production as a consequence for inhibiting metabolic flux from ER

Our data regarding sub-cellular localisation indicate that the endoplasmic reticulum (ER) is the site of the C4H enzyme, which converts Cinnamic acid to Coumaric acid. It has been hypothesized in

this study, that Ferulic acid formation could be suppressed by preventing *p*-Coumaric acid from leaving the ER. To test this hypothesis, we used Brefeldin A, a fungal fatty acid which can inhibit the protein transport from the endoplasmic reticulum (ER) to the Golgi complex (Helms & Rothman, 1992). As a result, *Ch*.PAL-3 overexpressor lines were treated with 5 µM Brefeldin A and cultured for seven and fourteen days prior to HPLC measurement (Figure 3.30).

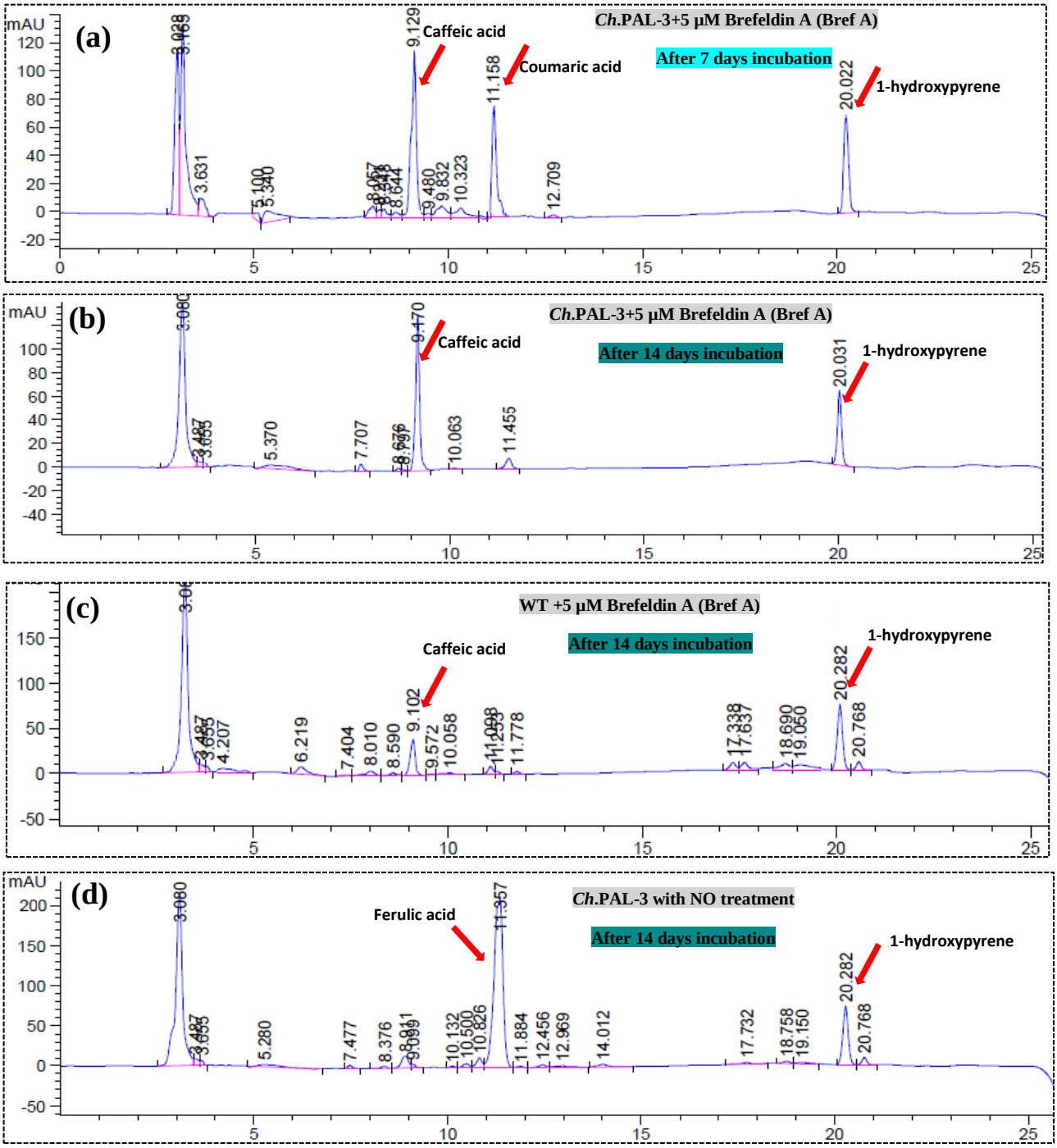


Figure 3.30. HPLC Chromatogram results of *Ch*.PAL-3 oevrexpressor line after treatment with Brefeldin A and seven days incubation (a) and fourteen days incubation (b), wild type after treatment with Brefeldin A (c) and *Ch*.PAL-3 with no treatment (d) and fourteen days incubation, Diode array detection (310 nm) and injection volume:10 µl

The HPLC findings after seven days incubation revealed no Ferulic acid but the formation of Caffeic acid and Coumaric acid (Figure 3.30.a), whereas the results after fourteen days incubation revealed only Caffeic acid production (Figure 3.30.b) with a final concentration of 87 μg ml^{-1}. This occurred when a significant level of Ferulic acid was identified in samples with no treatment after Fourteen days (Figure 3.30.d).

As a consequence, we may state that the HPLC findings validated our theory.

3.3. Expression analysis of PAL before and after treatment with precursors

According to pharmacological data trans-Cinnamic acid may operate as a feedback regulator of the expression and enzymatic activity of L-phenylalanine ammonia-lyase, the first enzyme in the phenylpropanoid pathway, (PAL). It has been proposed that trans-Cinnamic acid, a byproduct of the first enzyme in the phenylpropanoid pathway, Phenylalanine ammonia-lyase (PAL), may act as a signal molecule to modulate the flow into the route. Inhibiting PAL activity and gene transcription, Cinnamic acid given exogenously also causes the creation of a proteolytic enzyme inhibitor of PAL (Blount et al., 2000). To test the reported negative effect of Cinnamic acid on the expression of PAL genes, the qRT-PCR method was utilized using Ribosomal protein L25 and Elongation Factor (EF) primers as reference genes to evaluate the relative expression level of *Ch*.PAL overexpressors in the presence or absence of phenylalanine and Cinnamic acid (with the final concentration of 0.5mM) as the substrate and product of this gene.

For this purpose, two sets of PAL overexpressors lines were cultivated, one with phenylalanine/Cinnamic acid treatments and one without any treatment, then total RNA from transformants were extracted after seven days incubation. The q-PCR experiment has been done with three biological replicates followed by three technical replicates. The relative gene expression levels of exogenous *Cephalotaxus hainanensis* PAL and endogenous *Nicotiana tabacum* PAL were measured using the $2^{-\Delta\Delta Ct}$ technique.

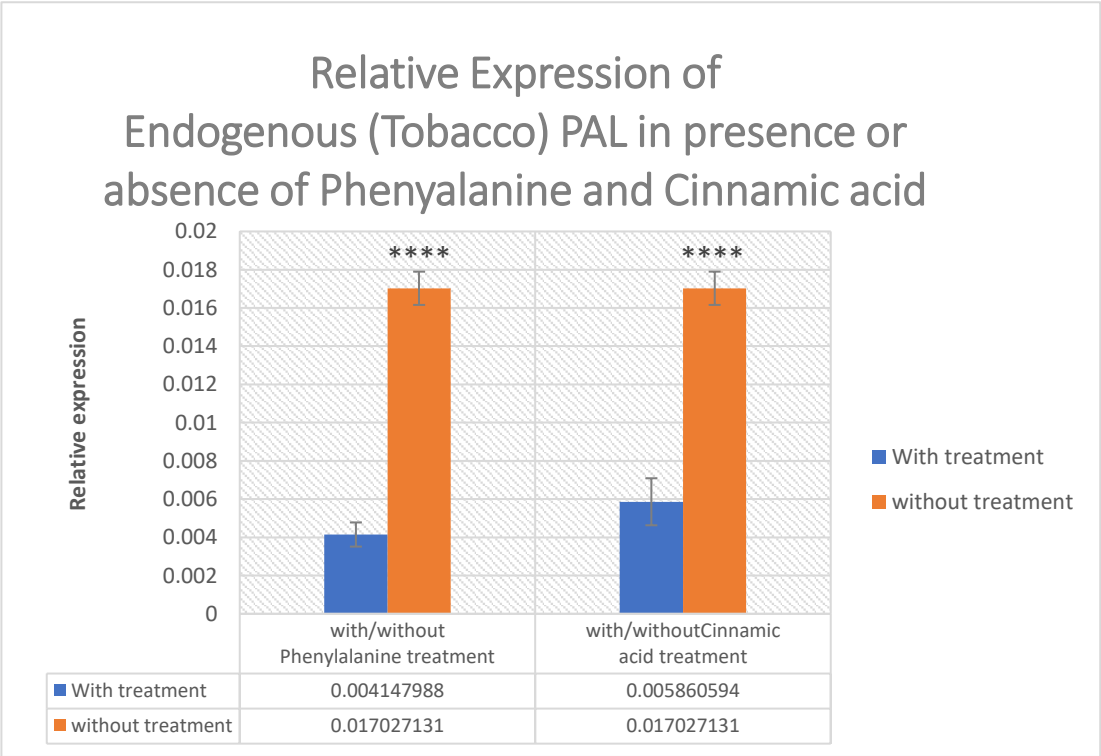


Figure 3.31. Relative gene expression analysis of endogenous (tobacco) PAL in *Ch*.PAL-3 overexpressor lines with or without Phenylalanine/Cinnamic acid treatment, Data represent the average of three biological replicates with three technical replicates in each experiment. Error bars show the standard error value. The asterisk indicates statistically significant differences between the control and cold treatment (Student's t-test; ****P ≤0.0001).

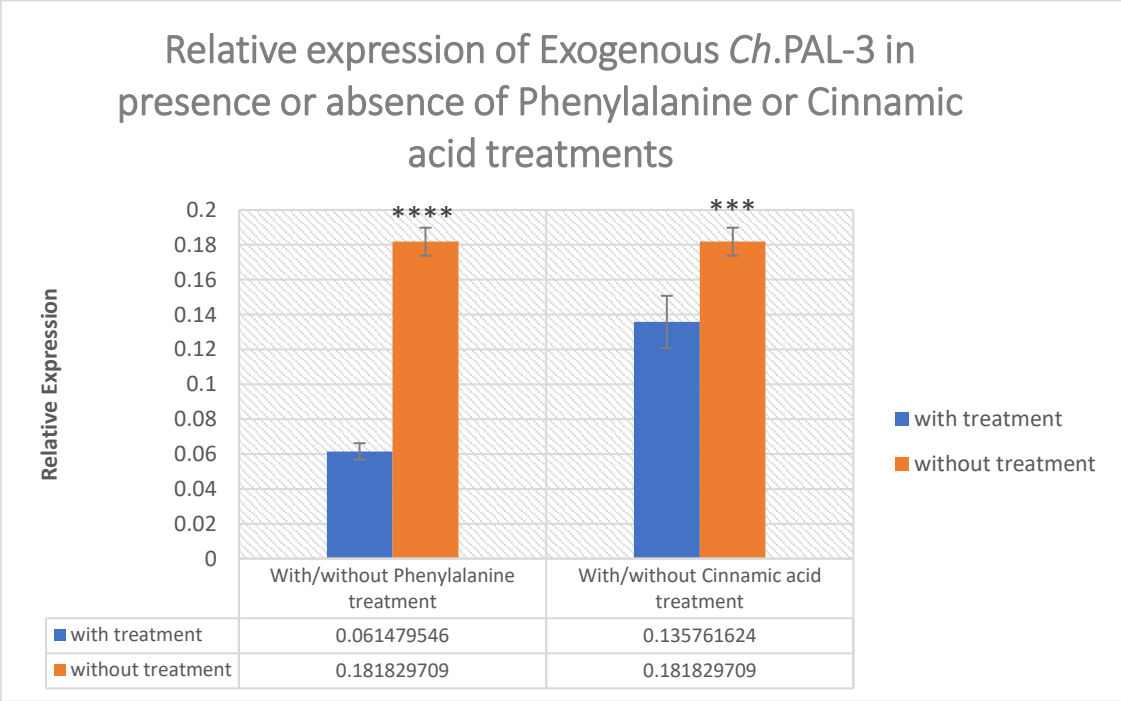


Figure 3.32. Relative gene expression analysis of exogenous (*Cephalotaxus hainanensis*) PAL in *Ch*.PAL-3 overexpressor lines with or without Phenylalanine/Cinnamic acid treatment, Data represent the average of three biological replicates with three technical replicates in each experiment. Error bars show the standard error value. The asterisk indicates statistically significant differences between the control and cold treatment (Student's t-test; ****P ≤0.0001 and ***P ≤ 0.00).

The results of the q-PCR experiment demonstrated a significant reduction in the relative gene expression of both exogenous *Ch*.PAL-3 and endogenous *Nicotiana tabacum* PAL in *Ch*.PAL-3 over expressed cell lines following treatment with either Phenylalanine or Cinnamic acid. These results were consistent with the hypothesis regarding the inhibitory effect of Cinnamic acid on the activity and expression of PAL which makes the downregulation of the PAL following treatment with Cinnamic acid, conceivable. Phenylalanine treatment also affects PAL genes to be downregulated, which may be due to cell-level buildup of Cinnamic acid. In other words, a higher Phenylalanine concentration would lead to a higher buildup of Cinnamic acid, which would then cause the downregulation of these genes. The intriguing aspect of this scenario is that the downregulation of exogenous *Ch*.PALs which are being regulated under the control of constitutive CaMV 35S promoter which is always on. Therefore, any changes in gene expression of exogenous PALs in presence of Phenylalanine or Cinnamic acid would be considered as a post transcriptional regulation.

3.4. Summary of results

Cephalotaxine alkaloids were discovered to have inhibitory activity against a variety of human tumor cell lines. Homoharringtonine, an esterified form of Cephalotaxine that was approved for the treatment of chronic myeloid leukemia in 2012, is an extremely expensive pharmaceutical compound due to a severe scarcity of its raw material, the plant genus *Cephalotaxus* (Cephalotaxaceae), which is the only source of this type of natural product. *Cephalotaxus hainanensis*, an endangered species of the *Cephalotaxus* genus, is the primary source of Homoharringtonine. Because of its extensive spectrum of herbal therapeutic capabilities, this exceedingly slow growing tree, which is so popular in traditional Chinese medicine, is on the verge of extinction as a result of unlawful over-exploitation. The restricted supply of Homoharringtonine, along with the rising demand for this medicine, necessitates alternate methods of its manufacture. The initial effort attempted to produce Homoharringtonine semi-synthetically by esterifying Cephalotaxine, was commercially named Omacetaxine mepesuccinatewhich was isolated from dried *Cephalotaxus* leaves. However, this procedure still requires a significant quantity of rare *Cephalotaxus* biomass. Plant cell fermentation is another option, although this process still resulted in a poor production rate in the suspension culture of *Cephalotaxus* leaves. Another option is metabolic engineering and heterologous production of this natural substance in a suitable biological

host. As a result, a comprehensive understanding of the Cephalotaxine metabolic production pathway is necessary.

Therefore, we focused our research on the sub-cellular localisation of enzymes involved in the formation of 1-Phenylethylisoquinoline scaffold (the general and unique precursor for Cephalotaxine alkaloids) to broaden our knowledge and improve our ability to logically design the metabolic engineering of Cephalotaxine alkaloids for future applications. In order to accomplish this, 21 functional genes (four PALs, four C4Hs, two 4CLs, five CCRs, three DBRs, one TyDC, one PPO, and one PR10) in this pathway were isolated from *Cephalotaxus hainanensis* cDNA and then were labelled with GFP by inserting them in to an expression vector pK7FWG2.0 which was later transformed in to *Nicotiana tabacum*. According to a sub-cellular localisation analysis of the functional genes in the 1-Phenylethylisoquinoline scaffold manufacturing route that were discovered by confocal microscopy in current research, this pathway involves a variety of cellular compartments. The first three enzymes, PAL, C4H, and 4CL, which are also the three enzymes that are shared in the general phenylpropanoid pathway and are resulted in production of p-Coumaroyl CoA, have the most transit in this route. These enzymes were found to be located in three different cellular compartments: the cytosol, endoplasmic reticulum, and peroxisomes. This raises the question of whether or not cells implement a strategy to transport the products of these three separate enzymes more effectively between various cellular compartments in the highly dynamic phenylpropanoid pathway which in discussion chapter we will go into greater detail about this query.

Another possible plastidic pathway for the production of p-Coumaroyl-CoA was also identified and confirmed by colocalisation study of the *Ch*.PAL-1 and *Ch*.4CL-1 isoforms with tpFNR-mEOS which is a plastid marker. The final stage of the production process for 4-HDCA, the scaffold for 1-Phenylethylisoquinoline that results from the Phenylalanine branch, takes place in the cytosol. Here, one may inquire about the 4CL enzyme's substrate, p-Coumaric acid presence in plastids despite of presence of C4H (the enzyme producing Coumaric acid) in Endoplasmic reticulum. How does this substance enter plastids? Or is there a mechanism in plastids that leads to Coumaric acid? These queries will be addressed in the discussion section.

The Tyrosine branch, which culminates in the final delivery of Dopamine, comprises a zigzag path linking two separate compartments: the cytosol and the plastids, according to the sub-cellular localization findings in this study. In plants, however, this zigzag approach is not so efficient in

terms of energy consumption. Therefore, it is possible to hypothesize that a plastidic TyDC exists, which greatly facilitates the pathway for the creation of Dopamine.

Metabolic activity of the PAL overexpressed cell lines was assessed in vivo utilizing high performance chromatography (HPLC). This was accomplished by feeding 0.5mM Phenylalanine to the same quantity of tobacco BY2 cell cultures' fresh weight, overexpressing *Ch*.PAL-GFP which were then harvested seven and fourteen days after inoculation. The HPLC results of four different *Ch*. PAL overexpressor lines revealed no Cinnamic acid production, but rather up to 120 μg ml^{-1} Ferulic acid and 40 μg ml^{-1} Coumaric acid production after seven days of incubation while following fourteen days of incubation, a substantial amount of Ferulic acid production was detected in cells, reaching a final concentration of 147 μg ml^{-1} whereas wild type cells produced very little metabolite in comparison to the overexpressor lines. In the process of converting Phenylalanine to Coumaric acid and Ferulic acid, endogenous tobacco BY2 genes were undoubtedly activated. This means that after fourteen days, Phenylalanine is transformed to Ferulic acid by activating three internal BY2 genes. C4H to turns the expected Cinnamic acid to Coumaric acid, 4-Coumarate 3-hydroxylase (C3H) that converts Coumaric acid into Caffeic acid, and Caffeic Acid 3-O-Methyltransferase (COMT) that turns Caffeic acid into Ferulic acid. The notion put out in this study is that PAL overexpression increases the concentration of Cinnamic acid in cells, which could be harmful to cells at high quantities and may account for the synthesis of Ferulic acid by PAL overexpression cell lines. The toxicity of Cinnamic acid for cells could also be the explanation of the reported negative feedback loop of Cinnamic acid caused by the expression of PAL at the entrance point into the phenylpropanoid pathway by the function of C4H (Blount et al., 2000). Due to this, the cells may decide to consume the Cinnamic acid by directing it toward the development of Ferulic acid via increasing the activity of neighboring genes such C4H, which could explain the production of Coumaric acid in samples that were collected after seven days. The activity of other genes, such as C3H and COMT, would probably be boosted by increasing the quantity of Coumaric acid, leading to a considerable amount of Ferulic acid being synthetized in samples that are fourteen days old.

In order to determine whether the subsequent C4H overexpression lines could consume the Cinnamic acid that is produced and converted to Ferulic acid by PAL overexpression, PAL and C4H overexpressed cell lines were co-cultured. The purpose of this experiment was to investigate the possibility to switch the pathway, so that more Coumaric acid would be produced instead of undesired Ferulic acid. To accomplish this, equal amounts of *Ch*.PAL-3 and *Ch*.C4H overexpressing cell lines were provided with Phenylalanine and co-cultured for fourteen days. Contrary to what was

expected, HPLC data showed that Ferulic acid was the most prevalent metabolite in all co-cultivations, even at higher concentrations (184 μg ml^{-1}) than when PAL overexpressors were grown individually. These results show that Ferulic acid synthesis is even more significant when *Ch*.PAL-3 and *Ch*.C4Hs are cultivated together than when *Ch*.PAL-3 is grown separately. The most fruitful co-culture, however, was between *Ch*.PAL-3 and *Ch*.C4H-1, which is remarkable because the Ch.C4H-1 was the least fruitful *Ch*.C4H over expressor when grown alone after being provided with 0.5 mM Cinnamic acid as a precursor and cultivated for seven days. In contrast, cocultivation of *Ch*.PAL-3 with *Ch*.C4H-4, which was the most prolific Ch. C4H over expressor of all, had the lowest Ferulic acid output. Based on these observations, we may conclude that the C4H cell lines have an inhibitory impact on the *Ch*. PAL overexpressor lines to varying degrees, depending on how active they are. This means that the capacity of the *Ch*.C4H overexpressed cell lines to repress the *Ch*. PAL overexpressor cells rises with their activity.

The presence or absence of the precursor Phenylalanine had no noticeable influence on *Ch*. PAL over expressor lines based on the HPLC findings. This result could be explained by the results of gene expression analysis in 3.3. Despite the fact that *Ch*.PAL was downregulated as a result of phenylalanine treatment, the metabolic outcome in treated and non-treated samples remained identical after fourteen days and the explanation was simply because there was still a greater amount of Phenylalanine that could be converted to Ferulic acid. That was the impetus for looking into the notion of getting cells to produce more Phenylalanine on their own rather than adding it to the medium manually. There is evidence that Sodium fluoride (NaF) can operate as a metabolic inhibitor of the EMP pathway by decreasing the activities of Aldolase and pyruvate kinase, diverting the pathway away from pyruvate production and toward phenylalanine accumulation (Y.-C. Li, 2014). That was the rationale behind the idea of treating *Ch*.PAL-3 over expressor line with 0.2 mM sodium fluoride (NaF). Ferulic acid production in treated cell lines increased noticeably over the course of fourteen days, reaching up to 358 μg ml^{-1}, according to HPLC analysis. These results demonstrated that directing carbon flow towards the shikimate pathway, and consequently toward production of phenylalanine, might assist in increasing the production of whole pathway. However, in *Nicotiana tabacum* the function of two endogenous C3H and COMT genes leads the pathway towards more production of Ferulic acid instead of P-Coumaroyl-CoA. Since Methyl jasmonate (MeJA) has been reported to be effective as an elicitor, activating the gene expression of related enzymes involved in the synthesis of plant secondary metabolites like PAL, which increases the production of Cephalotaxin in in suspension cultures of *Cephalotaxus mannii* (Y.-C. Li, 2014), treatment of

*Ch.*PAL-3 over expressor line with MeJA also revealed an increase in Ferulic acid production up to 238 μg ml^{-1}. These findings show how internal (BY2) C3H and COMT, which divert metabolic flow away from desirable path and toward Ferulic acid, operate with sturdiness. Our understanding of the sub-cellular localisation of the enzymes involved in this pathway thus come to handy in rerouting the process away from Ferulic acid. Our findings on sub-cellular localization indicate that the C4H enzyme, which convert Cinnamic acid into Coumaric acid, is located in the Endoplasmic reticulum (ER). Therefore, we postulated that we might be able to divert the pathway and reduce the production of Ferulic acid by preventing p-Coumaric acid from leaving the ER. To test this hypothesis, we used brefeldin A, a fungus-derived fatty acid that could stop protein transport from the endoplasmic reticulum (ER) to the Golgi complex (Helms & Rothman, 1992). HPLC results confirmed the accuracy of our hypothesis.

The HPLC results of *Ch.*PAL-3 overexpressor lines treated with 5 μM Brefeldin A after seven days of incubation showed NO synthesis of Ferulic acid but did reveal the development of Caffeic acid and Coumaric acids (Figure 3.30.a), whereas the HPLC results after fourteen days demonstrated only the production of Caffeic acid to the final concentration of 87 μg ml^{-1}. This happened when samples that had not received any treatment after fourteen days have shown a considerable amount of Ferulic acid (Figure 3.30.d).

The Brefeldin A experiment's capability to alter metabolism by manipulating intracellular structure successfully supported our initial theory that, if compartmentalization is a mechanism to guide the pathway in different directions, it should be feasible to change metabolism as a result.

4. Discussions

1-Phenethylisoquinoline scaffold is the unique substrate of PSS (the enzyme responsible for integrating the two pathways beginning with Phenylalanine and Tyrosine) which is exclusively found in *Cephalotaxus* specious. This compound's synthesis process shares numerous beginning stages with the overall phenylpropanoid pathway. Numerous studies have been undertaken over the last several decades on the regulatory evaluations of the phenylpropanoid pathway, indicating that several complex regulatory mechanisms particularly on the first three enzyme in this pathway, at multiple levels control the phenylpropanoid metabolic process. Although knowledge of these regulatory processes advances fundamental understanding of plant specialized metabolism, we still lack a comprehensive grasp of sub-cellular analyses of the enzymes in this intricate, highly dynamic biosynthetic machinery. Increasing the knowledge about these sub-cellular processes undoubtedly enhances our capacity to logically design plant metabolic pathways for future applications or the metabolic engineering of 1-phenethylisoquinoline scaffold as we believe that compartmentation and distribution are two inseparable essential elements for the molecules function and their efficient biosynthesis.

4.1. Biosynthesis of phenylpropanoids, a dynamic complexity

The stunning diversity of phenylpropanoids emerges from effective modification and amplification of a relatively small number of core structures originating from the shikimate pathway (Herrmann, 1995). Aromatic amino acid lyases connect the Phenylalanine and Tyrosine pool to biosynthetic pathways, which are occasionally demoted as secondary metabolism but which end up being just as important to plant life as photosynthesis or the citrate cycle. A group of enzymes, often arranged in superfamilies, such as oxygenases, ligases, oxidoreductases, as well as different superfamilies of transferases, accomplish the diversity and flexibility of the ensuing natural products, more especially the phenylpropanoid profile. Even though some of these enzymes could show overlap in specificities when tested in vitro, their developmentally and spatially regulated expression particularly leads to chemical phenotypes that are unique to particular tissues and plants. Other than Lignin and Flavonoid biosynthesis, metabolites generated from phenylpropanoid make up and contribute to several different biosynthetic branches. The development of analytical, chemical, and biomolecular methods, such as whole genome tilling arrays and the use of fluorescence biosensors, has led to several novel theories and insights on the intricate structure of pathways (Herbst et al., 2009). These

methods will serve as a crucial impetus for understanding plant biology (Yonekura-Sakakibara et al., 2008; Zeller et al., 2009). In order to get a complete understanding of the plant organism, they facilitate the quick annotation of genes, the structural and functional characterization of enzymes, the measurement of quantities of metabolites, and the discovery of signal transduction pathways in live cells. As a result, we narrowed down our research on the cellular localisation of the enzymes in the 1-phenethylisoquinoline scaffold production route.

The central phenylpropanoid pathway's genes are regularly found in multiple copies. For instance, the PAL genes have four isoforms in *Arabidopsis thaliana*, six isoforms in rice (*Oryza sativa*), nine in poplar (*Populus trichocarpa*), and five in Medicago (*Medicago truncatula*). Different isoforms have different localization and activity patterns. The *Arabidopsis* PAL1, PAL2, and PAL4 genes, for instance, are expressed at relatively high levels in stems throughout the latter stages of development, with PAL1 expression localized in the vascular tissue and PAL2 and PAL4 are both expressed in seeds. Olsen et al. (2008) demonstrated the involvement of PAL1 and PAL2 in Flavonoid biosynthesis, whereas the two additional proteins were hypothesized to be involved in Lignin production (Huang et al., 2010). 4CL has properties similar to PAL in *Arabidopsis*. 4-CL is encoded by four paralogs in *Arabidopsis* (Hamberger and Hahlbrock, 2004; Costa et al., 2005). According to Costa et al. (2005) and Ehlting (1999), 4-CL1 was suggested as the best option to drive the flow of metabolites into monolignols due to its robust expression in stem and root and high catalytic efficiency for conversion of p-coumaric acid into its CoA ester. However, the involvement of additional paralogs, particularly 4-CL2, cannot be completely ruled out. A wide variety of cell types express the 4CL3, which is mostly linked to Flavonoid production. Co-expression with Lignin biosynthesis genes was found for 4CL1, 4CL2, and 4CL4 (Ehlting et al., 1999; Li et al., 2015). The product of 4CL activity, p-Coumaroyl-CoA, is a key step in the phenylpropanoid pathway. It is a precursor for Monolignol, Coumarin, Stilbene, as well as (Iso)flavonoid biosynthesis (Kutchan et al., 2015). When exposed to various stimuli, some isoforms are thought to be variably regulated, which diverts flow to the specific metabolite production. Therefore, in this study we hypothesized that various 4CL isoforms may be implicated in various metabolic fluxes that occur in different cellular compartments which undoubtedly makes the process of channelizing metabolites from p-Coumaroyl-CoA toward various secondary metabolite synthesis more efficient.

4.2. Proposing a second biosynthetic pathway for production of p-Coumaroyl-CoA in plastids

The sub-cellular localisation results of this study (Figure 4.1) propose a second metabolic pathway for production of the significant molecule, p-Coumaroyl-CoA, in plastids by discovering of a second isoform of *Ch*.4CL in plastids, which was supported by colocalisation studies using the transient transformation of the plastid marker tpFNRmEos into the *Ch*.4CL-1 overexpressor line (Figure 3.8). Revealing the second distinct cellular compartments for this enzyme, which is the final enzyme in the general phenylpropanoid pathway's three shared steps, contributing in channelizing flux towards different phenylpropanoid biosynthetic pathways, could be complemented by the discovery the presence of one *Ch*.PAL isoforms, *Ch*.PAL-1, in plastids which was also confirmed by the colocalisation study with the plastid marker tpFNRmEos (Figure 3.4).

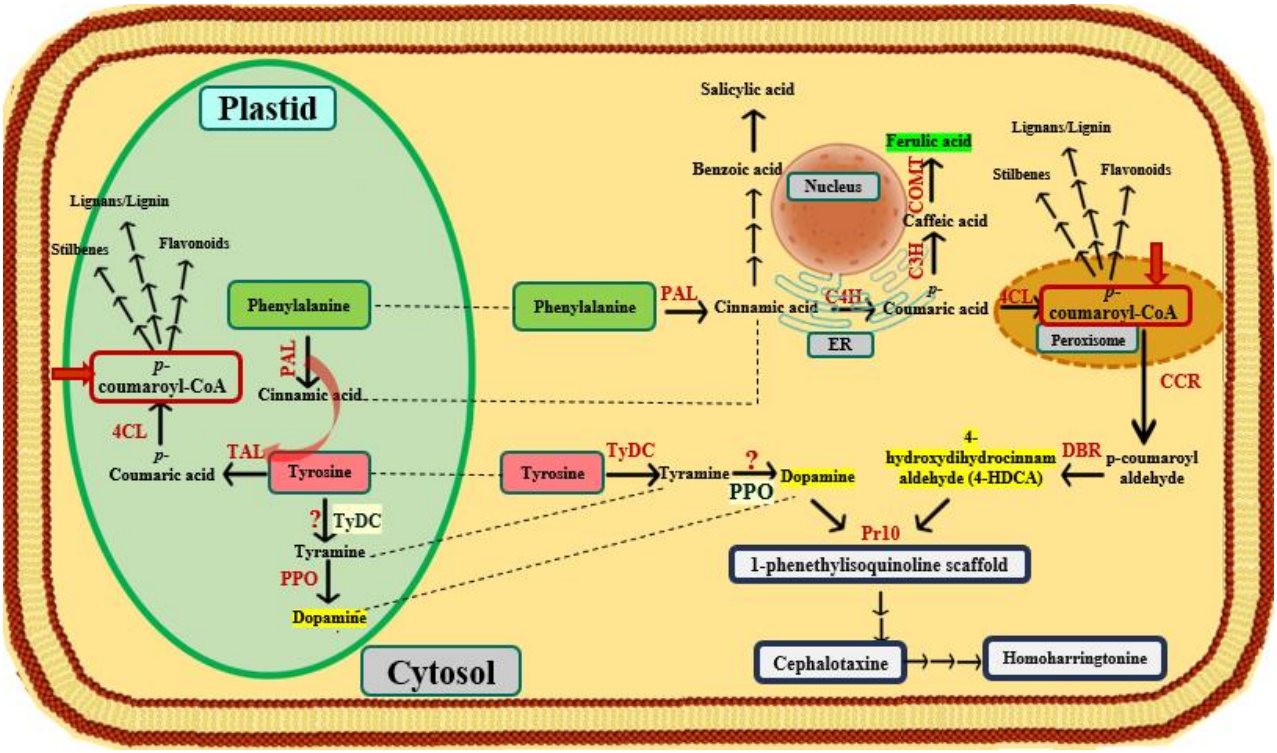


Figure 4.1. Schematic illustration of two distinct pathways producing p-Coumaroyl-CoA as indicated with red arrows

These results provide evidence for considering the plastidic PAL as a Tyrosine ammonia lyase (TAL) which could convert the Phenylalanine to Coumaric acid. The reason for considering PAL as a TAL is that both enzymes are members of the aromatic amino acid lyase family which could accept each other's substrates. Also, it has been reported that in monocotylic plants their activities reside

on the same polypeptide with similar catalytic efficiencies (Rosler et al., 1997). Therefore, by considering the plastidic PAL as TAL, p-Coumaric acid, the product of TAL could perform as the precursor for the plastidic 4CL to convert it to P-Coumaroyl CoA. This model undoubtedly makes the process of channelizing metabolites from p-Coumaroyl-CoA toward various secondary metabolite synthesis more efficient. These sub-cellular localization results therefore supported our theory suggesting p-Coumaroyl-CoA can be produced via two distinct pathways, one happening entirely in plastids and one occurring between cytosol, ER and Peroxisomes, increasing the efficiency of the process by which metabolites from this compound are directed into different secondary metabolite synthesis pathways.

4.3. Proposing a second biosynthetic pathway for production of Dopamine in plastids

According to the sub-cellular localization results in this study (Figure 4.1), the Tyrosine branch, which results in the ultimate delivery of Dopamine, consists of a zigzag route connecting two different compartments: the cytosol and the plastids. However, this zigzag strategy is not as energy-efficient when used to plants. The existence of a plastidic TyDC, which substantially improves the pathway for the production of Dopamine, can therefore be hypothesized. Unfortunately, our analysis was unable to validate the potential existence of another isoform of Tyrosine decarboxylase (TyDC) in plastids since it was difficult to extract all of the known isoforms of TyDC from *Cephalotaxus hainanensis* cDNA. However, the presence of such plastidic TyDC would allow cells to save energy by avoiding unnecessary transports across different cellular compartments. Another possibility suggested in this approach is the presence of a second Polyphenol oxidase (PPO) in Cytosol, which is also plausible owing to avoiding the zigzag pathway and therefore avoiding wasteful transport between distinct cellular compartments. Due to the difficulty in extracting another known isoform of Polyphenol oxidase (PPO) from *Cephalotaxus hainanensis* cDNA, we were unable to confirm the probable occurrence of another isoform of PPO in Cytosol which could be subject of further investigations regarding this topic in the future.

4.4. Metabolon structures, the cell´s strategy for efficiently transport products among different cellular compartment could affect the metabolism of cells

Based on the few shikimate pathway intermediates, phenylpropanoid metabolism produces a vast diversity of secondary metabolites (Vogt, 2010). Phenylalanine which is obtained by the shikimate pathway, serves as the starting point for the production of phenylpropanoids. According to this study's findings on sub-cellular localisation, three distinct cellular compartments are involved in this highly dynamic pathway that directs up to 30% of the carbon fixed by plants to the biosynthesis of lignin precursors, one of the secondary metabolites produced through this route. Transporting products among cytoplasm, where PAL enzymes were located, the endoplasmic reticulum, and finally the peroxisomes, where C4H and 4CLs were localized respectively, appears to be a bit costly for cells because cells need lots of energy conducting products in such a highly dynamic pathway. Furthermore, since this pathway uses a lot of biochemical resources, it is closely controlled both transcriptionally and post-transcriptionally (Weisshaar and Jenkins, 1998) (which was also confirmed by the qPCR analysis in 3.3) by the respective downstream products. Each branch of the route has well-coordinated gene expression. For instance, as it was already mentioned in section 3.3, trans-Cinnamic acid behaves as a feedback regulator of the expression and enzymatic activity of L-phenylalanine ammonia-lyase, the first enzyme in the phenylpropanoid pathway (PAL). This finding was in accordance with Blount et al. 2000, who created transgenic tobacco lines with varied activity levels of cinnamic acid 4-hydroxylase (C4H), the second enzyme in the process, by either sense or antisense expression of an alfalfa C4H cDNA. According to their findings, PAL activity and phenylpropanoid compound levels were decreased in the leaves and stems of plants with genetically down-regulated C4H activity. However, in plants whose PAL activity had been suppressed by gene silencing, C4H activity was not decreased. It has been found out that in plants with reduced C4H expression, the level of PAL activity is roughly proportional to the amount of C4H activity, but a decrease in PAL activity does not significantly affect the level of C4H activity. Plants that are overexpressing PAL due to a bean PAL transgene can also have their PAL activity reduced genetically by down-regulating C4H. These findings support the idea that feedback regulation of PAL detected by Cinnamic acid synthesis, at least in part, controls flow into the phenylpropanoid pathway.

Therefore, a kind of metabolic channeling between enzymes at the pathway's entry point provides a second level of regulation, in addition to transporting products more effectively across various cellular compartments could be proposed (Ralston & Yu, 2006).

According to a number of assessments, the collaborating enzymes from the phenylpropanoid pathway were thought to be arranged into complexes termed metabolons (Bassard & Halkier, 2018; Jørgensen et al., 2005; Laursen et al., 2015; Sweetlove & Fernie, 2013; Winkel-Shirley, 1999). Metabolons are multi-enzymatic complexes that are connected to the membranes. The term "metabolon" was initially used by Paul Srere to describe a supramolecular complex comprising successive metabolic enzymes and cellular structural components (Srere, 1985). The identification of metabolons within the cell is still difficult, despite significant technological breakthroughs in biochemistry, molecular biology, and genetics during the past 38 years from their first introduction by Srere in 1985. The majority of metabolon models are predicated on a dynamic, non-covalent aggregation of elements on the surface of the endoplasmic reticulum (ER). At the cellular level, optimizing biosynthesis is accomplished by organizing enzymes into metabolons. These structures offer direct transport of intermediates between succeeding enzymes, increasing local substrate concentration around the enzyme active center, minimization of highly biologically active and potentially toxic intermediates within the cell, and the coordination of reactions leading to various branches of pathways with shared enzymes or intermediates (Bassard & Halkier, 2018; Jørgensen et al., 2005). Biosynthetic pathways involved in both general and specialized metabolism describe the assembly of sequential enzymes into efficient dynamic supramolecular assemblies known as metabolons. (Bassard, Moller, & Laursen, 2017; Srere, 1987). This expands on the basic concept of compartmentalizing molecular-level biosynthetic activity. Dynamic enzyme reconfigurations perform a role in the creation of metabolons on-demand, which is required for metabolic diversity and flexibility (Laursen et al., 2016; Moller, 2010). The formation of this transient structure, known as a metabolon, when needed, could be an excellent strategy for cells to overcome the high costs of transporting products between various cellular compartments in the highly dynamic general phenylpropanoid pathway in order to produce the highly important p-Coumaroyl-CoA molecule, which contribute significantly to flux channeling towards different phenylpropanoid biosynthetic pathways like flavonoids, lignin and stilbene production.

The PAL-C4H pair was described as the first instance of metabolic channeling in the phenylpropanoid pathway. A purified microsomal fraction from potato tuber was shown to have a partial association with PAL, according to (Czichi & Kindl, 1975, 1977). A channeling theory was reinforced by the discovery that this fraction converted Phe into p-Coumaric acid more effectively than exogenously provided Cinnamate. As a result of additional research, PAL1, but not PAL2, was discovered to localize to the ER in tobacco (*Nicotiana tabacum*) (Rasmussen & Dixon, 1999).

Exogenous radiolabeled Cinnamate did not equilibrate with the pool of Cinnamate generated directly from PAL in tobacco cell cultures or microsomal membranes. Following C4H overexpression, both green fluorescent protein (GFP)-tagged PALs were partly relocalized to the ER (Achnine et al., 2004), and competition studies suggested that PAL1 had a greater affinity for ER binding sites. Although the findings suggested a loose association, colocalisation with C4H was validated by double immunolabeling and fluorescence resonance energy transfer (FRET). *Nicotiana benthamiana* has more recently provided evidence that PAL1 is localized close to the ER membrane (Bassard et al., 2012).

Membrane-bound proteins, particularly cytochrome P450 enzymes, are another example which are involved in the formation of metabolons in the instance of plant secondary metabolism (Jrgensen et al., 2005; Ralston and Yu, 2006). The primary protein fold of P450 enzymes often linked to the cytoplasmic surface of the endoplasmic reticulum (ER) via their N terminus (Bayburt and Sligar, 2002). Therefore, P450s can only move in two dimensions. They may be linked to membrane domains with a particular lipid composition or to proteins that build membranes. The metabolism of flavonoids and cyanogenic glucosides has also been supported by P450 anchoring of plant metabolons (Winkel, 2004; Jrgensen et al., 2005). The two *Arabidopsis* cytochrome P450s (P450s), CYP73A5 (C4H) and CYP98A3 (C3H) mediating para- and meta-hydroxylations of the phenolic ring of monolignols, were discovered to colocalize in the endoplasmic reticulum (ER) and to form homo- and heteromers. Chen et al. revealed interaction between poplar C4H and C3H, which benefits the activity of both proteins (Chen et al., 2011). Plant CYP73A5 and CYP98A3 on the other hand, seemed to be rapidly transmitting with the highly dynamic plant ER, whereas their lateral diffusion on the membrane's surface was limited (Bassard et al., 2012). Thus, the mobility of plant P450 enzymes appears to be mostly dependent on the actin/myosin cytoskeleton and the high ER flexibility, which has been reported as a noteworthy feature of plant cells (Sparkes et al., 2009). Fast streaming of the plant ER allows for faster exchanges between membrane-anchored enzymes and cytosolic proteins and metabolites. It might thus be a crucial role in improving plant metabolism. Another aspect that may limit the mobility of CYP73A5 and CYP98A3 is their participation in large homo/heteropolymers and/or more complicated supramolecular complexes. Such multimeric complexes may have limited mobility due to their size, but they may also be connected with membrane microdomains or membrane protein(s). The formation of P450 oligomers may potentially facilitate metabolite flux transfer and offer a platform for bridging soluble enzymes (Bassard et al.,

2012). In addition to the possible regulatory impact, P450 interaction in oligomers would enhance the local concentration of P450 enzymes and their products.

The presence of a PAL-C4H pair as the first instance of metabolic channeling in the phenylpropanoid pathway, as well as the formation of internal (*Nicotiana tabacum*) CYP73 (C4H) and CYP98 (C3H) oligomers in *Ch*.PAL overexpressor lines, could accurately describe Ferulic acid production in nearly all PAL over expressed cell lines. As demonstrated in Figure 4.2, supramolecular assemblies known as metabolons (between PAL and C4H and also between C4H and C3H) that were marked with purple color led the Cinnamic acid resulted from overexpression of PAL towards the formation of Caffeic acid. Caffeic acid accumulation in cells may boost COMT activity, resulting in Ferulic acid synthesis.

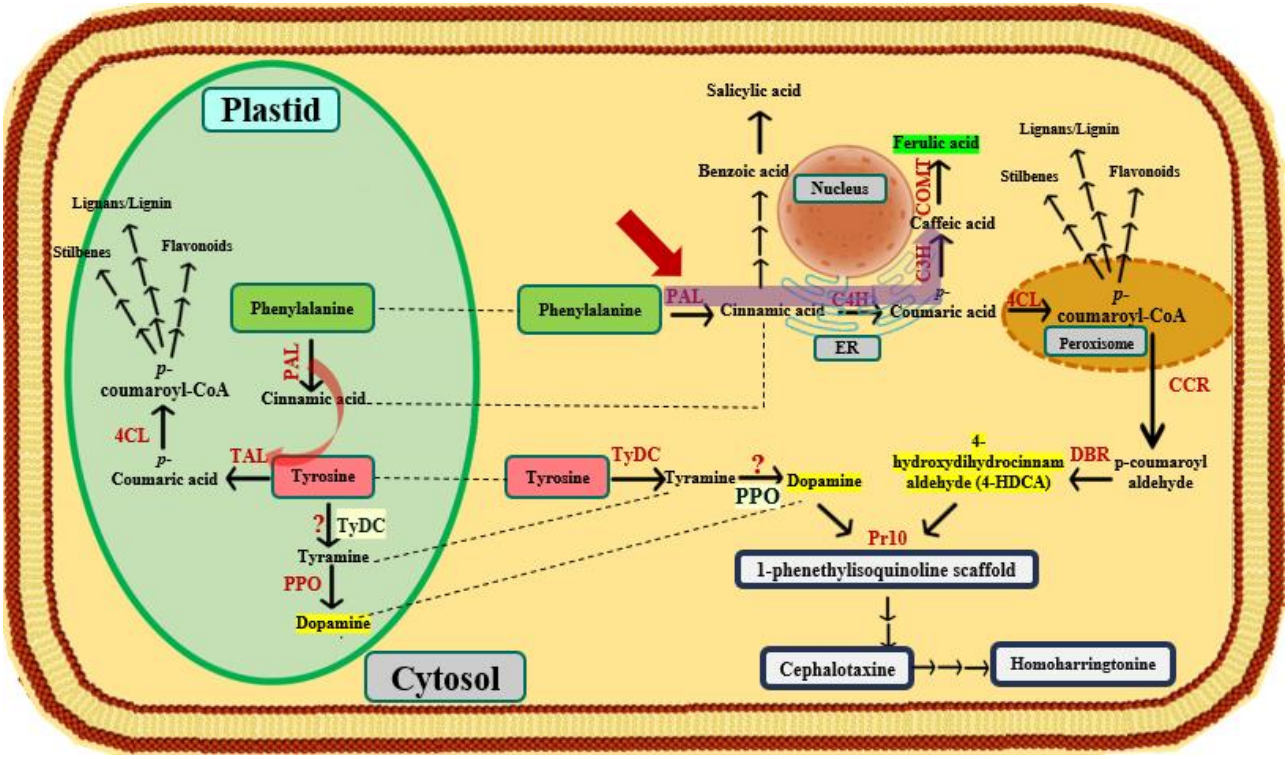


Figure 4.2. Schematic illustration of metabolite channeling among PAL and C4H and also between C4H and C3H as indicated with purple color

This interpretation is completely in accordance with the definition of "metabolons" (also known as "metabolite channeling" or "substrate channeling") which refers to a temporary structural-functional complex formed between sequential enzymes of a metabolic pathway, wherein the intermediates are maintained within the metabolon and the interconversions are catalyzed by sequential enzymes. Metabolons, which are actually enzyme super complexes, are formed by transient interactions

between proteins and surfaces that keep substrate channeling accessible. This allows the metabolon's constituent parts to associate and dissociate to control metabolic flow (Zhang & Fernie, 2021) (Figure 4.3).

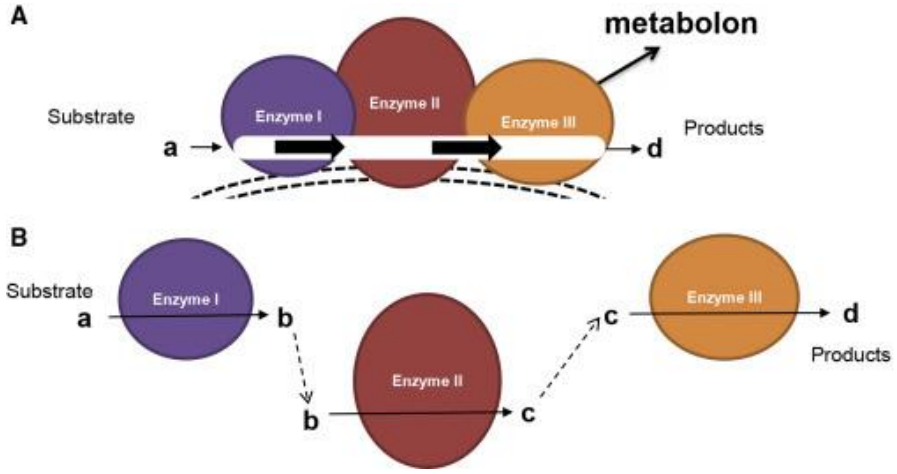


Figure 4.3. The substrate channeling association (Zhang & Fernie, 2021)

Up to now the results of this study demonstrated that compartmentalization is a mechanism to steer the pathway into different directions notwithstanding of transient metabolon formation for facilitating the transportation among different cellular compartment. Therefore, to proof of this concept we try to redirect the production of PAL over expressed cell lines by altering the cellular compartment localisation.

4.6. Sub-cellular compartmentalization drives metabolic flux towards possible intermediates or desired side pathways

In order to enhance the production of a particular metabolites, metabolic engineering seeks to redirect cellular molecules to a chosen biosynthetic pathway (Nielsen & Keasling, 2016). This objective is typically accomplished by altering the levels of enzyme expression, such as raising the enzymes required for product formation while lowering the expression of enzymes in competing, endogenous pathways. (Ajikumar et al., 2010; Keasling, 2010). But it can be difficult to produce a designed route at sustained high levels beside the fact that permanently activating a route, might

have negative effects, such as competing with other pathways and building up hazardous intermediates. Therefore, it may be necessary to use small-molecule inducers to activate enzyme expression at a specific time or to an intermediate level in order to avoid competing with key natural pathways by diverting flux to artificial routes (Lalwani et al., 2018). The burden of manufacturing these high enzyme loads can also make the level of overexpression required for optimal product synthesis harmful (Tomala & Korona, 2013). Therefore, in order to achieve a desired product, there would be significant benefits to boosting the flux through a desired pathway on demand without changing the levels of enzyme expression, by manipulating the cellular compartmentalization since natural metabolic systems compartmentalize enzymes into organelles in order to overcome comparable difficulties or, in certain cases, post-translationally stimulate activation. In this study, we present evidence that modifying cellular compartmentalization in *Nicotiana tabacum* cells overexpressing *Ch*.PAL genes would result in lowering the concentration of side metabolites and conducting the flux through a desired pathway.

Based on the sub-cellular localisation results in this study, we hypothesized that by blocking p-Coumaric acid from exiting the ER, we would be able to alter the route and lower the generation of Ferulic acid. Brefeldin A, a fatty acid produced from a fungus that may prevent protein transport from the endoplasmic reticulum (ER) to the Golgi complex (Helms & Rothman, 1992), was employed to test this theory. Brefeldin A treatment in *Ch*.PAL-3 overexpressor lines, whose most abundant metabolite was Ferulic acid, did not result in any Ferulic acid production. Therefore, we can state that our theory about using our sub-cellular localisation knowledge to shift the route away from the undesirable Ferulic acid and consequently toward lowering the production of this compound was confirmed by Brefeldin A treatment of *Ch*.PAL-3 overexpressor lines. In fact, the production of Coumaric acid and Caffeic acid after seven days, followed by the production of Caffeic acid alone after fourteen days, showed that the COMT enzyme was likely localized somewhere outside of the Endoplasmic reticulum (ER). By using Brefeldin A, Caffeic acid was prevented from leaving the ER which caused this molecule to accumulate in *Ch*.PAL-3 overexpressor lines.

In this study we demonstrated the concept that compartmentalization is a method to steer the pathway in different directions by successfully rerouting the pathway away from the production of Ferulic acid. This study's findings about the sub-cellular localisation of the enzymes participating in this pathway will undoubtedly aid future research into how to direct this extremely dynamic metabolic flux toward preferred side routes or potential intermediates.

5. Conclusion and Outlook

In this study, we attempted to save and preserve the functional genes in the production pathway of the distinctive Cephalotaxine alkaloids precursor, 1-phenethylisoquinoline scaffold, from cDNA of the endangered living fossil, *Cephalotaxus hainanensis*, by isolating them from its cDNA and expressing them stably in *Nicotiana tabacum*. The fluorescent labeling of these genes with GFP allows us to investigate the sub-cellular localisation of genes implicated in this production pathway. Sub-cellular localisation results in this study revealed the presence of two distinct pathways for Cinamoyl CoA production, one occurring entirely in plastids and the other occurring between the cytosol, ER, and Peroxisomes, increasing the efficiency of the process by which metabolites from this compound are directed into different secondary metabolite synthesis pathways. In addition, in this study two independent paths in two dissimilar cellular compartments, one in the cytosol and the other in the plastids, are suggested for converting Tyrosine to Dopamine which require additional investigation to be verified. According to the sub-cellular localisation findings of this study, several cellular compartments are engaged in the early stage of the phenylpropanoid synthesis pathway. The transient formation of metabolons, which acts as metabolic channeling across various cellular compartments among different enzymes at the pathway's entry point, has been proposed in this study as an excellent strategy for cells to overcome the high costs of transporting products between various cellular compartments in the highly dynamic general phenylpropanoid pathway. In addition to compartmentalization studies, we evaluated the in vivo activity of isolated *Ch.*PAL genes as the most critical enzyme at the phenylpropanoid pathway entry point, controlling carbon flow towards secondary metabolite formation. This was accomplished by overexpressing this enzyme under the control of the constitutive CaMV 35S promoter, resulting in the generation of a considerable amount of Ferulic acid through elevating the activity of two endogenous (tobacco) genes, C3H and COMT. Ferulic acid production was even increased further by sodium fluoride (NaF) and Methyl jasmonate (MeJA), although no Cinnamic acid was detected in these overexpressed cell lines. Not producing any Cinnamic acid could also be attributed to the formation of metabolons formed between sequential enzymes of the metabolic pathway, wherein intermediates are maintained within the metabolons, and interconversions are catalyzed by sequential enzymes.

By Using the cellular localisation results of this study, we attempted to reroute the production pathway of *Nicotiana tabacum* cells overexpressing *Ch.*PAL by changing compartmentalization. The outcomes of this experiment demonstrated that compartmentalization can be used as a strategy

to steer the biological pathways in different directions and to successfully reroute them away from the production of undesirable products. These has been achieved by using Brefeldin A, fungal derivative fatty acid which could inhibit protein transport from the endoplasmic reticulum (ER) to the Golgi complex (Helms & Rothman, 1992), to shifted the metabolic route away from production of unwanted Ferulic acid towards production of Caffeic and Coumaric acid. The findings of this study regarding the subcellular localisation of the enzymes involved in this pathway will surely benefit future research into how to direct this incredibly dynamic metabolic flux into favored side routes or possible intermediates.

According to gene expression analysis data in this study trans-Cinnamic acid may operate as a feedback regulator of the expression and enzymatic activity of PAL which could be regulated both transcriptionally and post-transcriptionally. Therefore, it is suggested that releasing PAL from feedback inhibition of Cinnamic acid would provide a powerful route to increase phenylpropanoid and flavonoid production. To reach this idea, this study proposes utilizing transgenic tobacco lines with Co-overexpression of the first and the second enzyme of the pathway, PAL and C4H, to reduce the inhibitory effect of Cinnamic acid´s accumulation by consuming it for Coumaric acid production in future studies. This study also suggests using well-characterized fungal strains such as *Saccharomyces cerevisiae* due to their less complicated structure and ability to synthesize Phenylalanine and Tyrosine due to harboring shikimate pathway, as a better host for heterologous synthesis of 1-phenethylisoquinoline scaffold for future investigations as they grow well on inexpensive carbon sources, and have had significant success with classic strain modification and process optimization tactics leading in multi-gram per liter yields of the necessary medicinal compounds.

6. References

Abdelkafi, H., & Nay, B. (2012). Natural products from Cephalotaxus sp.: chemical diversity and synthetic aspects. *Natural Product Reports*, *29*(8), 845. https://doi.org/10.1039/c2np20037f

Achan, J., Talisuna, A. O., Erhart, A., Yeka, A., Tibenderana, J. K., Baliraine, F. N., Rosenthal, P. J., & D'Alessandro, U. (2011). Quinine, an old anti-malarial drug in a modern world: role in the treatment of malaria. *Malaria Journal*, *10*(1), 144. https://doi.org/10.1186/1475-2875-10-144

Achnine, L., Blancaflor, E. B., Rasmussen, S., & Dixon, R. A. (2004). Colocalization of l-Phenylalanine Ammonia-Lyase and Cinnamate 4-Hydroxylase for Metabolic Channeling in Phenylpropanoid Biosynthesis. *The Plant Cell*, *16*(11), 3098–3109. https://doi.org/10.1105/tpc.104.024406

Ajikumar, P. K., Xiao, W.-H., Tyo, K. E. J., Wang, Y., Simeon, F., Leonard, E., Mucha, O., Phon, T. H., Pfeifer, B., & Stephanopoulos, G. (2010). Isoprenoid Pathway Optimization for Taxol Precursor Overproduction in *Escherichia coli*. *Science*, *330*(6000), 70–74. https://doi.org/10.1126/science.1191652

Allina, S. M., Pri-Hadash, A., Theilmann, D. A., Ellis, B. E., & Douglas, C. J. (1998). 4-Coumarate:Coenzyme A Ligase in Hybrid Poplar1. *Plant Physiology*, *116*(2), 743–754. https://doi.org/10.1104/pp.116.2.743

Amirkia, V., & Heinrich, M. (2014). Alkaloids as drug leads – A predictive structural and biodiversity-based analysis. *Phytochemistry Letters*, *10*, xlviii–liii. https://doi.org/10.1016/j.phytol.2014.06.015

Anantharaju, P. G., Gowda, P. C., Vimalambike, M. G., & Madhunapantula, S. v. (2016). An overview on the role of dietary phenolics for the treatment of cancers. *Nutrition Journal*, *15*(1), 99. https://doi.org/10.1186/s12937-016-0217-2

Barros, J., & Dixon, R. A. (2020). Plant Phenylalanine/Tyrosine Ammonia-lyases. *Trends in Plant Science*, *25*(1), 66–79. https://doi.org/10.1016/j.tplants.2019.09.011

Barros, J., Serk, H., Granlund, I., & Pesquet, E. (2015). The cell biology of lignification in higher plants. *Annals of Botany*, *115*(7), 1053–1074. https://doi.org/10.1093/aob/mcv046

Bassard, J.-E., & Halkier, B. A. (2018). How to prove the existence of metabolons? *Phytochemistry Reviews*, *17*(2), 211–227. https://doi.org/10.1007/s11101-017-9509-1

Bassard, J.-E., Richert, L., Geerinck, J., Renault, H., Duval, F., Ullmann, P., Schmitt, M., Meyer, E., Mutterer, J., Boerjan, W., de Jaeger, G., Mely, Y., Goossens, A., & Werck-Reichhart, D. (2012). Protein–Protein and Protein–Membrane Associations in the Lignin Pathway. *The Plant Cell*, *24*(11), 4465–4482. https://doi.org/10.1105/tpc.112.102566

Bayburt, T. H., & Sligar, S. G. (2002). Single-molecule height measurements on microsomal cytochrome P450 in nanometer-scale phospholipid bilayer disks. *Proceedings of the National Academy of Sciences*, *99*(10), 6725–6730. https://doi.org/10.1073/pnas.062565599

Beaudoin, G. A. W., & Facchini, P. J. (2014). Benzylisoquinoline alkaloid biosynthesis in opium poppy. *Planta*, *240*(1), 19–32. https://doi.org/10.1007/s00425-014-2056-8

Becker, H. (2000). Functions of Plant Secondary Metabolites and their Exploitation in Biotechnology. *Annual Plant Reviews*, *3*(2), 362. https://doi.org/10.1016/S0176-1617(00)80201-1

Blount, J. W., Korth, K. L., Masoud, S. A., Rasmussen, S., Lamb, C., & Dixon, R. A. (2000). Altering Expression of Cinnamic Acid 4-Hydroxylase in Transgenic Plants Provides Evidence for a Feedback Loop at the Entry Point into the Phenylpropanoid Pathway. *Plant Physiology*, *122*(1), 107–116. https://doi.org/10.1104/pp.122.1.107

Bocar, M., Jossang, A., & Bodo, B. (2003). New Alkaloids from *Cephalotaxus f ortunei*. *Journal of Natural Products*, *66*(1), 152–154. https://doi.org/10.1021/np0203178

Boerjan, W., Ralph, J., & Baucher, M. (2003). Lignin Biosynthesis. *Annual Review of Plant Biology*, *54*(1), 519–546. https://doi.org/10.1146/annurev.arplant.54.031902.134938

Bonamore, A., Rovardi, I., Gasparrini, F., Baiocco, P., Barba, M., Molinaro, C., Botta, B., Boffi, A., & Macone, A. (2010). An enzymatic, stereoselective synthesis of (S)-norcoclaurine. *Green Chemistry*, *12*(9), 1623. https://doi.org/10.1039/c0gc00036a

Bourgaud, F., Gravot, A., Milesi, S., & Gontier, E. (2001). Production of plant secondary metabolites: a historical perspective. *Plant Science*, *161*(5), 839–851. https://doi.org/10.1016/S0168-9452(01)00490-3

Buschmann, H., Green, P., Sambade, A., Doonan, J. H., & Lloyd, C. W. (2011). Cytoskeletal dynamics in interphase, mitosis and cytokinesis analysed through *Agrobacterium* -mediated transient transformation of tobacco BY-2 cells. *New Phytologist*, *190*(1), 258–267. https://doi.org/10.1111/j.1469-8137.2010.03587.x

Calabrese, J. C., Jordan, D. B., Boodhoo, A., Sariaslani, S., & Vannelli, T. (2004). Crystal structure of phenylalanine ammonia lyase: multiple helix dipoles implicated in catalysis. *Biochemistry*, *43*(36), 11403–11416. https://doi.org/10.1021/bi049053+

Capistrano, R., Vangestel, C., Wouters, A., Dockx, Y., Pauwels, P., Stroobants, S., Apers, S., Lardon, F., Pieters, L., & Staelens, S. (2016). Efficacy Screening of *Gloriosa Superba* Extracts in a Murine Pancreatic Cancer Model Using ^{18} F-FDG PET/CT for Monitoring Treatment Response. *Cancer Biotherapy and Radiopharmaceuticals*, *31*(3), 99–109. https://doi.org/10.1089/cbr.2015.1954

Caputi, L., Franke, J., Farrow, S. C., Chung, K., Payne, R. M. E., Nguyen, T.-D., Dang, T.-T. T., Soares Teto Carqueijeiro, I., Koudounas, K., Dugé de Bernonville, T., Ameyaw, B., Jones, D. M., Vieira, I. J. C., Courdavault, V., & O'Connor, S. E. (2018). Missing enzymes in the biosynthesis of the anticancer drug vinblastine in Madagascar periwinkle. *Science*, *360*(6394), 1235–1239. https://doi.org/10.1126/science.aat4100

Carocha, V., Soler, M., Hefer, C., Cassan-Wang, H., Fevereiro, P., Myburg, A. A., Paiva, J. A. P., & Grima-Pettenati, J. (2015). Genome-wide analysis of the lignin toolbox of *<scp>E</scp> ucalyptus grandis* . *New Phytologist*, *206*(4), 1297–1313. https://doi.org/10.1111/nph.13313

Chang, A., Lim, M.-H., Lee, S.-W., Robb, E. J., & Nazar, R. N. (2008). Tomato Phenylalanine Ammonia-Lyase Gene Family, Highly Redundant but Strongly Underutilized. *Journal of Biological Chemistry*, *283*(48), 33591–33601. https://doi.org/10.1074/jbc.M804428200

Chen, H.-C., Li, Q., Shuford, C. M., Liu, J., Muddiman, D. C., Sederoff, R. R., & Chiang, V. L. (2011). Membrane protein complexes catalyze both 4- and 3-hydroxylation of cinnamic acid derivatives in monolignol biosynthesis. *Proceedings of the National Academy of Sciences*, *108*(52), 21253–21258. https://doi.org/10.1073/pnas.1116416109

Ching, S. L. K., Gidda, S. K., Rochon, A., van Cauwenberghe, O. R., Shelp, B. J., & Mullen, R. T. (2012). Glyoxylate Reductase Isoform 1 is Localized in the Cytosol and Not Peroxisomes in Plant Cells.

Journal of Integrative Plant Biology, *54*(3), 152–168. https://doi.org/10.1111/j.1744-7909.2012.01103.x

Cochrane, F. C., Davin, L. B., & Lewis, N. G. (2004). The Arabidopsis phenylalanine ammonia lyase gene family: kinetic characterization of the four PAL isoforms. *Phytochemistry*, *65*(11), 1557–1564. https://doi.org/10.1016/j.phytochem.2004.05.006

Conn, H. O. (1981). Making sense out of nonsense. *Hepatology*, *1*(4), 370–372. https://doi.org/10.1002/hep.1840010416

Costa, M. A., Collins, R. E., Anterola, A. M., Cochrane, F. C., Davin, L. B., & Lewis, N. G. (2003). An in silico assessment of gene function and organization of the phenylpropanoid pathway metabolic networks in Arabidopsis thaliana and limitations thereof. *Phytochemistry*, *64*(6), 1097–1112. https://doi.org/10.1016/S0031-9422(03)00517-X

Courdavault, V., O'Connor, S. E., Oudin, A., Besseau, S., & Papon, N. (2020a). Towards the Microbial Production of Plant-Derived Anticancer Drugs. *Trends in Cancer*, *6*(6), 444–448. https://doi.org/10.1016/j.trecan.2020.02.004

Czichi, U., & Kindl, H. (1975). Formation of p-coumaric acid and o-coumaric acid from L-phenylalanine by microsomal membrane fractions from potato: Evidence of membrane-bound enzyme complexes. *Planta*, *125*(2), 115–125. https://doi.org/10.1007/BF00388698

Czichi, U., & Kindl, H. (1977). Phenylalanine ammonia lyase and cinnamic acid hydroxylases as assembled consecutive enzymes on microsomal membranes of cucumber cotyledons: Cooperation and subcellular distribution. *Planta*, *134*(2), 133–143. https://doi.org/10.1007/BF00384962

Davin, L. B., Jourdes, M., Patten, A. M., Kim, K.-W., Vassão, D. G., & Lewis, N. G. (2008). Dissection of lignin macromolecular configuration and assembly: Comparison to related biochemical processes in allyl/propenyl phenol and lignan biosynthesis. *Natural Product Reports*, *25*(6), 1015. https://doi.org/10.1039/b510386j

Dong, N., & Lin, H. (2021). Contribution of phenylpropanoid metabolism to plant development and plant–environment interactions. *Journal of Integrative Plant Biology*, *63*(1), 180–209. https://doi.org/10.1111/jipb.13054

Ehlting, J., Buttner, D., Wang, Q., Douglas, C. J., Somssich, I. E., & Kombrink, E. (1999a). Three 4-coumarate:coenzyme A ligases in Arabidopsis thaliana represent two evolutionarily divergent classes in angiosperms. *The Plant Journal*, *19*(1), 9–20. https://doi.org/10.1046/j.1365-313X.1999.00491.x

Ehlting, J., Shin, J. J. K., & Douglas, C. J. (2001). Identification of 4-coumarate:coenzyme A ligase (4CL) substrate recognition domains. *The Plant Journal*, *27*(5), 455–465. https://doi.org/10.1046/j.1365-313X.2001.01122.x

Ehrenworth, A. M., & Peralta-Yahya, P. (2017). Accelerating the semisynthesis of alkaloid-based drugs through metabolic engineering. *Nature Chemical Biology*, *13*(3), 249–258. https://doi.org/10.1038/nchembio.2308

Emiliani, G., Fondi, M., Fani, R., & Gribaldo, S. (2009). A horizontal gene transfer at the origin of phenylpropanoid metabolism: a key adaptation of plants to land. *Biology Direct*, *4*(1), 7. https://doi.org/10.1186/1745-6150-4-7

Fabricant, D. S., & Farnsworth, N. R. (2001). The Value of Plants Used in Traditional Medicine for Drug Discovery. *Environmental Health Perspectives*, *109*, 69. https://doi.org/10.2307/3434847

Ferrer, J.-L., Austin, M. B., Stewart, C., & Noel, J. P. (2008). Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. *Plant Physiology and Biochemistry*, *46*(3), 356–370. https://doi.org/10.1016/j.plaphy.2007.12.009

Finer, JohnJ., Vain, P., Jones, MarkW., & McMullen, MichaelD. (1992). Development of the particle inflow gun for DNA delivery to plant cells. *Plant Cell Reports*, *11*(7). https://doi.org/10.1007/BF00233358

FRESNO, M., JIMENEZ, A., & VAZQUEZ, D. (1977). Inhibition of Translation in Eukaryotic Systems by Harringtonine. *European Journal of Biochemistry*, *72*(2), 323–330. https://doi.org/10.1111/j.1432-1033.1977.tb11256.x

Fulda, M., Heinz, E., & Wolter, F. P. (1994). The fadD gene of Escherichia coli K12 is located close to rnd at 39.6 min of the chromosomal -map and is a new member of the AMP-binding protein family. *Molecular and General Genetics MGG*, *242*(3), 241–249. https://doi.org/10.1007/BF00280412

Gao, S., Yu, H.-N., Xu, R.-X., Cheng, A.-X., & Lou, H.-X. (2015). Cloning and functional characterization of a 4-coumarate CoA ligase from liverwort Plagiochasma appendiculatum. *Phytochemistry*, *111*, 48–58. https://doi.org/10.1016/j.phytochem.2014.12.017

Goujon, T., Sibout, R., Eudes, A., MacKay, J., & Jouanin, L. (2003). Genes involved in the biosynthesis of lignin precursors in Arabidopsis thaliana. *Plant Physiology and Biochemistry*, *41*(8), 677–687. https://doi.org/10.1016/S0981-9428(03)00095-0

Gould, S. J., Keller, G. A., Hosken, N., Wilkinson, J., & Subramani, S. (1989). A conserved tripeptide sorts proteins to peroxisomes. *Journal of Cell Biology*, *108*(5), 1657–1664. https://doi.org/10.1083/jcb.108.5.1657

Gross, G. G. (1981). *The Biochemistry of Lignification* (pp. 25–63). https://doi.org/10.1016/S0065-2296(08)60032-4

Hagel, J. M., & Facchini, P. J. (2010). Dioxygenases catalyze the O-demethylation steps of morphine biosynthesis in opium poppy. *Nature Chemical Biology*, *6*(4), 273–275. https://doi.org/10.1038/nchembio.317

Han, R. (1994). Highlight on the studies of anticancer drugs derived from plants in china. *STEM CELLS*, *12*(1), 53–63. https://doi.org/10.1002/stem.5530120110

Hartung, E. F. (1954). History of the Use of Colchicum and related Medicaments in Gout: With Suggestions for Further Research. *Annals of the Rheumatic Diseases*, *13*(3), 190–200. https://doi.org/10.1136/ard.13.3.190

Helms, J. B., & Rothman, J. E. (1992). Inhibition by brefeldin A of a Golgi membrane enzyme that catalyses exchange of guanine nucleotide bound to ARF. *Nature*, *360*(6402), 352–354. https://doi.org/10.1038/360352a0

Herrendorff, R., Faleschini, M. T., Stiefvater, A., Erne, B., Wiktorowicz, T., Kern, F., Hamburger, M., Potterat, O., Kinter, J., & Sinnreich, M. (2016). Identification of Plant-derived Alkaloids with Therapeutic Potential for Myotonic Dystrophy Type I. *Journal of Biological Chemistry*, *291*(33), 17165–17177. https://doi.org/10.1074/jbc.M115.710616

Hommel, M. (2008). The future of artemisinins: natural, synthetic or recombinant? *Journal of Biology*, *7*(10), 38. https://doi.org/10.1186/jbiol101

https://www.lifetechnologies.com. (n.d.).

http://www.qiagen.com. (n.d.).

http://www. roboklon.com. (n.d.).

Huang, J., Gu, M., Lai, Z., Fan, B., Shi, K., Zhou, Y.-H., Yu, J.-Q., & Chen, Z. (2010). Functional Analysis of the Arabidopsis *PAL* Gene Family in Plant Growth, Development, and Response to Environmental Stress . *Plant Physiology*, *153*(4), 1526–1538. https://doi.org/10.1104/pp.110.157370

Huang, M. T. (1975). Harringtonine, an inhibitor of initiation of protein biosynthesis. *Molecular Pharmacology*, *11*(5), 511–519.

Hudlicky, Tomas., & W. Reed, Josephine. (2007). *Evolution of Design and Methods for Natural Products*.

Hu, Q., Min, L., Yang, X., Jin, S., Zhang, L., Li, Y., Ma, Y., Qi, X., Li, D., Liu, H., Lindsey, K., Zhu, L., & Zhang, X. (2018). Laccase GhLac1 Modulates Broad-Spectrum Biotic Stress Tolerance via Manipulating Phenylpropanoid Pathway and Jasmonic Acid Synthesis. *Plant Physiology*, *176*(2), 1808–1823. https://doi.org/10.1104/pp.17.01628

Hu, W.-J., Kawaoka, A., Tsai, C.-J., Lung, J., Osakabe, K., Ebinuma, H., & Chiang, V. L. (1998). Compartmentalized expression of two structurally and functionally distinct 4-coumarate:CoA ligase genes in aspen (*Populus tremuloides*). *Proceedings of the National Academy of Sciences*, *95*(9), 5407–5412. https://doi.org/10.1073/pnas.95.9.5407

Hu, Y., Li, W. Ch., Xu, Y. Q., Li, G. J., Liao, Y., & Fu, F.-L. (2009). Differential expression of candidate genes for lignin biosynthesis under drought stress in maize leaves. *Journal of Applied Genetics*, *50*(3), 213–223. https://doi.org/10.1007/BF03195675

Hyun, M. W., Yun, Y. H., Kim, J. Y., & Kim, S. H. (2011). Fungal and Plant Phenylalanine Ammonia-lyase. *Mycobiology*, *39*(4), 257–265. https://doi.org/10.5941/MYCO.2011.39.4.257

Ilari, A., Franceschini, S., Bonamore, A., Arenghi, F., Botta, B., Macone, A., Pasquo, A., Bellucci, L., & Boffi, A. (2009). Structural Basis of Enzymatic (S)-Norcoclaurine Biosynthesis. *Journal of Biological Chemistry*, *284*(2), 897–904. https://doi.org/10.1074/jbc.M803738200

Jin, J., Wang, J.-X., Chen, F.-F., Wu, D.-P., Hu, J., Zhou, J.-F., Hu, J.-D., Wang, J.-M., Li, J.-Y., Huang, X.-J., Ma, J., Ji, C.-Y., Xu, X.-P., Yu, K., Ren, H.-Y., Zhou, Y.-H., Tong, Y., Lou, Y.-J., Ni, W.-M., … Chen, S.-J. (2013). Homoharringtonine-based induction regimens for patients with de-novo acute myeloid leukaemia: a multicentre, open-label, randomised, controlled phase 3 trial. *The Lancet Oncology*, *14*(7), 599–608. https://doi.org/10.1016/S1470-2045(13)70152-9

John, I. (2000). Wink M. 1999. Biochemistry of plant secondary metabolism. Annual plant reviews, Volume 2. 374pp. Sheffield: Sheffield Academic Press Ltd. £85 (hardback) and Functions of plant secondary metabolites and their exploitation in biotechnology. Annual plant reviews, Volume 3. 370pp. Sheffield: Sheffield Academic Press Ltd. £79 (hardback). *Annals of Botany*, *86*(1). https://doi.org/10.1006/anbo.2000.1196

Jørgensen, K., Rasmussen, A. V., Morant, M., Nielsen, A. H., Bjarnholt, N., Zagrobelny, M., Bak, S., & Møller, B. L. (2005). Metabolon formation and metabolic channeling in the biosynthesis of plant natural products. *Current Opinion in Plant Biology*, *8*(3), 280–291. https://doi.org/10.1016/j.pbi.2005.03.014

J Patterson, B Chamberlain, & D Thayer. (2004). *Finch tv version 1.4. 0. Publ. by Authors*.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A.,

Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

Karimi, M., Inzé, D., & Depicker, A. (2002). GATEWAY™ vectors for Agrobacterium-mediated plant transformation. *Trends in Plant Science*, *7*(5), 193–195. https://doi.org/10.1016/S1360-1385(02)02251-3

Kawasaki, T., Koita, H., Nakatsubo, T., Hasegawa, K., Wakabayashi, K., Takahashi, H., Umemura, K., Umezawa, T., & Shimamoto, K. (2006). Cinnamoyl-CoA reductase, a key enzyme in lignin biosynthesis, is an effector of small GTPase Rac in defense signaling in rice. *Proceedings of the National Academy of Sciences*, *103*(1), 230–235. https://doi.org/10.1073/pnas.0509875103

Kawatra, A., Dhankhar, R., Mohanty, A., & Gulati, P. (2020). Biomedical applications of microbial phenylalanine ammonia lyase: Current status and future prospects. *Biochimie*, *177*, 142–152. https://doi.org/10.1016/j.biochi.2020.08.009

Keasling, J. D. (2010). Manufacturing Molecules Through Metabolic Engineering. *Science*, *330*(6009), 1355–1358. https://doi.org/10.1126/science.1193990

Kenrick, P., & Crane, P. R. (1997). The origin and early evolution of plants on land. *Nature*, *389*(6646), 33–39. https://doi.org/10.1038/37918

K.M. Chaves, S., M. Feitosa, C., & da S. Araújo, L. (2016). Alkaloids Pharmacological Activities - Prospects for the Development of Phytopharmaceuticals for Neurodegenerative Diseases. *Current Pharmaceutical Biotechnology*, *17*(7), 629–635. https://doi.org/10.2174/1389201017071605032015 41

Kong, K.-H., Lee, J.-L., Park, H.-J., & Cho, S.-H. (1998). Purification and characterization of the tyrosinase isozymes of pine needles. *IUBMB Life*, *45*(4), 717–724. https://doi.org/10.1080/15216549800203122

KOUKOL, J., & CONN, E. E. (1961). The metabolism of aromatic compounds in higher plants. IV. Purification and properties of the phenylalanine deaminase of Hordeum vulgare. *The Journal of Biological Chemistry*, *236*, 2692–2698.

Kuo, M.-C., Chang, S.-J., & Hsieh, M.-C. (2015). Colchicine Significantly Reduces Incident Cancer in Gout Male Patients. *Medicine*, *94*(50), e1570. https://doi.org/10.1097/MD.0000000000001570

Lacombe, E., Hawkins, S., van Doorsselaere, J., Piquemal, J., Goffner, D., Poeydomenge, O., Boudet, A.-M., & Grima-Pettenati, J. (1997). Cinnamoyl CoA reductase, the first committed enzyme of the lignin branch biosynthetic pathway: cloning, expression and phylogenetic relationships. *The Plant Journal*, *11*(3), 429–441. https://doi.org/10.1046/j.1365-313X.1997.11030429.x

Lalwani, M. A., Zhao, E. M., & Avalos, J. L. (2018). Current and future modalities of dynamic control in metabolic engineering. *Current Opinion in Biotechnology*, *52*, 56–65. https://doi.org/10.1016/j.copbio.2018.02.007

LANG, X.-D., SU, J.-R., LU, S.-G., & ZHANG, Z.-J. (2013). A taxonomic revision of the genus <i>Cephalotaxus</i> (Taxaceae). *Phytotaxa*, *84*(1). https://doi.org/10.11646/phytotaxa.84.1.1

Larsson, S., & Ronsted, N. (2013). Reviewing Colchicaceae Alkaloids – Perspectives of Evolution on Medicinal Chemistry. *Current Topics in Medicinal Chemistry*, *14*(2), 274–289. https://doi.org/10.2174/1568026613666131216110417

Laursen, T., Møller, B. L., & Bassard, J.-E. (2015). Plasticity of specialized metabolism as mediated by dynamic metabolons. *Trends in Plant Science*, *20*(1), 20–32. https://doi.org/10.1016/j.tplants.2014.11.002

Lauvergeat, V., Lacomme, C., Lacombe, E., Lasserre, E., Roby, D., & Grima-Pettenati, J. (2001). Two cinnamoyl-CoA reductase (CCR) genes from Arabidopsis thaliana are differentially expressed during development and in response to infection with pathogenic bacteria. *Phytochemistry*, *57*(7), 1187–1195. https://doi.org/10.1016/S0031-9422(01)00053-X

Lavhale, S. G., Kalunke, R. M., & Giri, A. P. (2018). Structural, functional and evolutionary diversity of 4-coumarate-CoA ligase in plants. *Planta*, *248*(5), 1063–1078. https://doi.org/10.1007/s00425-018-2965-z

Lazar, T. (2005). Biochemistry and molecular and biology of plants B. Buchanan, W. Gruissem and R. L. Jones (eds), American Society of Plant Physiologists (distribution through Wiley & Sons), xxxix + 1367 pp., £100 ($175), ISBN 0-943088-37-2 (hardback); £75 ($135), ISBN 0-943088-39-9 (paperback) (2002). *Cell Biochemistry and Function*, *23*(2), 148–148. https://doi.org/10.1002/cbf.1131

Lee, E.-J., & Facchini, P. (2010). Norcoclaurine Synthase Is a Member of the Pathogenesis-Related 10/Bet v1 Protein Family. *The Plant Cell*, *22*(10), 3489–3503. https://doi.org/10.1105/tpc.110.077958

Liang, B., Li, C., Ma, C., Wei, Z., Wang, Q., Huang, D., Chen, Q., Li, C., & Ma, F. (2017). Dopamine alleviates nutrient deficiency-induced stress in Malus hupehensis. *Plant Physiology and Biochemistry*, *119*, 346–359. https://doi.org/10.1016/j.plaphy.2017.09.012

Li, J., Lee, E.-J., Chang, L., & Facchini, P. J. (2016). Genes encoding norcoclaurine synthase occur as tandem fusions in the Papaveraceae. *Scientific Reports*, *6*(1), 39256. https://doi.org/10.1038/srep39256

Liu, Q., Gao, T., Liu, W., Liu, Y., Zhao, Y., Liu, Y., Li, W., Ding, K., Ma, F., & Li, C. (2020). Functions of dopamine in plants: a review. *Plant Signaling & Behavior*, *15*(12), 1827782. https://doi.org/10.1080/15592324.2020.1827782

Li, Y.-C. (2014). Enhanced cephalotaxine production in Cephalotaxus mannii suspension cultures by combining glycometabolic regulation and elicitation. *Process Biochemistry*, *49*(12), 2279–2284. https://doi.org/10.1016/j.procbio.2014.10.005

Lois, R., Dietrich, A., Hahlbrock, K., & Schulz, W. (1989). A phenylalanine ammonia-lyase gene from parsley: structure, regulation and identification of elicitor and light responsive cis-acting elements. *The EMBO Journal*, *8*(6), 1641–1648. https://doi.org/10.1002/j.1460-2075.1989.tb03554.x

Lundström, J., & Agurell, S. (1971). Biosynthesis of mescaline and tetrahydroisoquinoline alkaloids in Lophophora williamsii (Lem.) Coult. *Acta Pharmaceutica Suecica*, *8*(3), 261–274.

Lü, S., & Wang, J. (2014). Homoharringtonine and omacetaxine for myeloid hematological malignancies. *Journal of Hematology & Oncology*, *7*(1), 2. https://doi.org/10.1186/1756-8722-7-2

Mann, J. D., & Mudd, S. H. (1963). Alkaloids and Plant Metabolism. *Journal of Biological Chemistry*, *238*(1), 381–385. https://doi.org/10.1016/S0021-9258(19)84008-5

Miao, Y., Lang, X., Li, S., Su, J., & Wang, Y. (2012). Characterization of 15 Polymorphic Microsatellite Loci for Cephalotaxus oliveri (Cephalotaxaceae), a Conifer of Medicinal Importance. *International Journal of Molecular Sciences*, *13*(9), 11165–11172. https://doi.org/10.3390/ijms130911165

Miura, S., Kasuya-Arai, I., Mori, H., Miyazawa, S., Osumi, T., Hashimoto, T., & Fujiki, Y. (1992). Carboxyl-terminal consensus Ser-Lys-Leu-related tripeptide of peroxisomal proteins functions in vitro as a minimal peroxisome-targeting signal. *The Journal of Biological Chemistry*, *267*(20), 14405–14411.

Miyazawa, S., Osumi, T., Hashimoto, T., Ohno, K., Miura, S., & Fujiki, Y. (1989). Peroxisome targeting signal of rat liver acyl-coenzyme A oxidase resides at the carboxy terminus. *Molecular and Cellular Biology*, *9*(1), 83–91. https://doi.org/10.1128/MCB.9.1.83

Morita, H., Nagakura, Y., Hosoya, T., Ekasari, W., Widyawaruyanti, A., Mori-Yasumoto, K., Sekita, S., & Hirasawa, Y. (2010). Cephastigiamide A, and Antiplasmodial Activity of Cephalotaxus Alkaloids from Cephalotaxus harringtonia Forma Fastigiata. *HETEROCYCLES*, *81*(2), 441. https://doi.org/10.3987/COM-09-11870

Morita, H., Yoshinaga, M., & Kobayashi, J. (2002). Cephalezomines G, H, J, K, L, and M, new alkaloids from Cephalotaxus harringtonia var. nana. *Tetrahedron*, *58*(27), 5489–5495. https://doi.org/10.1016/S0040-4020(02)00521-5

Morris, J. S., Caldo, K. M. P., Liang, S., & Facchini, P. J. (2021). PR10/Bet v1-like Proteins as Novel Contributors to Plant Biochemical Diversity. *ChemBioChem*, *22*(2), 264–287. https://doi.org/10.1002/cbic.202000354

Nagata, T., Nemoto, Y., & Hasezawa, S. (1992). *Tobacco BY-2 Cell Line as the "HeLa" Cell in the Cell Biology of Higher Plants* (pp. 1–30). https://doi.org/10.1016/S0074-7696(08)62452-3

Nett, R. S., Lau, W., & Sattely, E. S. (2020). Discovery and engineering of colchicine alkaloid biosynthesis. *Nature*, *584*(7819), 148–153. https://doi.org/10.1038/s41586-020-2546-8

Nick, P. (2016). Life versus 'biomass'—why application needs cell biology. *Protoplasma*, *253*(5), 1175–1176. https://doi.org/10.1007/s00709-016-1014-7

Nielsen, J., & Keasling, J. D. (2016). Engineering Cellular Metabolism. *Cell*, *164*(6), 1185–1197. https://doi.org/10.1016/j.cell.2016.02.004

Oyebode, O., Kandala, N.-B., Chilton, P. J., & Lilford, R. J. (2016). Use of traditional medicine in middle-income countries: a WHO-SAGE study. *Health Policy and Planning*, *31*(8), 984–991. https://doi.org/10.1093/heapol/czw022

Parry, R. J., Chang, M. N. T., Schwab, J. M., & Foxman, B. M. (1980). Biosynthesis of the Cephalotaxus alkaloids. Investigations of the early and late stages of cephalotaxine biosynthesis. *Journal of the American Chemical Society*, *102*(3), 1099–1111. https://doi.org/10.1021/ja00523a029

Paudler, W. W., Kerley, G. I., & McKay, J. (1963). The Alkaloids of Cephalotaxus drupacea and Cephalotaxus fortunei. *The Journal of Organic Chemistry*, *28*(9), 2194–2197. https://doi.org/10.1021/jo01044a010

Paul, A. G. (1973). Biosynthesis of the peyote alkaloids. *Lloydia*, *36*(1), 36–45.

Pérard-Viret, J., Quteishat, L., Alsalim, R., Royer, J., & Dumas, F. (2017). *Cephalotaxus Alkaloids* (pp. 205–352). https://doi.org/10.1016/bs.alkal.2017.07.001

Petrovska, B. (2012). Historical review of medicinal plants′ usage. *Pharmacognosy Reviews*, *6*(11), 1. https://doi.org/10.4103/0973-7847.95849

Piquemal, J., Lapierre, C., Myton, K., O'connell, A., Schuch, W., Grima-pettenati, J., & Boudet, A.-M. (2002). Down-regulation of Cinnamoyl-CoA Reductase induces significant changes of lignin profiles in transgenic tobacco plants. *The Plant Journal*, *13*(1), 71–83. https://doi.org/10.1046/j.1365-313X.1998.00014.x

Polturak, G., & Aharoni, A. (2018). "La Vie en Rose": Biosynthesis, Sources, and Applications of Betalain Pigments. *Molecular Plant*, *11*(1), 7–22. https://doi.org/10.1016/j.molp.2017.10.008

Powell, R. G., Weisleder, D., Smith, C. R., & Rohwedder, W. K. (1970). Structures of harringtonine, isoharringtonine, and homoharringtonine. *Tetrahedron Letters*, *11*(11), 815–818. https://doi.org/10.1016/S0040-4039(01)97839-6

Qian, Y., Lynch, J. H., Guo, L., Rhodes, D., Morgan, J. A., & Dudareva, N. (2019). Completion of the cytosolic post-chorismate phenylalanine biosynthetic pathway in plants. *Nature Communications*, *10*(1), 15. https://doi.org/10.1038/s41467-018-07969-2

Qiao, F., Cong, H., Jiang, X., Wang, R., Yin, J., Qian, D., Wang, Z., & Nick, P. (2014). De Novo Characterization of a Cephalotaxus hainanensis Transcriptome and Genes Related to Paclitaxel Biosynthesis. *PLoS ONE*, *9*(9), e106900. https://doi.org/10.1371/journal.pone.0106900

Raes, J., Rohde, A., Christensen, J. H., van de Peer, Y., & Boerjan, W. (2003). Genome-Wide Characterization of the Lignification Toolbox in Arabidopsis . *Plant Physiology*, *133*(3), 1051–1071. https://doi.org/10.1104/pp.103.026484

Ragg, H., Kuhn, D. N., & Hahlbrock, K. (1981). Coordinated regulation of 4-coumarate:CoA ligase and phenylalanine ammonia-lyase mRNAs in cultured plant cells. *Journal of Biological Chemistry*, *256*(19), 10061–10065. https://doi.org/10.1016/S0021-9258(19)68741-7

Ralston, L., & Yu, O. (2006). Metabolons involving plant cytochrome P450s. *Phytochemistry Reviews*, *5*(2–3), 459–472. https://doi.org/10.1007/s11101-006-9014-4

Rasmussen, & Dixon. (1999). Transgene-mediated and elicitor-induced perturbation of metabolic channeling at the entry point into the phenylpropanoid pathway. *The Plant Cell*, *11*(8), 1537–1552. https://doi.org/10.1105/tpc.11.8.1537

Renault, H., de Marothy, M., Jonasson, G., Lara, P., Nelson, D. R., Nilsson, I., André, F., von Heijne, G., & Werck-Reichhart, D. (2017). Gene Duplication Leads to Altered Membrane Topology of a Cytochrome P450 Enzyme in Seed Plants. *Molecular Biology and Evolution*, *34*(8), 2041–2056. https://doi.org/10.1093/molbev/msx160

Ritter, H., & Schulz, G. E. (2004). Structural Basis for the Entrance into the Phenylpropanoid Metabolism Catalyzed by Phenylalanine Ammonia-Lyase. *The Plant Cell*, *16*(12), 3426–3436. https://doi.org/10.1105/tpc.104.025288

Robin, J.-P., Dhal, R., Dujardin, G., Girodier, L., Mevellec, L., & Poutot, S. (1999). The first semi-synthesis of enantiopure homoharringtonine via anhydrohomoharringtonine from a preformed chiral acyl moiety. *Tetrahedron Letters*, *40*(15), 2931–2934. https://doi.org/10.1016/S0040-4039(99)00327-5

Ro, D.-K., & Douglas, C. J. (2004). Reconstitution of the Entry Point of Plant Phenylpropanoid Metabolism in Yeast (Saccharomyces cerevisiae). *Journal of Biological Chemistry*, *279*(4), 2600–2607. https://doi.org/10.1074/jbc.M309951200

Rohde, A., Morreel, K., Ralph, J., Goeminne, G., Hostyn, V., de Rycke, R., Kushnir, S., van Doorsselaere, J., Joseleau, J.-P., Vuylsteke, M., van Driessche, G., van Beeumen, J., Messens, E., & Boerjan, W.

(2004). Molecular Phenotyping of the *pal1* and *pal2* Mutants of *Arabidopsis thaliana* Reveals Far-Reaching Consequences on Phenylpropanoid, Amino Acid, and Carbohydrate Metabolism. *The Plant Cell*, *16*(10), 2749–2771. https://doi.org/10.1105/tpc.104.023705

Saito, K., & Matsuda, F. (2010). Metabolomics for Functional Genomics, Systems Biology, and Biotechnology. *Annual Review of Plant Biology*, *61*(1), 463–489. https://doi.org/10.1146/annurev.arplant.043008.092035

Schattat, M. H., Griffiths, S., Mathur, N., Barton, K., Wozny, M. R., Dunn, N., Greenwood, J. S., & Mathur, J. (2012). Differential Coloring Reveals That Plastids Do Not Form Networks for Exchanging Macromolecules. *The Plant Cell*, *24*(4), 1465–1477. https://doi.org/10.1105/tpc.111.095398

Sewalt, VJH., Ni, W., Blount, J. W., Jung, H. G., Masoud, S. A., Howles, P. A., Lamb, C., & Dixon, R. A. (1997). Reduced Lignin Content and Altered Lignin Composition in Transgenic Tobacco Down-Regulated in Expression of L-Phenylalanine Ammonia-Lyase or Cinnamate 4-Hydroxylase. *Plant Physiology*, *115*(1), 41–50. https://doi.org/10.1104/pp.115.1.41

Sheng, X., & Himo, F. (2019). Enzymatic Pictet–Spengler Reaction: Computational Study of the Mechanism and Enantioselectivity of Norcoclaurine Synthase. *Journal of the American Chemical Society*, *141*(28), 11230–11238. https://doi.org/10.1021/jacs.9b04591

Shi, R., Sun, Y.-H., Li, Q., Heber, S., Sederoff, R., & Chiang, V. L. (2010). Towards a Systems Approach for Lignin Biosynthesis in Populus trichocarpa: Transcript Abundance and Specificity of the Monolignol Biosynthetic Genes. *Plant and Cell Physiology*, *51*(1), 144–163. https://doi.org/10.1093/pcp/pcp175

Shiva Hemmati. (2015). Phenylalanine ammonia-lyase through evolution: A bioinformatic approach. *Trends in Pharmaceutical Sciences*, 10–14.

Sivakumar, G. (2013). Colchicine Semisynthetics: Chemotherapeutics for Cancer? *Current Medicinal Chemistry*, *20*(7), 892–898. https://doi.org/10.2174/092986713805219073

Sofowora, A., Ogunbodede, E., & Onayade, A. (2013). The role and place of medicinal plants in the strategies for disease prevention. *African Journal of Traditional, Complementary and Alternative Medicines*, *10*(5). https://doi.org/10.4314/ajtcam.v10i5.2

Srere, P. A. (1967). Enzyme Concentrations in Tissues. *Science*, *158*(3803), 936–937. https://doi.org/10.1126/science.158.3803.936

Srere, P. A. (1985). The metabolon. *Trends in Biochemical Sciences*, *10*(3), 109–110. https://doi.org/10.1016/0968-0004(85)90266-X

Stander, E. A., Papon, N., & Courdavault, V. (2021). Puzzling Out the Colchicine Biosynthetic Pathway. *ChemMedChem*, *16*(4), 621–623. https://doi.org/10.1002/cmdc.202000633

Steiner, U., Schliemann, W., & Strack, D. (1996). Assay for Tyrosine Hydroxylation Activity of Tyrosinase from Betalain-Forming Plants and Cell Cultures. *Analytical Biochemistry*, *238*(1), 72–75. https://doi.org/10.1006/abio.1996.0253

Stuible, H.-P., Büttner, D., Ehlting, J., Hahlbrock, K., & Kombrink, E. (2000). Mutational analysis of 4-coumarate:CoA ligase identifies functionally important amino acids and verifies its close relationship to other adenylate-forming enzymes. *FEBS Letters*, *467*(1), 117–122. https://doi.org/10.1016/S0014-5793(00)01133-9

Sudhir Kumar, Glen Stecher, & Koichiro Tamura. (2016). *MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Molecular biology and evolution, 33(7)* (pp. 1870–1874).

Sweetlove, L. J., & Fernie, A. R. (2013). The Spatial Organization of Metabolism Within the Plant Cell. *Annual Review of Plant Biology*, *64*(1), 723–746. https://doi.org/10.1146/annurev-arplant-050312-120233

Swinkels, B. W., Gould, S. J., Bodnar, A. G., Rachubinski, R. A., & Subramani, S. (1991). A novel, cleavable peroxisomal targeting signal at the amino-terminus of the rat 3-ketoacyl-CoA thiolase. *The EMBO Journal*, *10*(11), 3255–3262. https://doi.org/10.1002/j.1460-2075.1991.tb04889.x

Szczebara, F. M., Chandelier, C., Villeret, C., Masurel, A., Bourot, S., Duport, C., Blanchard, S., Groisillier, A., Testet, E., Costaglioli, P., Cauet, G., Degryse, E., Balbuena, D., Winter, J., Achstetter, T., Spagnoli, R., Pompon, D., & Dumas, B. (2003). Total biosynthesis of hydrocortisone from a simple carbon source in yeast. *Nature Biotechnology*, *21*(2), 143–149. https://doi.org/10.1038/nbt775

Teutsch, H. G., Hasenfratz, M. P., Lesot, A., Stoltz, C., Garnier, J. M., Jeltsch, J. M., Durst, F., & Werck-Reichhart, D. (1993). Isolation and sequence of a cDNA encoding the Jerusalem artichoke cinnamate 4-hydroxylase, a major plant cytochrome P450 involved in the general phenylpropanoid pathway. *Proceedings of the National Academy of Sciences*, *90*(9), 4102–4106. https://doi.org/10.1073/pnas.90.9.4102

Tomala, K., & Korona, R. (2013). Evaluating the Fitness Cost of Protein Expression in Saccharomyces cerevisiae. *Genome Biology and Evolution*, *5*(11), 2051–2060. https://doi.org/10.1093/gbe/evt154

Torrens-Spence, M. P., Chiang, Y.-C., Smith, T., Vicent, M. A., Wang, Y., & Weng, J.-K. (2020). Structural basis for divergent and convergent evolution of catalytic machineries in plant aromatic amino acid decarboxylase proteins. *Proceedings of the National Academy of Sciences*, *117*(20), 10806–10817. https://doi.org/10.1073/pnas.1920097117

Tsukamoto, T., Hata, S., Yokota, S., Miura, S., Fujiki, Y., Hijikata, M., Miyazawa, S., Hashimoto, T., & Osumi, T. (1994). Characterization of the signal peptide at the amino terminus of the rat peroxisomal 3-ketoacyl-CoA thiolase precursor. *The Journal of Biological Chemistry*, *269*(8), 6001–6010.

van Acker, R., Vanholme, R., Storme, V., Mortimer, J. C., Dupree, P., & Boerjan, W. (2013). Lignin biosynthesis perturbations affect secondary cell wall composition and saccharification yield in Arabidopsis thaliana. *Biotechnology for Biofuels*, *6*(1), 46. https://doi.org/10.1186/1754-6834-6-46

Vanholme, R., Demedts, B., Morreel, K., Ralph, J., & Boerjan, W. (2010). Lignin Biosynthesis and Structure. *Plant Physiology*, *153*(3), 895–905. https://doi.org/10.1104/pp.110.155119

Vogt, T. (2010). Phenylpropanoid Biosynthesis. *Molecular Plant*, *3*(1), 2–20. https://doi.org/10.1093/mp/ssp106

Wang, C., Guo, Z., Huang, X., & Huang, L. (2016). Development and Characterization of 15 Microsatellite Markers for *Cephalotaxus fortunei* (Cephalotaxaceae). *Applications in Plant Sciences*, *4*(5), 1500129. https://doi.org/10.3732/apps.1500129

Wang, Y., Subrizi, F., Carter, E. M., Sheppard, T. D., Ward, J. M., & Hailes, H. C. (2022). Enzymatic synthesis of benzylisoquinoline alkaloids using a parallel cascade strategy and tyrosinase variants. *Nature Communications*, *13*(1), 5436. https://doi.org/10.1038/s41467-022-33122-1

Weigel, D., & Glazebrook, J. (2006). Transformation of *Agrobacterium* Using the Freeze-Thaw Method. *Cold Spring Harbor Protocols*, *2006*(7), pdb.prot4666. https://doi.org/10.1101/pdb.prot4666

WENGENMAYER, H., EBEL, J., & GRISEBACH, H. (1976). Enzymic Synthesis of Lignin Precursors. Purification and Properties of a Cinnamoyl-CoA: NADPH Reductase from Cell Suspension Cultures of Soybean (Glycine max). *European Journal of Biochemistry*, *65*(2), 529–536. https://doi.org/10.1111/j.1432-1033.1976.tb10370.x

Winkel-Shirley, B. (1999). Evidence for enzyme complexes in the phenylpropanoid and flavonoid pathways. *Physiologia Plantarum*, *107*(1), 142–149. https://doi.org/10.1034/j.1399-3054.1999.100119.x

Xiong, Z., Wang, L., Sun, J., Jiang, X., Cong, H., Sun, H., & Qiao, F. (2022). Functional characterization of a Colchicum autumnale L. double-bond reductase (CaDBR1) in colchicine biosynthesis. *Planta*, *256*(5), 95. https://doi.org/10.1007/s00425-022-04003-0

Xu, D., Lin, H., Tang, Y., Huang, L., Xu, J., Nian, S., & Zhao, Y. (2021). Integration of full-length transcriptomics and targeted metabolomics to identify benzylisoquinoline alkaloid biosynthetic genes in Corydalis yanhusuo. *Horticulture Research*, *8*(1), 16. https://doi.org/10.1038/s41438-020-00450-6

Yi, X., Gao, L., Wang, B., Su, Y.-J., & Wang, T. (2013). The Complete Chloroplast Genome Sequence of Cephalotaxus oliveri (Cephalotaxaceae): Evolutionary Comparison of Cephalotaxus Chloroplast DNAs and Insights into the Loss of Inverted Repeat Copies in Gymnosperms. *Genome Biology and Evolution*, *5*(4), 688–698. https://doi.org/10.1093/gbe/evt042

Yoshinaga, M., Morita, H., Dota, T., & Kobayashi, J. (2004). Bis-cephalezomines A–E from Cephalotaxus harringtonia var. nana. *Tetrahedron*, *60*(36), 7861–7868. https://doi.org/10.1016/j.tet.2004.06.070

Yu, R., Zhu, J., Wang, M., & Wen, W. (2015). Biosynthesis and regulation of terpenoid indole alkaloids in Catharanthus roseus. *Pharmacognosy Reviews*, *9*(17), 24. https://doi.org/10.4103/0973-7847.156323

Zhang, Y., & Fernie, A. R. (2021). Metabolons, enzyme–enzyme assemblies that mediate substrate channeling, and their roles in plant metabolism. *Plant Communications*, *2*(1), 100081. https://doi.org/10.1016/j.xplc.2020.100081

Zhao, Q., & Dixon, R. A. (2011). Transcriptional networks for lignin biosynthesis: more complex than we thought? *Trends in Plant Science*, *16*(4), 227–233. https://doi.org/10.1016/j.tplants.2010.12.005

Zheng, W., Qiu, L., Wang, R., Feng, X., Han, Y., Zhu, Y., Chen, D., Liu, Y., Jin, L., & Li, Y. (2015). Selective targeting of PPARγ by the natural product chelerythrine with a unique binding mode and improved antidiabetic potency. *Scientific Reports*, *5*(1), 12222. https://doi.org/10.1038/srep12222

Zhong, J., Qing, J., Wang, Q., Liu, C., Du, H., Liu, P., Du, Q., Du, L., & Wang, L. (2022). Genome-Wide Identification and Expression Analyses of the 4-Coumarate: CoA Ligase (4CL) Gene Family in Eucommia ulmoides. *Forests*, *13*(8), 1253. https://doi.org/10.3390/f13081253

Zhou, R., Jackson, L., Shadle, G., Nakashima, J., Temple, S., Chen, F., & Dixon, R. A. (2010). Distinct cinnamoyl CoA reductases involved in parallel routes to lignin in *Medicago truncatula*. *Proceedings of the National Academy of Sciences*, *107*(41), 17803–17808. https://doi.org/10.1073/pnas.1012900107

Ziegler, J., & Facchini, P. J. (2008). Alkaloid Biosynthesis: Metabolism and Trafficking. *Annual Review of Plant Biology*, *59*(1), 735–769. https://doi.org/10.1146/annurev.arplant.59.032607.092730

7. Appendix

Gene sequences used in this study:

>*Ch*PAL1 (MN398159), Number of Nucleotides: 2235 bp
ATGTTTTCAATGGCGGAGGTTGCAGCAGCAGCAGCTAGCTATGATCATTTAAACTGGTCTAAAGCCGC
GAAGTTGTTGGTAGGCAGCCATGTGAAAGAAGTGAGGAACATGGTCGAAGAGTTTCGTGGAATCAAG
ACCGTATCTTTGGGAGGATCGAACCTGAGCATAGCCCAGGTTGCTGCAATCGCTCGCAGGCCCGAGG
TGAAGGTGCTGCTAGATGAAAAACATGCCAAGAGCAGAGTCGATGAAAGCTCAAATTGGGTTATCAA
CAACATCTGCAAGGGCACAGACACCTATGGAGTCACAACTGGATTCGGCGCCACGTCGCACAGACGA
ACGGACCAGACGTTCGAGTTGCAGAAAGAATTGATACGGTTCTTAAATGCCGGAGTCGTGGGCGACG
ATGGGGTCAACTGTTTGTCTGCGTCTGCTACCAGGGCTGCCATGCTGGTGCGAACGAATACGCTCATG
CAGGGCTACTCTGGTATCAGGTGGGAAATTTTGGAAGCACTGCAGAAGTTGATGGATGCCAACGTTA
CGCCTAAGTTGCCTTTGAGGGGAACTATTACTGCGTCTGGAGACCTGGTGCCTCTGTCCTACATTGCG
GGGCTCCTCACCGGAAGGGGTAATTGTGTTGCTGTTGCGGAAGATGGGACTGAGATGAGTGCCGATG
AGGGTTTGAAATTGGCTGGGGTTTGTAAGGCCTTCGAGCTGTATCCCAAGGAAGGGCTGGCGCTCGTC
AATGGAACTGCGGTAGGTGCCGCGGTTGCTTGTACTGTTTGTTATGATGCCAATGTTTTGGCTGTTTTT
GCAGAGGTGGCTTCTGCCATGTTCTGCGAGGTGATGCAGGGTAAACCAGAGTTCACAGATCCTTTGAC
ACACAAGCTGAAGCACCATCCAGGCCAAATGGAAGCCGCAGCAGTCATGGAGTGGCTTCTAGACGGG
AGCTCTTACATCAAATCAGCAGCCACACTGCACGAAACGGACCCCTTAAAAAAGCCCAAGCAAGACC
GCTACGCCCTCCGCACCTCCCCTCAGTGGCTGGGCCCCCAGATCGAAGTCATCCGCATGGCCACGCAC
GCAATCAACAGAGAAATAAATTCTGTTAACGACAATCCTCTTATAGACGTAGCCCGAGACAAAGCGC
TGCACGGGGGCAACTTTCAGGGCACACCTATTGGCGTCTCCATGGACAACGTTCGCCTGGCCTTGGCC
GCCATAGCCAAGCTCATTTTCGCTCAGTTCTCCGAACTCGTCAACGACTACTACAACAACGGCCTGCC
GTCAAATCTTTCGGGTGGCGCAAACCCTAGCCTGGACTACGGCCTCAAGGGCGCAGAGATAGCCATG
GCCTCGTATACTTCCGAGATCCAGTATCTGGCAAATCCTGTAACAACCCATGTCCAGAGCGCAGAGCA
ACATAATCAGGATATAAACTCCCTGGGGCTGATTTCTGCCAGAAAGACTGCTGAAGCCATAGAAATA
CTAAAGCTCATGGTGTCCACATATCTGGTTGCTTTATCGCAGGCCATCGATCTCCGTCATCTGGAAGA
GAATTTCCATGGAGGCGTAAAGCAGATGGTATGCCAAGCGGCCAGGAAGACGCTGTATTCAACCACA
GAGGGCTTGTTGTTACCGTCCAGATTCTGCGAGAAGGAGCTCCTGAAAGTTGTGGACAGGCTTCATGT
TTTCACCTACATAGAAGACCCGGCTAACCCTTCTTCGCCATTAATGTTACAGCTTCGCCAAGTGCTGGT
GGAACAAGCGCTGAAATCACTGGACGTCATGAAGGATTTCCCTGTACTTACCACCATATCTCTGTTCG
AAGAAGAGCTCAAAAGAAAACTAGCCGTGGAGGTTCCACTGATACGCCAAAATTACGACAATGGAC
ATTTCGACGTCCACAACCAAATTCAGAATTGCAGAACATATTCGCTGTATGAATTTGTGCGGGAGGAG
CTTGGAACCAGTTTGCTGTGCGGACCTACAGGGAAAACGCCTGGTGAGGACATTGAAAAAGTATTTG
TGGCAATTTCTGAAGGGAAGCTGGATGGGTGTCTGATGGAGTGCTTGGATGGATGGAATGAATCCCC
AGGGCCTTTTGTTGAAGCGAAGAAAAATAATTTATGCGAAGTCCTGAAAACTAACAATGCATGCGGC
TGGAGCTGGTTCCAGCATATTGGAGGGCAGGAAGTGAATGGAGGAAAGGGCTACTGGCTTCTCTCTA
TCATTTGA
>*Ch*PAL2 (MN398160), Number of Nucleotides: 2364 bp
ATGGAGACTTCTCTGAATCAGAACCAGAACGAGAACAGGGGTGTTAGTCTGAGAGAGGCGTGTAGAC
CGCCTACGTTACCTGTGCCGGTGGGCGGGGTGAATCCCATCCACGTGGAAGGGAAAGGCTACGTGAG
CAATCCCCCTCACTGGAAAAAGGCGGCGGAGGCCATGCAATGCAGTCACTTCGATGAAGTGAAGGAG
ATGATCGGGCGGTTCCACGGTGCACAGAAGGTCGTCCTGCGTGGCACCACGTTAACCGTGGCTGAGG
TGAGCGCCATCACCCGCCGACCCGAGGTGAGGGTGGAGTTGGACGAGGCGACCGCGAAACAACGGG
TGGAGGAGAGCGCTACGTGGGTGGCCAACAACATCGCCAAGGGGACGGACACGTACGGGGTCACAA
CGGGCTTCGGAGCCACCTCTCACCGGCGAACGAACAAAGCGGTCGATCTGCAGAGAGAGCTGATCCG
TTTCCTCAATGCGGGAGTGTTGGGCAAAGGCATCAAGGACTCCCTGCAGTTGCCCTGCGAGTACACCA
AGGCGGCCATGCTCGTCAGGACCAATACTCTAATGCAGGGCTATTCTGGCATCCGATGGGAGGTTCTG
CAGGCTCTGCAGAAGCTCGTGGACTGTAACATCACTCCCAAAGTGCCTCTCAGAGGAACTATCACGG
CCTCCGGCGACCTGGTGCCACTCTCCTACATTGCCGGGCTGCTCACCGGGCGGCCCAATTCGAGAGCC
GTGACGGCCGAGGGTGAAGAGGTGGAGGCGCTGGAAGCGCTGAGTAGAGCGGGCATCGTTGGAGGG
CCCTTCGAATTGCAGCCCAAAGAGGGGCTAGCGCTTGTCAACGGCACCGCAGTCGGTTCTGCGGTGG
CGGCTTCCGTGTGCTTCGACGCCAACGTGTTAGCCGTCCTGGCGGAGATTCTCTCGGCCCTGTTCTGCG
AGGTGATGCAGGGGAAGCCCGAGTTTACTGACCCGTTAACTCACGAGCTGAAGCATCACCCGGGCCA
GATTGAAGCGGCGGCGGTGATGGAATATCTCCTGGAAGGCAGCGACTACGTGAAGGAGGCGAAGCG
TGTGCAGGAGACGGACCCGCTGAGCAAACCGAAGCAGGACCGTTACGCGCTGCGGACGTCCCCGCAG

TGGCTGGGCCCGCAGATAGAAGTGATCCGCATGGCCACGCACTCCATCGAGCGAGAGATCAATTCGG
TGAACGACAACCCGTTGATCGACGTAGCGCGCGACATGGCACTCCACGGCGGCAACTTCCAGGGCAC
GCCCATCGGGGTCTCCATGGACAACACCAGAATTGCGCTGGCGGCGGTGGGGAAGCTTATGTTTGCG
CAGTTTTCCGAACTGGTGTGCGATTACTACAACAACGGGCTGCCTTCCAATCTGAGCGGAGGACCAAA
CCCCAGCCTGGACTACGGGTTCAAGGGCGCCGAGATCGCCATGGCGTCCTACACTTCGGAGCTGCAG
TATCTGGCCTCCCCGGTGACCACTCACGTACAGAGCGCAGAGCAACACAACCAGGACGTGAATTCGC
TGGGCCTGATTTCGGCAAGGAAGACGGCGGAGGCCATCCAAATACTGAAGCTGATGTCCGCCACTTA
TCTAGTGGCCCTGTGCCAGGCCGTGGACCTGCGGCACTTGGAGGAAAACATGCGAGCCGTGGTGAAG
CATGTTGTGTCGCAGACGGTGAGGAAGACCCTGTACGCGGCCGAAGACGGAAACCTGCTTCCCACTC
GCTTCTGCGAGAAGGAACTCCTACAAATAATCGAGCATCAGCCCGTCTTCTCCTACATCGACGACCCC
TCCAACCCCTCATACGCCCTCATGCTTCAGCTGCGAGAGGTCCTCGTTGATGAAGCCCTCAAATGCCC
TCTCGAAGAGAACGACCAAAAACTGGGAATGCCCCTGCTATTTCGCAGGATCCCCATCTTCCAGGAG
GAGTTAAAAGCTCGGCTGGACGAAGAGGTCCCCAAGGCCCGCCACAGATTCGACGCGGGAGACTTCC
CCATTCCCAACAGAATTAAGAAATGCAGGACTTACCCCATTTATAAATTTGTAAGATCCGAATTGGGA
ACAGAAATGCTGAGCGGGGCGAAAGTGCGAAGTCCCGGCGAAGACATTGATAAAGTGTACAACGGC
ATTTGCGAGCCGGAGAAGATGTTTGGAGAGGTGCTCGTCAAATGTCTGGAAACCTGGCGCGGATCGG
CAGGGCCCTTCACTCCGCGGCCTCTACATCCTCCACAGTCACCTGCGGGGTTCAATGCTTCGTATTGGT
CATGGTTTGACACCGCCAAATCGCCCTCCGCCACCAGCGGCAGAGGGTTCTGGAACAATTCCTCCCAC
CAAGTGCTCTGA

>*Ch*PAL3 (MN398161), Number of Nucleotides: 2172 bp

ATGGTAGCGGGTGTGGATCTGATGGCCATGCAGCTCAATAATGGCAATGGATTTAGTCACGTCAATTC
AGGCGATGGTTTCTGCGTGGATTTTCAAAATGGCGGGTCAGGGGACCCGTTGAATTGGGGCGAGGCC
TCCCGGGGTGTGCAGGGCAGCCACTACGATGAAGTGAAAGACATGGTGGGGTCCTTCTTCGGAGCCA
AAGAGGTGTCCATTGAAGGAAAGTCGCTATCTGTTGCAGATGTGGCCGCGGTGGCCAGGAGAGCCGA
CGTGAAGGTGAGATTAGACGCGGCAACGGCCAAGTCCCGGGTGGATGAGAGCTCCAACTGGGTTATG
CACAATATGGCCAAAGGCACGGACACATATGGGGTCACCACGGGCTTTGGGGCCACCTCCCACAGGC
GCACCGACCGCGGCGCGGACCTTCAGAAGGAGCTCATCCGCTTTCTCAACGCGGGTGTTTTTTGTAAG
AACGACAACGTGCTTCCCGCGGACACCACTCGGGCGGCTATGCTTGTCCGGACGAACACTCTGATGC
AGGGGTACTCGGGGATCCGGTGGGGGATTCTGGAGGCCATTCAGGACTTTCTGAATGCAGGGATAAC
GCCCAAGTTGCCCTTGAGGGGCACCATCACTTCCTCGGGTGACCTCGTTCCATTGTCTTACATTGCCGG
GCTGGTAACCGGGAGGCCCAATTCCAGAGCCTTGACCCGGGACGGTAAAGAACTAACGGCACTCGAG
GCCCTGAACCGGATCGGGTTGGACAAACCATTTGAACTCCAGCCCAAGGAAGGGCTAGCCATTGTGA
ATGGGACGTCCGTCGGGGCCGCCCTGGCTGCTACCGTCTGCTTTGACGCCAACGTGCTCTGCCTGTTG
GCGGAAACCACGTCTGCCATGTTCTGTGAGGTAATGCAAGGGAAGCCAGAGTTTGCAGACCCGTTGA
CGCACCGGCTTAAGCACCACCCGGGGCAAATAGAGGCCGCTGCTATCATGGAATACATTCTGGAGGG
CTCATCCTACATGAAAGCCGCAGCGAAAAAGCACGAGACGAACCCGCTACAGAAGCCCAAGCAGGA
TAGGTACGCTCTGCGAACCTCGCCACAATGGCTGGGCCCGCAGATAGAGACAGTCCGGGCAGCAACT
CACATGATTGGGCGCGAGATAAACTCCGTGAACGACAACCCAGTGATCGACGTGGCCCGCGACAAGG
CCCTCCATGGTGGCAACTTCCAGGGCACCCCCATCGGCGTCTCGATGGACAACCTTCGCCTGGCCATC
TCCGCCGTGGGTAAGCTCATGTTCGCCCAGTTTTCAGAGTTGGTGAACGATTTTTATAACAACGGCCT
CCCTTCGAACCTGAGCGGCGGGGCCAATCCGAGCCTGGACTACGGATTAAAAGGCGCGGAGATCGCC
ATGGCCTCCTACACCTCGGAGCTCCTCTACCTGGCCTCCCCCGTCACCTCCCACGTCCAGAGCGCCGA
GCAGCACAACCAGGACGTCAACTCGCTGGGCCTCATCTCCGCTCGCAAGTCCGCGGAGGCAGTCGAC
ATCCTCAAGCTCATGAGCGCCACCTACTTAATCGCCCTCTGCCAGGCCGTCGACCTCCGCCATCTCGA
AGAAAACATGCAGAGCACCGTCAAACAAGTGGTCTCCACAGTCGCCAAAAAAACGCTCACCAGCACC
CACAATGGAGAGCTCCTGAACAGCCGCTTCTGCGAAAAAGAGCTACTCCAGGTGGTCGAGAGCCTGC
CGGTGTTCGGCTACATCGACGACCCGTGCAGTCCCAACTACCCCTTAATGCAGAACCTGAGGCAAGTT
CTCGTCCGCCGGGCTCTCGACGACTCCAACGCCTCCGTCTTCAACAAGATTCCCGTTTTCGAGAAAGA
GCTGAACGAGGCACTACAATCAGAGGTGGCTCGGGCCCGTCACGACTATTACGAAAAGGAAATGAGT
GCCCTGCCAAACAGAATCAACGACTGCCGCTCTTATCCTCTTTACGACTTCGTCCGAAACGAGATCGG
CACCCGTCTTTTGTGCGGGGAACGGGCCACGTCCCCAGGCGAGTACATCGAAAAGGTTTACGCCGCC
ATCTGCGGAGGCAAAGTGATTGCTCCTTTGATTAAATGCTTGGATGGTTGGAACGGAGCTCCTGGTCC
TTTCCCCATAGGTTGA

>*Ch*PAL4 (MN398162), Number of Nucleotides: 2139 bp

ATGGCAGCAACACAAGACTTTACAGAAGGAAATCTCTGCCTGGACAACGGAAAAGACAACGGCGCG
TTGCCCTACGACCCCTTGAACTGGGTAGCCGCAGCAGATTCAATGAAGGGCTCCCATCTGGACGATGT
GAAGCAAATGATAGAAGAGTACGGCTGCCCGGTGGTGAAACTCGAGGGCGCCGGGCTGACAGTCGCT
CAAGTAGCCGCGGTGGCCAGGCGAGCCGTCCGGGTGGAGCTCGAACTCGCCGCCCGGCCGCGGGTCG

AGGCGAGCAGCAAGTGGGTGATGGACAGTATGGCGAACGGAACCGACAGCTACGGGGTAACAACAG
GGTTCGGCGCAACCTCACATAGGCGGACCCGCAATGGCGAAGAGCTGCAAAAGGAGCTGATCCGGTT
CCTGAACGCCGGCATGTTTGGGTCGGGCGAAACAAATGTGCTACCATTGGAGACCTCCAGGGCGGCG
GTTCTCGTTAGGGCGAATACGCTTCTGCAAGGGTTTTCAGGAATCCGGTGGGATCTCATAGACGCCAT
GGGCAAGCTTTTGAATGCGGGCGTAACCCCACGGTTGCCGCTAAGGGGGACTATAACGGCGTCGGGC
GATTTGGTCCCTCTGAGTTACATCGCGGGGGTTTTGACAGGCCGGGCGAATGCCAGGGCTGTCGGCCC
GGACGGGTCTGTGATGTCGTCGGCTGAGGCTCTGAAGGCCGCGGGGGTCGTGGACGGGCCGTTTGAG
CTGCGGCCCAAGGAAGGGCTGGCGCTGGTGAACGGAACCGCCGTCGGGTCGGGGCTTGCGTCCACGG
TTCTGTACGATGCCAACATTTTGGTGCTTCTGGCGGAGGTGGCTTCGGCGGTTTTCTGTGAGGTGATGC
AGGGGAAGCCGGAATTTACCGACCATTTGACGCATCGGTTGAAGCATCACCCAGGGCAGATTGAGGC
GGCCGCCATTATGGAGTATGTGTTGGAGGGGAGCTCCTACATGAAGGCTGCTGCCAAGTTGCATGAG
ATGGACCCGTTGAAGAAGCCCAAACAGGACAGATATGCATTGAGGACTTCTCCACAGTGGCTCGGTC
CTTTAGTGGAAGTGATTCGCATAGCCACTCGAATGATCCAACATGAAATTAACTCTGTGAATGACAAT
CCATTGATAGATGTGGCCCGAAACAAGGCCCTACACGGTGGTAACTTTCAAGGCACTCCCATTGGTGT
TTCCATGGACAACACTCGTCTTGCCCTCGCTGCCATTGGTAAGCTCATGTTTGCCCAATTCAGCGAGCT
TGTAAATGACTTCTACAACAATGGTTTACCCTCCAACCTCAGTGGTGGCCCCAACCCAAGCTTGGACT
ATGGATTTAAAGGCGCTGAAATAGCAATGGCATCCTACACCTCTGAGCTACAATTTTTAGCCAATCCT
GTCACAAACCATGTTCAAAGCGCAGAACAACACAACCAGGATGTCAACTCCCTCGGCCTCATCTCAG
CCCGCAAGACAGCCGAGGCTGTTGAGATACTAAAATTAATGTCATCAACATACATAGTGGCTCTTTGT
CAAGCTGTCGACCTCCGACACCTCGAGGAAAACATGCAGTCCACTGTAAAACAAACTGTGGCGCAAG
CATGCAAGAAAACATTAACCATGAACCTCCAAGGCGAGCTCCTTCCTTCACGTTTCTGCGAAAAGGAC
CTCCTCAAGGTCGTTGAAAGAGAGCCTGTTTTCTCCTACATCGATGACCCTTGTAGTGCCTCATACCCT
CTCTTCCAAAAACTTCGACAAGTCTTGGTCGAACATGCCTTACAGAACACCGACAAGGAAAAGGACA
CCAACACCTCAATTTTTCATAGGATCAATGCATTTGAAGAGGCGCTGAAAGAAAGTCTAGTAACTGA
GGTTGAACAAACTCGTGAGGCCTTTGACAAAGGTGTCGCACCAGTGCCCAACAGAATCAAGGAATGT
CGATCGTACCCACTTTATGAGTTCGTGAGACATGAACTTGAAACCTCAATGTTATCGGGAACAAAGAG
CAACTCCCCCGGGCAAGACTTTGATAAAATGTTGCTTGCAATCAATGAAGGCAAGCTAGTGGCACCCT
TGCTCAAGTGTGTGGAAGGTTGGAATGGCG CCCCAATCCCAATTTAA

>*Ch*C4H1 (MW691947), Number of Nucleotides: 1506 bp

ATGGAGATGGATTTGTTTAACGGCCTACTCGCCCTGTTCGTTCTGCTCGTTGGTGCCATATTTTTATCC
AAGCTAACTTCTAAGAAGCTCAGCCTTCCACCAGGTCCCTTGGCGATGCCCATATTTGGCAACTGGCT
GCAAGTTGGGGACGATTTGAATCACCGCAATCTCAGCGTCCTCGCCAAAAAGTATGGCGAAATATTTC
TGCTAAAGATGGGGCAGAGAAATTTGGTGGTGGTTTCGTCGCCAGATTTGGCGACAGAGATGTTGCA
TACGCAGGGAGTGGAGTTCGCGTCTCGTACCCGCAACGTGGTGCACGACATCTTTACCGGCAAAGGG
CAAGACATGGTGTTTACCGTTTACGGCGAACACTGGAGAAAGATGCGGAGAATCATGACACTGCCGT
TTTTTACTAACAAGGTTGTTCAGCAGTACAGATTTGCGTGGGAGGACGAAATTTCGCGTGCGGTCGAG
GATATCAAGAAGAGGCCGGAATCTTCCACTACCGGAATTGTGATCCGCCGGAGGTTGCAGCTCATGA
TGTATAATATTATGTATAAGATGATGTTTGATCGGAGATTTGAGAGTGAGGAGGATCCGCTGTATTTG
AGGCTCAGGGCGCTTAACGGAGAAAGAAGCCGCCTGGCACAGAGTTTTGATTTTAATTATGGAGATT
TTATTCCTATTCTGAGGCCATTTCTGAGAGGGTACCTGAAGATCTGCAAGGAGGTCAAGGAAACGAG
GCTTGCCTTGTTTAAGGATTACTTCGTTGACGAACGAAAGAAATTGGTAAGCACTAATGGCGCTGTGA
ATATGAATGAGAAGGTTGGCATTGATCACATCTTGGAAGCCCAGGAAAAAGGTGAAATCAATGAGGA
TAATGTCCTCTACATTGTGGAGAACATCAACGTCGCAGCGATTGAGACAACCCTGTGGTCAATGGAAT
GGGGCATTGCAGAGATTGTGAATCACCCTGAAATTCAACAAAAAATCCGGGTTGAGCTTGAGACTGT
TCTTGGTAAAGGTGTTCCAATTACAGAGCCAGACATCGTCAAATTGCCTTACCTTCAAGCTGTAGTAA
AGGAAAGCTTGAGGTTGCATATGGCAATCCCATTGCTAGTTCCCCACATGAACCTCAACGATGCGAA
GCTTGGTGGCTATGACATACCGGCTGAGAGTAAAATCCTGGTCAACGCTTGGTGGCTTGCAAACAATC
CTGAATGGTGGGACAAACCTGATGAGTTTATCCCTGAGAGGTTCTTGGAGGAGAAGCTTGAGGCTAG
TGGAAATGATTTCAGATATTTACCATTCGGTGTTGGTAGGAGAAGCTGCCCCGGAATTATCTTGGCTC
TGCCCATTTTGGCCTTGTCCATTGGAAGGCTGGTTCAGAACTTTGAGCTCTTGCCCCCACCTGGGCAAT
CCAAGGTTGACTTAAGTGAGAAAGGAGGGCAGTTCAGCTTGCATATTCTTAACCACTCAGTTGTAGTG
GCTAAACCCAGAGCATGA

>*Ch*C4H2 (MW691948), Number of Nucleotides: 1539 bp

ATGGCGGACGTACAACAAGGATCTCTTGCTGTGCTAGAAAAATCCCTCGTAGCTCTGTTCGTGGTTGT
GGTGGGCGCTGTGCTGGTTAATAAGCTCAGAGGTAAAAAGTTGAAGCTCCCGCCGGGTCCCTTTGCA
GTGCCCATCTTCGGAAACTGGCTGCAAGTGGGAGATGATCTCAATCACAGGAACCTCACCGATTTGGC
CAGGAAGTTTGGCAAAATATTTTTGCTGAAGATGGGGCAGAGGAATCTGGTGGTGGTATCGTCGCCT
GATCTTGCCAAGGAGGTGCTGCACACTCAGGGTGTCGAGTTCGGGTCTCGAACCCGTAACGTAGTGTT

CGATATCTTTACGGGCAAGGGGCAGGACATGGTGTTCACAGTGTATGGAGAACACTGGAGAAAAATG
AGGAGGATCATGACCGTTCCGTTTTTTACTAACAAGGTGGTTCAGCAATATCGATTTGCGTGGGAAGA
GGAGATCGGGCGTGTTGTGGAGGACGTGAGATCCCGTCCTGAAGCTTCGACGTCAGGTATCGTTATAC
GGAGGCGGCTGCAGCTGATGATGTATAACATCATGTACAGGATGATGTTTGATAGACGCTTCGAGAG
CGAGAACGATCCCCTCTTTCTAAAGCTCAGGGCCTTGAACGGAGAGAGAAGTCGCCTCGCTCAGAGC
TTCGATTACAACTATGGAGATTTCATCCCCATTCTCAGGCCCTTTCTCAGAGGGTATCTCAACGTCTGC
AAAGAGGTCAAGGAAAGACGCCTCAATCTATTCAAGGACTACTTCGTCGAAGAGCGCAGGAAGCTTG
GCAGTAGCACGAGTTGTATGAATAACGGAGAGCTCAAGTGTGCCATGGATCATATCTTGGATGCGCA
GAAGAAAGGAGAGATCAACGAGGACAATGTCTTGTACATTGTGGAGAACATTAACGTTGCAGCAATT
GAGACAACACTGTGGTCAATTGAATGGGGAATCGCAGAGTTGGTGAACCACCCACACATCCAACAGA
AAGTACGTGAAGAGTTGGATAATGTGCTTGGGCAAGGAGTCCAAATCACAGAGCCAGACACCACACG
ACTGCCTTATCTACAAGCAGTGATAAAGGAGACTCTACGCTTACGCATGGCAATCCCCTTACTTGTGC
CGCACATGAACCTCCATGATGCCAAGCTTGGTGGCTATGACATCCCCGCAGAGAGCAAGATACTTGT
GAATGCTTGGTATTTGGCCAACAACCCTACCGAATGGAATAACCCTCAACAATTTCGCCCCGAGAGAT
TCCTTGAGGAAGAAAAGAACACAGAAGCTAATGGTAATGATTTTAGATACCTTCCATTTGGTGTGGGT
CGCAGGAGTTGCCCGGGAATTATTTTGGCACTTCCCATATTAGCACTCTCCATTGGACGGCTTGTTCA
AAATTTTGAGCTTCTACCTACACCCGACCACACCCATGTTGATACCACTGAAAAGGGTGGACAATTCA
GCCTTCATATTCTCAACCACTCTCTCATCATTGCCAAGCCCAGAGTTTACTAG

>*Ch*C4H3 (MW691949), Number of Nucleotides: 1455 bp

ATGGCTTCCTTGCTGAACCAGATTTATCTCTCTGTATTTTCACTCAGTGATCACCTCAAACAGCTTGTT
TTAGAATTTCCCTTCCAAATATTTGTAGGTCTATTGGTGACGTTGTTAGTTATAGCCAGAATTGGTAAG
AAAAGCTCCAGACAGCCTCCTGGCCCGTTGTGTTTACCCATATTTGGAAACTGGTTGCAAGTGGGTAA
CGATTTAAACCACAGAAACCTTATGAACATGGCAAAAAAGTATGGGGATGTTTTCATGTTGAGGTTG
GGGTGCAGAAATTTAGTGGTGGTGTCATCTCCTGAGCTTGCCAAACAAGTTCTCCATACCCAGGGCGT
CGAATTCGGTTCGAGGAAGAATAATCTCGTGTTCGATATTTTCACTGGCAATGGCCAAGACATGGTGT
TTACTGTCTATGGAGAGCACTGGCGAAAGATGCGAAGAATCATGACAGTACCCTTCTTTACGCAGAA
AGTTGTCCAGCAGTACAGACAAGGGTGGGAGGAGGAGATCCAGAGAGTCGTCGAAGATGTTCAGGC
GAATGAAATGGCGTTACAGAGCGGTATTGTTATAAGAAAACGGCTTCAGCTTATGCTTTATAATGTTA
TGTACCGCATGATGTTTGATCGCAGGTTCGAGTCTCAGGAGGATCCGCTGTTCGTTCAGGCGACAGCT
TTTAATGCGGAGAGAAGCCGACTGGCACAGAGTTTTGATTACAATTATGGGGATTTCATTCCCATCTT
GAGGCCTTTCCTTAGAGGGTATCTGAAAAAGTGCAGGGATCTGCAGAGCAAGAGGCTTGCATTCTTC
AACGACATTTTCATCCAAGAACGCAGGAAGATATTAGCGAATGGGGAGCAGATGAACAGGTGTGCCA
TAGACCACATATTGGATGCAGAGAAGAAGGGAGAAATAAACGAGCAAAACGTTCTTTACATAGTGGA
GAATATCAATGTGGCGGCGATAGAGACGACTCTGTGGTCAATGGAGTGGGCTCTTGCGGAGCTGGTG
AACCATCCGGAAGTGCAGGGGAAGGTAAGAAGCGAGTTGGAGAGAGTTGTGGGAAAGGGGGTGGCC
GTGAAGGAAGAGCACATAGAGAAGTTGCCATATCTGCAAGCAGTGGTGAAGGAGACGCTGAGGCTG
CACACGCCCATTCCGCTGCTGGTTCCGCACATGAACTCGGACGACGCCTCTCTTGGCGGATTCGCCAT
TCCTAAGGAGAGCAAGGTGGTAGTGAATGCGTGGTGGCTGGCCAACAATCCGGAGTGGTGGAACAAT
TCCTCGGAATTCAGGCCTGAAAGATTTATGGAAGAGGAGGAGGAGACGGAAGCGGCGGTAGGAGGC
AAAGTGGACTTCCGTTTTCTTCCGTTTGGAATGGGGAGGAGGAGTTGCCCTGGCATTATCCTCGCTCT
GCCTGTGTTAGGCCTCGTAATAGGCAGATTATTGCAGAACTTTGAGCTGATTGTGCCGCCGGGTGAGC
AGAAGATCGATGTAAGTGAAAAAGGTGGGCAGTTCAGTCTTCACATCGCCCGCCACTCCACCGTCCG
CTTCCGCCTTCTGTGA

>*Ch*C4H4 (MW691950), Number of Nucleotides: 1500 bp

ATGGATTTGTTTAATGGCCTGGTCGCTCTGTTCGTTGTGCTCGTTGGTGGCATATTTTTATCCAAGTTG
AGGTCCAAAAAACTGAGGCTGCCGCCAGGTCCATTGGCGGTGCCCATATTTGGCAACTGGCTGCAAG
TCGGGGACGATTTGAATCACCGCAACCTCAGCGACTTCGCCAGAAAGTATGGCGAAATATTTCTGCTG
AAGATGGGGCAGAGGAACCTGGTGGTGGTGTCGTCGCCAGATTTGGCGAAGGAGGTGCTGCACACGC
AGGGGGTGGAGTTCGGGTCTCGAACCCGCAACGTGGTGTTCGACATCTTTACCGGCAAAGGACAGGA
CATGGTGTTTACCGTTTACGGCGAACACTGGAGAAAGATGCGAAGAATCATGACGGTGCCGTTTTTCA
CTAACAAGGTCGTTCAGCAGTACAGATTTGCGTGGGAGGACGAAATACAGAGTGCGGTCGAGGATAT
CAAGAAGCGGCCGGAGTCGTCCACGACCGGAATCGTGATCCGCCGGAGGTTGCAGCTCATGATGTAT
AACATTATGTATAAGATGATGTTTGATAGGAGATTTGAGAGCGAGGAGGATCCCTTGTTTTTGAGGCT
CAAGGCGCTCAATGGAGAAAGGAGCCGCTTGGCACAGAGTTTTGATTATAATTACGGAGATTTTATTC
CTATTCTGAGGCCATTTCTGAGAGGGTACCTGAAGATCTGCAAGGAGGTCAAGGAAAGGAGGATCGC
CTTGTTTAAGGATTACTTCGTCGAGGAACGGAAGAAATTGGCAAGCATTAATGGCCCTGTGAACATG
AATGAGAAGGTTGCAATCGATCACATCTTGGAAGCCCAGGACAAAGGTGAAATCAACGAGGATAATG
TTCTGTACATTGTGGAGAACATCAACGTTGCAGCAATTGAAACAACTCTATGGTCCATGGAATGGGGC

ATTGCAGAGATTGTGAACCACCCTGAAATTCAGCAGAAGATCCGGGATGAGCTTGACTCTGTTCTTGG
TAGAGGTGTTCCAATTACAGAGCCAGACACCATCAAATTGCCTTACCTTCAAGCTGTAATCAAGGAAA
CCATGAGGTTACATATGGCAATCCCATTGCTAGTTCCCCACATGAACCTTAATGATGCGAAGCTTGGT
GGCTATGACATACCAGCTGAGAGTAAAATCCTGGTCAACGCTTGGTGGCTTGCTAACAATCCTGAATG
GTGGGACAGACCTGATCAGTTTATTCCTGAGAGGTTTTTGGAGGAGAAGCTTGAAGCCAGTGGAAAT
GATTTCAGATATTTACCATTTGGTGTTGGTAGAAGAAGCTGTCCTGGAATCATCTTGGCTCTGCCCATT
TTGGCCTTGGCCATTGGAAGGCTGGTTCAGAACTTTGAGCTCTTGCCACCACCCGGGCAATCCAAGGT
TGATGTAGCTGAGAAGGGAGGACAGTTCAGCTTGCACATTCTTAACCACTCAGTTGTAGTGGCCAAAC
CCAGAGCATGA

>*Ch*CCR1 (MZ833461), Number of Nucleotides: 975 bp

ATGACTATGGACAATCAGACAGGAGCTGGCCAGACTGTGTGTGTTACAGGAGCTGGAGGGTTCATTG
CTTCATGGCTTGTCAAGCTGCTTTTGGAAAGAGGTTATACTGTGAGAGGAACAGTTCGCAATCCAGAG
GATGCGAAGAATGCACATCTCAAACATCTGGAGGGAGCAGAAGAAAGGCTGATACTGGTAAAAGCG
GATCTTCTGGACTACAACAGCCTTGTAGAAGCAATCAATGGTTGCCAAGGAGTATTCCATGTTGCTTC
TCCAGTTACAGATGACCCAGTTCAAATGGTAGAGCCTGCTGTAAATGGGACCAAAAATGTTCTGGAT
GCCTCTGCACAGGCCGGGGTACGACGTGTGGTGTTCACTTCTTCCATTGGAGCAGTTTACATGGACCC
AAAGAGGGATTATGATTCAGTTGTTGATGAGAGCTGCTGGAGCAACCTTGATTTTTGTAAGGATACCA
AGAACTGGTATTGCTATGGGAAGACGGTGGCAGAGCAGGCTGCATGGGAGAGGGCAAAGGAGAAGG
GACTAGACCTAGTAGTGGTAAATCCATGTGTAGTTCTGGGCCCACTTCTGCAATCTGCCATGAATGCC
AGCACACTCCACATACTCAAATACTTGACCGGTTCAGCTAAAACTTATGCCAACTCAGTTCAGGCGTA
TGTCCACGTCAGAGACGTTGCAGAGGCCCACATCTTGGTTTTTGAGAACACGTCAGCTTCGGGGCGAT
ATTTGTGTGCTGAGAGTGTTCTCCACCGTGGTGACGTGGTCGACTTGCTGGCAAAGATGTTCCCACAG
TACCCTATTCCTACCAGGTGCTCGGATGAAGTGAAACCAAGGATGAAGCCTTACAACTTTTCCAATCA
AAAGCTAAAGGACATTGGATTGGACTTCACACCAGTTAAAGAATGCTTGTATGAAACAGTGCTATGC
TTGCAGGAGAAGGGCCATATTCCCAAGTGA

>*Ch*CCR2 (MZ833462), Number of Nucleotides: 966 bp

ATGGAGGAAGTCTGTGTGACAGGAGGCACAGGATTTATTGCTGCTTATCTCATTCGTGGTCTGCTCGA
AAAGGGCAACACAGTTCGCACCACTGTTCGCAATCCAGATAATGTGGAGAAGATTGGGTATCTATGG
GATCTGCCTGGTGCAAAGGAGAGACTGAAAATAATGAAAGCAGACCTGCTAGAAGAAAGTAGTTTTG
ATGATGCAGTGGATGGTGTAGATGGGGTCTTTCATACTGCTTCCCCAGTCTTAGTACCATATGATCAC
AACATCAAGGAGACCTTAATTGATCCTTGTGTGAATGGCACCATCAATGTCCTCAAGTCTTGTTCAAG
ATCACCCTCTGTGAAGAGGGTGGTGCTTACATCCTCCTGCTCATCTATCAGATATGACTACAATACCC
AGCAGATTTCACATTTAGACGAGTCTCACTGGAGCAATCCAGATTACTGCAAACAGTTCAATCTTTGG
TATGCATTTGCAAAAACTATTGCTGAGAAAGAGGCCTGGAAGTTTGCAGAGGAGCAGGGGCTCAATC
TGGTTGTGCTAAATCCATCATTTGTTATTGGACCGCTGCTGGCTCCTGAACCTACCAGTACTCTTTACT
TAATCCTAAACATATTGAAAGGAGGAAATAACAATACATACCCAAATATGAGGTTGGGTTTTGTTCAT
ATAGATGATGTAATAACAGCTCATATACTGGCAATGGAGGTACCATCAGCATCAGGAAGAATTATCT
GCTCTGGTGACGTTGCTCACTGGGGAGAAATTGTGAAGATGCTCAAGGAAAAATATCCAATGTACCC
AATTGCAGACAAATGTGGAGCAGAACAAGGCAATGACACACCTCACACCATGAACACTAGTAAGATT
AGAAGCCTTGGATTTGGCAGCTTTAAGAGCCTTGAGGAGATGTTTGAAGACTGCATTAGAAGTTTCCA
AGAGAAGGGGCTGCTGTAG

>*Ch*CCR3 (MZ833463), Number of Nucleotides: 987 bp

ATGCAAGCAAAAAAGCATCAATCCAGTTTCGAAACCGTGTGCGTGACAGGGGCAGGAGGGTTTTTGG
CCTCTTGGCTAGTCAAATTACTTCTCCAAAGCGGTTATAACGTCCGCGGCACTGTCCGAGACCCCGGT
GATGCAAAGAATGCACATTTGAAAAACATTGAAGGGGCAAAGGACAGGCTAAAACTTCTAAAGGCG
GACTTATTGGACTATGACTTACTTTTCACAGCAATTGATGGGTGTATAGGAGTTTTCCATACGGCATG
CCCAGTTCCTTCTCACAGATTGTCCAACCCAGAGGCAGAGATGGTAAATCCTGCTGTAAAGGGAACTT
TGAATGTGTTGAAGGTGTGTTCAGTGGCAAAGGTGAAGCGTGTAATTATGACGTCGTCTTTAGGTGCT
ATTCATGTAAACCCAAATCGACCTATGGGAACACTTGTAGATGAGAGCTGCTGGTCAGATCCTGAATA
CTGTAAGACAACTCAGAGTTGGTATCTCATGTCAAAAACAGTCTCAGAACGGGACGCTTGGAACTGT
GCAAAGGAGAATGGCCTTGACCTTGTCACCATTTGTCCATCTTGGATTTTTGGGCCTATGTTACAGCC
AATTATTAATGCCAGTTCTCTAGGTCTTATCAAGCTTCTGACTGGTGATTTAGAGAGACGCGATAACA
AAGTACGCCATATTGTAGATGTCAGAGATTGTGCAAAAGCACACCTACTTGCTTATGAAGCTCCTTCT
GCAGCAGGACGATATCTATGCTCAGCTCATACAATTAATACTAAGGAACTGGTTGACATTCTTCATAG
GCTTTACCCACAATATACTTATCCCAAGGATTTTGTGGATGTAGAGTTGAACAGATCAGGAGTTGAGC
GATTATCAAATAGTAAACTGCAAGAATTGGGTTTGGAGTTTATGAAGTTGGAAAAGACACTTGTTGAC
TCTGTCGAATGCTTCTTCCAGAAGAATATTTTGAAGTGA

>*Ch*CCR4 (MZ833464), Number of Nucleotides: 903 bp

ATGCCTGTGTTTTCTCACTTGGATGAAAGAAATGGCAGAGGGAAAGCTGTGTGTGTAATGGATGCCTC
CAGCTATGTAGGCCTCTGGATTGTGAAGGGCCTTCTACACAGGGGTTACACTGTGCATGCCACTGTCC
AGAATAGAGGTGAAGTTGAATCTTTAATGAAATTGAGTGGGGAGCGGTTGAAGATCTTGTATGCAGA
TATGTTGGACTATCACAGTATTGTGGATGCACTCAGTGGATGTTCTGGCCTTTTCTACACATTTGATCC
CGCTCAATATGATGAGATGATGGCTGAAGTTGAGGTGAGGGCAGCCCACAATGTGTTGGAAGCCTGT
GCTCATACTGAAACAATTGAAAAAGTGGTATTCACGTCTTCATTGGCTGCAGTAGTTTGGGGAGATGA
CAGAAATTCAATTTCTGACCTCCATGAAAAACATTGGAGCGATGCAATCCTTTGCAGAAAATTGAAGC
TGTGGTATGCACTGGCCAAGACACTGTCAGAAAAGACAGCATGGGCCCTGGCAATGGACAGAGGAGT
GAATATGGTGACAATCAATGCAGGTCTAGTTGTTGGGCCAGGCTCTGCATACAAAACCTCAGGATCC
ACCATTGCATATCTTAAAGGTGCTGCACAGATGTATGAGAATGGAATGTTAGCAAGTGCAGATGTAA
AATTTGTAGCAGAAGCCCACATCTCTGCATTTGAGAACCCCTCTGCATATGGGAGGTATATATGTTTC
AACCAGATAGTTAACAATGCACAGAATGCTGACAATCTTGCTGAGAGTTTGAGACCACTCATACCATT
TCCTGACAGATGTGAAGATTCAAGTGTATACCAACAACGATTGAACAACAAGAAGCTTAGTGGTCTC
ATGGGTGGATACACAAAGCAGTAG

>*Ch*CCR5 (MZ833465), Number of Nucleotides: 969 bp

ATGGATAACAAAGAGACTGTATGTGTGACTGGAGCTAGTGGGTTCATTGGTTCATGGGTGCTTCGTTT
GCTCTTGGAGGCTGGTTACTCTGTGCATGCCTCTGTTCAGGACTTGGGAAATGAGAAGGAGACCAAGC
ACTTGGAATCCATGGAAGGAGCCAAAGAGAGGCTGAAATTGTTTCAACTAGACCTGCTGGATTACGA
TTCTATTGAGGCTGCCATTAGTGGGTGTGTGGGTGTTTTTCACCTAGCATCCCCCTGCATTGTGGATGA
AGTCAAGGACCCACAAGCTCAGCTCTTGGATCCTGCTATTAAGGGTACTATGAATGTCTTGCAAGCTG
CCCACAAAGCAAAAGTGAAACGTGTGGTGTTAACCTCATCGATCTCTGCCATCACTCCCTGTCCAAAC
TGGCCTGCTGATGTGCCCAAGGATGAGAACTGTTGGACTGATCTTGATTACTGCGAAAAAAATGGGA
TTTGGTATCCTGCATCAAAAACACTTGCTGAAAAAGCGGCTTGGGAATTTGCCAAGGAGACAGGTTTG
GATGTGGTTGTCATCAATCCAGGTACCGTGCTGGGACCTATTTTACCTCCAACCACGAATGCTAGTAT
GGCCATGTTTCTCCAACTTCTTCAAGGCTGCACGGATGGCTATGCAAACTTTTATATGGGTTGTGTACA
TGTAAAAGATGTTGCAAAAGGACATATCTTACTATATGAAACTCCATCAGCATCTGGAAGGCACCTTT
GTGTGGAAGCAACAACTCATTGGAGTGACTTTGCTGATATAGTAGCAAAACTCTATCCTGAATATAAA
ATACACAGGTTCACAGAAGTGACACAACCTGGTCTGTTAAGAGTTCAAAATGCTTCAAAAAAATTGG
TTGACCTAGGGCTAATTTTTACACCCATGGAACAAGTTATCAAAGATTCAGTTTCCAGTCTAAAAGAA
AAAGGCTTTTTGAATTGA

>*Ch*CCR6 (MZ833466), Number of Nucleotides: 981 bp

ATGACAGGGAAGCTAGTTTGTGTTACAGGGGCATCTGGCTTCATTGCCTCGTGGCTGGTCAAGCTTCT
CCTCGATAGAGGCTATGTTGTCCGAGCTACTGTTCGTGATTTAGGTAATCCTGACAAGACAAGGCATT
TATCTGCACTAGAAGGAGCCAAAGAAAGACTTCAGCTGGTGAAAGCAAATTTACTTGAAATGGGATC
CTTTGATGCAGCTATAGATTCATGTGAAGGTGTTTTTCACACTGCATCTCCATTTTTCCATGGTGTCCA
AGATCCCCAGGCTGATTTGCTTGAGCCAGCAGTTAAGGGTACTCTCAATGTTCTTGACACATGTGCCA
AAACATCCTCTGTGAAGAGGGTGGTAGTAACATCCTCTATGGCAGCAGTAGCATACAACAGGAGGAG
TCGAGCTCCAGATGTAGTGGTTGATGAAACTTGGTTTAGTGATCCAGAATATTGTAGGGAAGTGCAGG
GCTGGTATCTGCTATCAAAGACATTGGCAGAAGAAACTGTATGGAAAATGGCAAAGGAGAAGGGCAT
TGATATTGTCACAATAAATCCAGGAATGGTGATTGGCCCTCTCTTGCAGCCTACTCTGAATACAAGCT
GTGCTGCAATTTTACAACTGATGAATGGTTCTTCCACATTTCCAAATTTAACGTTTGGATGGGTGAGTG
TTAAGGATGTCGCTGAAGCACATATTCTCGCCTTTGAAGTACCATTTGCAAAAGGAAGACATCTTTTA
GTCGACGAGGTTGCTCACTACTCAGAAATAGTGAAGATCTTGAGGAAGCTTTATCCAGGCTGTTCACT
TCCGACAAAGTCTGCAGATGATAAACCATTTCCCCCTACCTTTAAAGTATCTAAAGAAAAGGTAGAG
AAATTGGGTTTGAAGTACACTCCTATTGAAGAGGCTCTCAGACAGACTGTTGAAAGCTTGATGGAAA
AGAAGATGCTTCAGTTGGAGAGTAATATGTGA

>*Ch*CCR7 (MZ833467), Number of Nucleotides: 999 bp

ATGGAAATAGAAGCAAAGAAAAGGATGAATCAATCGAGGGTCGAAACTGTGTGTGTTACAGGTGCA
GGAGGGTTTGTGGCATCATGGATCGTTAAATTACTTCTCCAACGTGGTTATAATGTCCGCGGCACTGT
CCGAGACCATGGTGATGCAAAGAATGAACATTTGAAGAACCTAGAAGGGGCAAAGGAGGGGCTAAA
TATTCTAAAGGCAGACTTGTTGGACTATGACTCAATTTTAGCAGCAATAGATGGGTGTAGTGGAGTTT
TCCACACGGCATGCCCGGTTCCTTCTGGCAGAGTGTCTGACCCAGAGGCAGAGATTGTTAATCCTGCT
GTAAAGGGAACTTTGAATGTATTGAAGGCATGCTCTGTGGCAGGGGTGAAGCGTGTAATTATGACTTC
ATCTTTAGCTGCCGTTTATCTAATCCAAAACCGACATATCAATAAACTAGTAGATGAGAGCTGCTGGT
CAGATCCAGAATATTGTAAGACAACACAGAATTGGTATTTTATGTCAAAAACAGTTTCAGAACAGGA
GGCCTGGACATTTTCAAAGGAGAATGGCCTTGACCTTGTCACTGTTTGTCCATCTTTGACTTTGGGACC
TATGTTACAGCCAACAACCAATACTAGTTCTCAATATTTTAGCCAGCTTCTGACCGGCAATTCAGAGA
GATGTGAAAATTTGGTACGCAATATTGTAGATGTTGGAGATTGTGCAAAAGCCCACCTACTTGCTTAT

GAAACTCCTGCTGCAGCAGGACGATATCTGTGTACAGGTCATACAATCAAAAACAAGGAACTGGTTG
ACATCCTTCAAAGGCTTTACCCACAATATACCTACCCCAAAGTCTTTGTGGATGCAGAGTCTTATTTAT
CAGGACCTGAACGACTGTCTGATGGTAAACTACGAGATTTGGGTTTGGAGTACGTGAAGTTGGAAAA
AACACTTGTTAACACTGTTGAATGCCTCCAAAAGAGAAGTATTTTGAAGTGA

>*Ch*CCR8 (MZ833468), Number of Nucleotides: 1044 bp

ATGGCGATGGTGAAGTCTGCCGAGTGCCATTTACCGCAAGGGTGGGATGCGAAGGGATTCGATACCG
GAAAAACGGATGTGTTGCCTAAGGTAGTCTGTGTCACAAATGGTAGCTCGTATTTGGGGCTCTGGATT
GCCATACATTTGCTGCGACGAGGATACCTGGTGCGTCTTACAGTCGACTCCACAGACGAGCTTCACAA
GCTGAGGGAGATGGAGGAATTTGCAATGCATTCGAATCGAGTGATGGGAGTTGTGGGCAATATTGCT
GTTGACGACGCTTCGACGCTCTCTCAAATGTTTGATGGGTGTTATGGAGTGTTCCATACTTCCTCTTTT
ATTGATCCTCATGGTGTCTCTGGTTTCACTGAGCATATGGTAAATTTGGAAGTAAAAGGAGCGGAGAA
GGTGGTGGAAGCCTGCTCAATAACATCTTCTGTAAGAAGGCTGGTATTTACATCTTCGTTGGCAGCTT
GCATCTGGCACACAGAAGGGACTGCCTCTTCTTTTGTGGTGGATGAGAGATGCTGGAGTGACCCAAAT
ATATGCAGAGAGAAAAAGTTATGGTTTGCACTGTCAAAGACAATGGCAGAGAAAGCAGCCTGGAGA
GCGGCTGGAGAAAGAGGAGGCCTCAACATGGTGACCATCTGCCCTGCCCTCCTCACCGGCCATGCCTT
TTCTTCTTCCAATTCCACATCCTCCATTGCTTACCTCAAAGGTGGAAGAGAAATGTATGAGAAAGGTG
TTCTGGCAACAGTAGATGTGCGCAGGGCAGCAGAGGCCCATATCTGTGTATATGAAGAAATGGGTGG
TGGAGGAGGAGCGAGTGGAAGATATATTTGCTTCGATACGCTCATAAATACATCATCAGATGCCCTG
GATTTGGAGAAGAAATTGAAGATGAACATGGGGTTTTCTCAATTAATTGACAGAGATAATTATAACG
AGAATGGGACATTTTTGTTTGGGGTAGATAGGATAGTGAATGCAAAGCTGGCCAGAGTAATTGCGGC
ATCCGTCCGCATCAATTCCTGCAGCCGATGA

>*Ch*CCR9 (MZ833469), Number of Nucleotides: 966 bp

ATGGAGGAGAAGAGGGTGTGTGTGACAGGAGGGAGTGGATTCATAGGATCATGGCTCGTTCGTAAAT
TACTGCAGCGTGGATATATTGTTCATGCAACTCTCCAAAATCTAGCAGACCCAAATGAAAAAAACCAT
CTTCAAGCCTTGGAAGGAGCCAACGAAAGACTCAAATTATTCCAACTCGATGTTTTGGACTATGATTC
AGTGGCTGCTGCAGTCACTGGCTGCCAGGGTGTCTTCCACTTAGCCTCTCCCTGCACACTTGAAGAAA
TCAAGGATCCCCAGGTTGAGCTGTTGGATCCAGCAGTTAAAGGGACGTTGAACGCACTTGAAGCATG
TCGAAGAGCCAATGTAAAGCGTATTGTTATCACCTCTTCGATTTCTTCCTTGGTTCCCAATCCTACATG
GCCTCCCAATACTGTGATCAATGAGAATTCTTGGACGGATATCGAATACTGTAAGCGTGATAAGATTT
GGTATCCAGTGTCAAAGACTCTTGCAGAGAAAGCAGCATGGGAATATGTCCACACCCACAAAATGGA
TGTGGTAGCTGTTCATCCTTCCACTTGTTTGGGACCTCTTTTACAACCCAAATTGAATGCCAGCTCTGC
TGTTTTACTGAAGCTATTACAAGGCTCACAAGATACTCAGTCAAATTATTGGTTAGGATGCGTACACG
TGGAAGATGTTGCAAAAGCACATATTTTACTTTATGAAACACCATCAGCATCTGACCGTCATCTTTGC
ACCAATGGTATAATTCATTTCAGTGATTTTGCAGAGGTAGTAGCCAAGATCTGCCCTGAATACCATAT
CCACAGGTTTAAAGAAGAAACCCAGCCCGGATTATTGAGATGTAAAGATGCAGCCAAAAAGCTCATT
GACTTAGGACTTTGTTTCACTCCCATAGAGAAGGCCATTCAAGAATCCGCTGCTAGCATGAAAGAGA
GAGGGTTTCTTCATTAG

>*Ch*CCR10 (MZ833470), Number of Nucleotides: 975 bp

ATGCAGATGGCGGTGACAGAAGCTGTGTGTGTTACGGGAGCCAGTGGCTTCATTGGCTCATGGGTTGT
TCGTTTGCTCTTAGAACGGGGTTATACCGTGCACGCTACCGTTCAAAGTCTTGAAAATTGGAAGGAAA
CGAAGCACTTGGAGGCAATGGAAGGAGCCAAAGAAAGGCTGAAGTTATTTCAAATGGATTTGCTGGA
TTACGAGTCAATTGAGGCGGCCATTAATGGCTGCACGGGTGTTTTTCATTTGGCAATGCCCAATACTA
TTGAGGCCGTAGAGGATCCAGAAAAGCAACTCATGGATCCTGGCATCAAGGGTACCTTGAACGTATT
GGAAGCAGCCCACAAAGCCAAAGTAAAGCGCGTGGTGCTGACCTCTTCTGTATCCGCCTTCATTCCCA
ACCCCAAATGGCCTGCTGAAACCCCCTTGGACGAGACCTGCTGGACTGACCTCGACTACTGCAAACA
AAATGGGATTTGGTACCCTGTGGCGAAAACACAAGCTGAAAAGGCAGCTTGGGAATTTAGTAAAGAG
AAAGGGTTGGATGTAGTTGCTATTAACCCGGGGACTGCTCTGGGCACCATTCTGCCTCCTGATTTCAA
TGCCAGCCTGGCCATGATCGTACGCCTTGTCAATGGTAACAAAGAGGAGTATGGTAACTTTTATATGG
GTTGTGTTCATGTAAGGGATATTGCAAAATCTCAAGTACTACTATATGAAACACCCTCTGCATCAGGA
CGTCACCTGTGTGTTGAATCTATAACTCATTGGAGCGACTTTGCAGAGTTGACAGCGAAACTATATCC
AGAATATGATGTGCCCAAGTTCACTGGTGTTACTCAACCTGGGTTGGCGCGTGTGAAGAATGCACCAA
AAAAGTTGATTGATCTTGGGGTCGAATTTATGCCTATGGAGCAGATCATTAAAGACAGTGTTTCATCA
TTGAAAGAGAAAGGTTTTCTCAAGTAA

>*Ch*4CL1 (MZ833459), Number of Nucleotides: 1515 bp

ATGGCGGGATGCGAAGGGGAAAATGCCGAGTTTCTGTTTAGATCCAAACTTCCGGACATCGAGATAC
CTAATCATCTTCCCCTGCATAAATATTGTTTCGAGCAGCTGGCGGAATTTGCAGAGAGGCCGTGCCTC
ATCGACGGAGCGAGCGGGAAAATATTCACCTACGGTGAGGTGGATTTGATATCCCGCAAGGCGGCTG
CCGGTCTGGCCAAGTTGGGACTCCAACAAGGGCAAGTGGTGATGCTCTTGCTTCAGAATTGCGTGGA

ATTCGCATTTGTGTTCTTGGGGGCGTCCATCAGGGGCGCCATTGCCACAACGGCCAATCCATTTTATA
CTCCGGGTGAGATCGCCAAGCAGGTGAAGGCCTCGGGCGCCCGCATTATCGTTACCCAGGCGGCATA
CGTTGACAAACTGGGCGATCTGACCCGAGAGGATTGTGATTTGACTGTTATCACCATTGACAACCCTC
CCGAGGGGTGCAACCCCATCTCCGTCCTAACGGAGGCGGACGAGAGCGAGTGCCCCTGTGTGGACAT
CCACCCGGACGATGTGGTGGCCTTGCCTTACTCTTCGGGCACCACAGGCCTCCCCAAGGGCGTCATGC
TCACGCACAAGGGATTGGTGTCCAGCGTCGCTCAGCAGGTCGACGGCGACAACCCCAATCTCTATTTC
CATTCGGAAGACGTGATCCTGTGCGTGCTTCCTTTGTTTCATATCTACTCTCTGAACTCGGTGCTCCTC
TGCGCTCTGAGAGTGGGGGCCGCCATGCTCGTCATGCAGAAATTCAACATCGTGGCGTTGCTGGAGCT
GATCGACAAGTACAAGGTCACAGTGGCTCCGTTCGTTCCTCCTATTGTTCTGGAGATCACGAAAAACC
CAATCGTTGCTAACTACGACGTTTCTTCAATACGGCTAATTATCTCGGGCGCTGCGCCCCTAGGGAAG
GAGCTGGAGGACGCCCTCAGGGCGCGCTTCCCGGGGGCCAAGTTCGGACAGGGATACGGCATGACGG
AAGCGGGGCCGGTGTTAGCAATGAATCTTGCATTTGCAAAGGAGCCCTTCCCTGTAAAGTCAGGCTCC
TGCGGAACGGTCGTTCGGAACGCCCAACTAAAGATCATAGACACGGAGACCGGTACGTCTCTACCAC
ATAACCAAGCCGGTGAAATCTGCATCCGTGGACCCCAAATTATGAAAGGGTATCTGAACGACCCAGA
AGCCACGGTACGAACCATCGACAAAGAGGGGTGGCTGCATACCGGCGACGTTGGCTATATTGACGAC
AACGAAGAAATTTTCATCGTGGACAGAGTGAAAGAGCTTATAAAATATAAAGGCTATCAGGTCGCGC
CTGCTGAACTGGAAGCCATTCTGGTAAACCACCGATCAATTGTGGACGCAGCAGTGGTGCCTCAGAA
GAATGAAGCGGCAGGAGAGGTTCCGGTGGCGTTTGTCGTCAAGTCTGAGGGAGCGGACATCAGCGAG
CAAGAAATAAAAGATTTCATTGCAAAACAAGTGGTTTTCTACAAGAGATTGTACAAGGTTTATTTCGT
GGATTCAATTCCCAAGTCTCCCTCCGGCAAGATCTTGAGAAAGGATTTGAGAGCCAGACTTGTAGTAG
TTACTTAA

>*Ch*4CL2 (MZ833460), Number of Nucleotides: 1671 bp

ATGATTGAGGTGGACGAGCTGCCTCCTCCTCCTCCTCCTCCTCCTGCTCCCATTATTAATGATAAT
TCGTATGTGTACAGATCAAAGCTCCCGGACATAGATATCCCCTTCCATCTTCCTCTACACGCCTATTGT
TTGGAGAATGCAAATGAATTCGGAGACAGGGTGTGCTTCATACAAGCCTCCTCCGGTAAGGTCTATAC
TTACGGAGATGTGGATTTAATATCGAGGAAAGCGGCCTCCGGGCTGGCTAAATTGGGCATTGGTAAA
GGGGATGTGGTGATGCTGCTGCTCCACAATTGCCCCGAGTTTGCCTTTCTATTCCTGGGAACCTCCATC
GCAGGCGGCGTCACCACCACCGCTAATCCCTTCTACACCCCGGCGGACATTGGCAAGCAGCTGCGTG
CCTCCAACGCCCGCCTCCTCGTCACGCACGCCGCCTACGTTGAGAAGCTTATGGACTTCCTCCCCCAC
CTGCAGGTGCTTACCATCGACACCCCGCCCGAGGGGTGCGCCCACATTTCGGTGGTGTTGGAAGCAG
AAGAGCAGGAGTGCCCCGCCGTGGACATCCAGCCAGATGATGCGGTGGCGTTGCCTTACTCGTCCGG
AACCACGGGGCTGCCCAAGGGAGTGATGTTGACACACAAGGGCCTCCTTTCGAGCGTGGCCCAGCAG
GTGGACGGCCTCAACCCCAATCTCTATCTGCACTCGGAGGACGTGGTGCTCTGCGTCCTCCCGCTTTTT
CATATCTATTCGCTCAACTCTGTGCTGCTCTGCTCCTTGAGGGCAGGCTCCACCATCCTGCTCATGCAC
AAGTTTGAAATCGCCACTCTGCTCCACCTCATTCACACCTACAAGGTGACCGTGGCGCCTCTCGTCCC
TCCCATTGTGCTTGCCATTGCCAAGAATCCTATGCTCCATCACCACGACCTCACCTCCGTTCGGATTCT
CCTCTCTGGGGCCGCGCCTCTCGGGAAGGACCTCGAACACGCGCTCATCACTCGACTTCCGTCAGCTA
CGTTTGGACAGGGATACGGTATGACAGAAGCAGGGCCCGTGTTGTCAATGAGCCTCGCATTCGCCAA
GGAGCCCTTCCCTGTCAAGTCCGGCTCCTGTGGGACGGTTGTTCGCAATGCGCAGATGAAGATCGTTG
ACCCGGACACCGGAGAGTCCCTTCCCTGTAACAAACATGGGGAAATATGCATCCGAGGACCCCAAAT
TATGAAAGGTTATCTCAACGATCCGGAGGCCACAGCCAGAACTATTGACAAAGATGGGTGGCTGCAT
ACCGGAGACATTGGATATATAGACGAGGACGATGAAGTTTTCATCGTGGACAGAGTGAAAGAACTCA
TCAAATACAAGGGATTCCAGGTGCCTCCGGCTGAATTAGAATCAATTCTCATCACCCATCCATCAATC
GCAGACGCGGCAGTTGTACCTCAAAAGGATGAAGCGGCAGGAGAGGTGCCAGTAGCTTTTGTCGTCA
GATCCAACGGATTTCATCTCACAGAAGATGAAATTAAGCAATTCGTGGGTAAACAGGTGGTATTTTAC
AAAAAACTGCATAAGGTCTATTTCATCCATGCAATACCCAAGTCTCCTTCTGGAAAAATTCTTCGAAA
GAATTTGAGAGCCAAGCTTGTTGCCCAAGTAACGGCCAAATGA

>*Ch*DBR1 (MZ833471), Number of Nucleotides: 1047 bp

ATGGAAACGAAAGAAGTGATCAATAGACAAATTACCTTCAAGGATTACCTTGTCGGTGAACCTAAGG
AGAGTGATTTTATGCTGAAGAGTTGTACACTTTCCTTAAAACTAGAAGAAGGATCTCAAGACGTGCTT
GTAAAGAACTTATATTTTTCTATCGATCCATACCAACGAGCTCGCATGAAATACTACAGCAGTTCACA
CAAAGCTCCTTTTGCATCTCCTCTCACTCCGGGTTCGGTCATTGATACTTATGGTGTGGCCAAAGTCAT
TGTCTCAAATCACCCAGAGTTCAAAGAAGGTGACTTGGTAGTTGGTGTGATTGGAACAGAAGACTAT
AGCATCGTACCTGGAGCAATCCTATTAACAAAGATCAAGTTCACAGACATACCTCTTTCATACTATGT
TGGTGTTCTAGGCTTCACTGGGTTTACTGCTTATGCTGGATTGTATAATGTATGCTCTCCTCAAAAAGG
TGACCGGGTTTTCATCTCTGCAGCCTGTGGAGCTGTTGGCCATCTAGCTGGACAATTTGCAAAGCTGC
AAGGATGCTATGTTGTGGGTAGTTCTGGTAGCAAGCAAAAGGTTGATTTACTAAAAAATAAATTTGGC
TTTGACGACGCTTTCAATTACAAGGAAGAACCTGACTTGAATGCTGCCCTCAAGAGGTGTTTTCCCAG

TGGAATAGATATATACTTTGAGAATGTGGGTGGAGACCTGCTTGAAGCTGTACTAGAGAATATGAAT
GTCCATGGCCGAATCACTGTCTGTGGGGCCATCTCACAGTATAACAAAGACAAACCAGATGGCGTAA
CTGACCTCATGACATTGGTCTACAAGCGCATCAGAATGCAGGGTTTCCTTGCAGCAGATTACTTAAAT
TCTTTCCCTGAGTTTGTTGAAAAAGCTAGAGATTATCTTCTTGAAGGAAAAATAAAATGCCTTGAAGA
CATTTCTGATGGTCTGGAAAGTGTACCTTCTGCTTTTGTTGGCTTATTCAATGGCCAAAATGTAGGCAA
GAAAGTTGTTCGCCTCCATCATCCATAG

>*Ch*DBR2 (MZ833472), Number of Nucleotides: 1032 bp

ATGGCAAGCAATGAAGTGAGTAATAAGCAAGTAATCTTGAAGGATTATGTTAATGGAGTGCCGAAGG
AGTCTGATATGTTGGTAAAATATACCAAACTTCCATTGCAATTGAAAGATGGATCACAAGATGTTGCT
GTAAAGAACTTGTATTTGTCATGTGATCCTTACATGGGAGGCCGCATGAAAAATATCACTTCTTACTT
CACAGCTTTTACTCCTGGATTGGTCATCAATGGTTATGGTGTCTCCAAAGTTATAATTTCAAATCATCC
AGAATTTAAACAGGGGGATTTAGTAGCTGGTTCTGTTGGATGGGAAGAGTACAGCATAATACCAGAG
GGAAGGGGCTTGCAAAAGATCATACATACAGATGTACCTCTTTCATACTATTTAGGAGTTTTAGGAAT
GCCTGGACACACAGCTTATGTTGGTTTTTATAAAGTGAGCTCCCCTAAACCAGGAGAAACGGTATACG
TCTCTGCAGCTGCTGGAGCAGTTGGTCAGCTTGTTGGCCAATTTGCCAAGTTGCACGGTTGTTATGTTG
TTGGGAGTGCTGGAAGCAAAGACAAGGTAGACGTGCTGAAAAATAGGCTAGGTTTTGATGATGCATT
CAATTACAAGGAGGAACCGGACTTAAATGCTGCTCTTAAGAGGTACCTCCCAAATGGGATAGACATC
TATTTTGAGAATGTGGGTGGTGATATGCTTGAAGCAGTTCTTGAAAACATGAACATCCATGGCCGTAT
TTCTCTTTGTGGAATGATTTCTCAATATAATTTGGAGGAGGCAAGAGGTATAAGGAACCTATCACAAC
TAATAATAAAGCGCATAAAAATGGAAGGTTTCATTGAACCTGATTTTTTGCATACATACCCAGAGTAT
GTTGAGGAGGTAAGAGGTTATATCAAAGAAGGGAAAATAGTTTATGTTGAAGATATTGCTGAGGGTC
TGGAGAATGCACCGTCCGCATTTGTTGGTCTCTTTCACGGTAAAAACATTGGAAAACAGGTTGTTCGC
ATATGTGACGAATGA

>*Ch*DBR3 (MZ833473), Number of Nucleotides: 1035 bp

ATGGCAAGCAATGAAGTGAGTAATAAGCAAGTAATATTGAAGGATTATGTTACTGGAGTGCCGAAGG
AGTCTGATATGTTGGTAAAATATACCAAAATTTCATTGCAATTGAAAGATGGATCACAAGATGTTGCT
GTAAAGAATTTGTATTTGTCATGTGATCCTTACATGCGAGGCCGCATGAAAAATATCACTTCTTACTTC
ACAGCTCCTTTTACTCCTGGATTGGTCATCAATGGTTATGGTGTCTCCAAAGTTATAATTTCAAATCAT
CCAGAATTTAAAAAGGGGGATTTAGTAGCTGGTTTTGCTGGATGGGAAGAGTACAGCATAATACCAG
AGGGAAGGGGCTTACAGAAGATCATACATACAGATGTACCTCTTTCGTACTATTTAGGAGTTTTAGGA
ATGCCTGGACACACAGCTTATGTTGGGTTTTTTAAAGTGAGCTCACCTAAACCAGGAGAAACGGTATA
CATCTCTGCAGCTGCTGGAGCAGTTGGTCAGCTTGTTGGCCAATTTGCCAAGTTGCACGGTTGTTATGT
TGTTGGGAGTGCTGGAAGCAAAGACAAGGTAGACATGCTGAAAAATAAACTAGGTTTTGATGATGCA
TTCAATTACAAGGAGGAACTGGACTTAAATGCTGCTCTTAAGAGGTACCTTCCAAATGGGATAGACAT
CTATTTTGAGAATGTGGGTGGTGATATGCTTGAAGCAGTTCTTGAAAACATGAATATCCATGGCCGTA
TTGCTGTTTGTGGAATGATTTCACAATATAATTTGGAGGAGGAAAGAGGTATAAGGAACCTATCACA
ACTAATAATAAAGCGCATAAAAATGGAAGGTTTCCTTGAACCTGATTTTTTGCATACATACCCAGAGT
TTGTTGAGGAGGTAAGAGGTTATATCAAAGAAGGGAAAATAGTTTATGTTGAAGATATTGCTGAGGG
TCTGGAGAATGCACCGTCCGCATTTGTTGGTCTCTTTCACGGTAAAAACATTGGAAAACAGGTTGTTC
GCATATGTGACGAATGA

>*Ch*DBR4 (MZ833474), Number of Nucleotides: 1032 bp

ATGGGAAAAGAGGTTGTGAACAAGCAAGTAATCTTCAAAGATTACATTATAGGGGTACCAAACGAAT
CTGATATGGAGCTCAGGTCCAGCACTATTGCCTTGGAGGTGAAGGAGGGGACACATGATGTTCTTGTC
AAGAATTTGTACCTGTCCTGTGATCCTTATATGCGAGGGAGAATGAGAAACTACCACAATTCTTACAT
ACCTCCCTTCACCCCTGGAACAGTTGTACAAGGTTATGTGGTGGCTAAAATGATACTTTCAGATCACC
CAGATTTTGAAGAGGGTGACTTAGTATCCGGTATTTCTGGATGGGAAGAGTACAGCATAATCCCGAA
AGGAAAGAACTTAACAAAGATCACATATACTGATTTACCCCTTTCATACTTTGTGGGGATTTTAGGAA
TGCCTGGATTAACAGCTTATGCTGGATTCTTTGAGGTTTGCACTCCTAAAAAGGGAGAGCATGTTTTT
GTCTCTGCAGCTTCTGGAGCAGTTGGCCAGCTTGTTGGACAATTTGCAAAGTTGATGGGTTGCTATGT
TGTCGGGAGTGCTGGTAGTAAGGAGAAGGTGGATTTACTGAAAAATAAGCTCGGCTTTGATGATGCT
TTCAACTATAAGGAGGAGCATCACTTGGATGCTACTCTGAAAAGGCATTTTCCAGAAGGAATCGACA
TCTATTTTGACAATGTGGGGGGTAGCATGCTTGATGCAGTACTTATAAATATGAAAGTCCATGGTCGT
ATCGCTGTTTGTGGTATGATATCTCAACATAGTTTTCAAGATGCACAAGCCATACATAACTTGAATAT
GGTAATCCCCAAGCGCATAAGCCTCCAAGGTTTCCTTCAGTCCGATTACTTGCACTTGTACCCGCAGT
TTCTTGAGAATGTGATAAATTATATCAAAGAAAGGAAGCTAGTTTACATTGAAGATATTGCTGTAGGC
CTCGAAAATGCACCTGCAGCTTTTGCAGGACTATTTTATGGTAAAAATGTTGGCAAGCAAGTAATTTG
TATTTCTGATGGATGA

>*Ch*DBR5 (MZ833475), Number of Nucleotides: 1047 bp

ATGGAAACGAAAGAAGTGATCAATAGACAAATTACCTTCAAGGATTACCTTGTTGGTGAACCTAAGG
AGAGTGATTTTATGCTGAAGAGTTGTACACTCTCCTTAAAACTTGAAGAAGGATCTCAGGATGTGCTT
GTACAGAACTTATATTTGTCTATCGATCCATATCAACGATCTCGCATGAAATACTATAGCAGTTCACA
CAAAACTCCTGCTGCATCTCGTCTCACTCCGGGTTTGGTCATTGATACCTATGGTGTGGCCAAAATCAT
TGTCTCAAATCACTCAGAGTTCAGAGAAGGTGACTTGGTAGTTGGTTCTATTGGAATGGAAGACTATA
GCATCATACCTGGAGCAATCCTATTAACAAAGATCAAGTTCACAGACATACCTCTTTCGTACTATGTT
GGTGTTTTAGGTATGACTGGGTTAACTGCTTATGCTGGATTGCATAATGTATGCTCTCCTCAAAAAGG
TGACCGAGTTTTCATCTCTGCAGCCTGTGGATCTGTTGGCCATCTAACTGGACAATTTGCAAAGCTGC
AAGGATGCTACGTTGTGGGTAGTTCTGGTAGCAAGCAAAAGGTTGATTTGCTAAAAAATAAGTTTGG
ATTTGATGACGCTTTCAATTACAAGGAAGAACCTGATTTGAATGCTACCCTCAAGAGGTGTTTTCCCA
CTGGAATAGATATATACTTTGAGAATGTGGGTGGAGAGCTGCTTGAAGCTGTACTAGAGAATATGAA
TGACCATGGCCGAATTGCTGTCTGTGGGGCCATCTCACAATATAACAAGGACAAACCAGAGGGTGTA
ATAGACCTCTTGAAGTTGGTCTACAAGCGCATCAGAATGCAGGGTTTTCTCGTAGTAGATTACTTGAA
TTCATTCCCTGAGTTTGTTGAAAAAGCTAGAGAGTATCTTCTTGAAGGGAAAATACAATGTCTTGAAG
ACATTTCTGATGGTCTGGAGAGTGTACCTTCTGCTTTTGTTGGCTTATTCCAAGGTCAAAATGTAGGCA
AGAAAGTTGTTCGCCTCCATCATCTGTAG

>*Ch*TyDC1 (MZ833476), Number of Nucleotides: 2073 bp

ATGGATAGCATGGATTTGTTGATGGATTTGTTGGGAGGTGTACACATGAATGGCAAGGATTTATCAGG
GTTTGGGCATATAAATGATAAGGATTTACCAAGAAATGATCATATAAATGGAAAAGATGAATACAAT
GGTCATGTGAATGGTAAAGAATTGCCCATCAATGGTCATATATATGGCAAGGATTCATGCAATGGTCA
TATAAATGGCAAGGATTCATGTAATGGTCATATAAATGGCAAGGATTCATGCAATGGTCATATAAATT
GTAAGAATTTACTAAGCAATGGTCAAGTAAATGGGAAGAATTTACCCAGAAATGACAAGGATGTGCA
AAGCAATGGTAGCATGATTGGAACTGATGAATGGAATGGTTATATGAATAGAAAGGATTTGCCAAGC
AATGGCCATATGAATGGTAAGGATGAATCTAGTGGTCGTATGAATGGTAAGAATGAATACAATGGTC
ATATGAATGAAAACAATTTGCCAAACAATGGCCTTATGAATGGAATGGATGAATTCAATGCTCATAT
GAATAGAAATGATGAATTCAATGATCATATGAATGGTAAGGATGAATGTGAGGGCTCAAACAGTCAC
AATTATCTTACCAAGTTATTAGACATTGAAGAGTTCAAAAAGCATGCTCACACTATGGTCGACTTTAT
TGCCGATTATTACACCAAAGTAGAGAGCTTCCCAGTGCTCAGCCAAGTAAAGCCTGGATATTTAAATT
CTTTGATACCTGATAGTGCTCCACAACATCCTGATTCAGTGGAGAGTCTTCTTTGTGATGTGAAAGAG
AAGATAATACCTGCATTGACACACTGGCAAAGTCCCAATTTTTATGCATATTTCCCTGCAGGGCTGAG
CACTGCATCATTTATGGGTGAGATGTTGAGTAGTGGCTTGGCTATAGTTAGTCTGACATGGATGACAT
CTCCTGCTGCCACTGAACTTGAAACCATAGTTTGTGATTGGCTTGGCAAAATGCTCAACCTGCCTGAG
CCTTTTCTCTCATCTTCAGGTAAAGGTGGTGGTGTTATTCAAGGATCAGCAAGTGAGGCAGTTTTGGTT
TCATTATTAGCTGCAAGGAACAAGGCTTTGAAGAAATTTGGCACAGATTCTGTTGACAAGCTTGTAGT
GTATATATCAGATCAGACACACTGCTCAATTGAAAAGGCCTGCAAAATTGCAGGAATATCTCAAAAC
AATCTTAGAGTGCTGCCAACAAGTTCTTTTACAAATTATGCACTTTTCCCTGAGATTGTTGCTCAATTT
ATAATGAAAGATATTAGTGATGGTTTGCACCCAACATTCTTATGTGCTACGATTGGAACGACTTCATC
AACAGCTGTGGATCCTCTTCAAAAACTTACTAAATTAGCTCGGGAACATGACATGTGGTTTCATGTGG
ATGGTGCTTATGGCGGTGCTGCTTGTATTTGCCCTGAATATCAACACTATCTTGTGGGGGTAGAGGAT
GCCGACTCTTTTGATATGAATCCCCATAAATGGCTTTTCACCAACTTCGATTGTTCTATACTTTGGGTA
AAGGACTCAAAGGCTTTGGTTGAAGCATTATCGATGAATCCAGAATATTTAACGAATAAGGTATCAC
AATCTCAACAAGTGATAGATTATAAGGATTGGCAAATACCACTTGGAAGTCGGTTTAGGTCACTAAA
GGTGTGGATGGTGTTGCGTCTCTATGGAATTTCAAAGTTACAAACATTTATTCGATATCAAATTTCTCT
TGCACAACATTTTGAAAGTCTTGTAATCAATGATTATAGATTTGAGGTCATGGCTCCTTGCACTTTTGC
ATTGGTTTGTTTTCGAATTTTACCTCTGCCAAATGATTCTGACAATGGACGTACTTTAAATTCACGTCT
ATTGGATGCTGTAAATGACAGTGGGCTTATTTACTTATCACATACGGTTTTATCAGGGAAATATGCTTT
GCGTTTTTGTGTGGGTGCGGCATTAACAAGCATACATCATGTAAATGGTGCATGGAAACTTCTTCAAG
ACCAGGCAACAATTTTATTAAGAGAGAATAATTAA

>*Ch*TyDC2 (MZ833477), Number of Nucleotides: 1550 bp

ATGGGAAGGAACTTTGCTGAAGGGAGTGAGAGTGGGAGTCTGAAGAGGGCATTTGACGTGGAGGAG
TTCAGAGAAAATGCCCACAAAATGGTAGATTTTATTGCCGACTACTACAGAGATATCCACAAGTTCCC
TGTTCGCAGCCAAGTTCAGCCTGGATATTTGCAAACTCTTATACCTGATAATGCTCCGGAATATCCTG
AAACATTGGAGAGTGTTCTAACTGATGTACAGAACAAGATCATACCTGGGATCACCCATTGGCAGAG
CCCTAACTTTTTTGCCTTCTATCCTTCAAATAGCAGTACTGCTGGGCTTCTAGGAGAAATGCTCAGTGG
TGGTTTTAATATTGTTGGTTTTAGCTGGATAACATCTCCTGCTGCAACTGAGCTTGAAGCTATAGTTCT
GAACTGGCTCGGAAAAATGTTAAAGCTTCCCGAGGCTTTTCTTTTACCTGGTAAAGGTGGTGGTGTAA
TTCAAGGTACAGCAAGTGAAGCTGTTTTGGTTGCATTGTTAGCAGCACGGGACAAAGCATTGAAGGG
GGATGGTGTCAAATCACTTGAAAAGCTTGTTGTATATGCATCAGATCAAACACATTCTGCATTGCAGA

AAGCTTGTATGATTGCAGGGATATGCAAATCAAATTTTAGAGTATTGCCGACCAATGCTTCTACAAAC
TATGCACTTTCTCCTGATGTTCTTCAAAAAGCAATTTTAGATGATATATATGCTGGATTTTGTCCCTTTT
TTCTATGTGCAACGGTTGGTACCACTTCATCAACGGCTGTGGATTCTTTGAAAGAAATTGGGAAAATA
TCTATGAATCATGGTATGTGGTTCCATGTTGATGCTGCATATGCTGGTAGTGCCTGTATTTGTCCTGAA
TATCAACACTATCTTGATGGTGTGGAGGCAGCAGATTCTTTCAATATGAATGCCCATAAGTGGCTTCT
TACAAATTTTGATTGCTCTGCTCTTTGGGTTAAGGACTCAAATGCTTTGGTTGAAGCTCTCTCTACACA
ACCAGAATTTTTGAGGAACAAGGCTTCGGAATCCAAGCTAGTAGTTGACTATAAAGATTGGCAAATA
CCCCTTGGCCGTCGATTTAGGTCCTTAAAGTTGTGGATGGTTCTGCGCCTGTATGGTGTGTCAAATTTA
CAAACTTATATTCGGAACCATATTGATCTTGCCAAGCATTTTGAGGAACTCATAAGAGAAGACTCCAG
ATTTGAGATTATGACACCCCGCACATTTGCTTTGGTGTGCTTCCGTGTTTTGCCTCCTCTAGATGATCC
AGACAATGGCTTCACTTTGAACTCACAATTGTTGGATGTTGTAAATAGTAGCGGACACATTTATTTAT
CACATACGGTTCTTGCTGGGATATACACATTGCGTTTTGTTGTGGGAGCCCCATTAACTCAAGAAGAA
CATGTCAAATCTGCATGGAAAATACTCCAAGACAAAGCATCTATTCTTCTAAAGACTTATGCATTACT
CAAGTGA

>*Ch*PPO1 (MZ833479), Number of Nucleotides: 1674 bp

ATGGCTCCTACTTATCCCGACTTCACCAGATGCCACAAGCCCTTAGACGTGCCCCCCTATGCAAAGGA
TCTTAATTGCTGCCCTCCTTATCAAGGCGAAACGCCCGTGCCCTTCACCGTCCCGAAGGACCTGCCAA
TGCGCACCCGGGTCTCTGCGCATCGAGTGGATGAAAAGTATGTGGCGAAATACGCCCGGGCCGTGCA
GCTGATGAAGGACCTCCCGCCAGACGACCCCAGGAACTTTACCCAGCAGGCCAACGTCCACTGCGCC
TACTGCCAGGGCGCTTACTACATGGGCAACACAGACAAGCCCTTCATGGTCCACAACGGCTGGTTCTT
TCTCCCCTGGCACCGCTGGTATCTCTACTTCCACGAGCGCATGTTGGCAAAACTGGTCGGCGACGACA
CCTTCGCGCTGCCTTTCTGGAACCACGACTATGAACCGGGCATGACCATTCCGCCCATGTACTTAGAC
AAAAGCCTCTCGCTCTTTCACGACAAGAGGAACCCCGATCATGGTCCCTCCTCCATTGCAGACCTCAG
CTTCTACATGCCCGACGGCTACAGCAAATCGCACTTCCTTTCTTCCTCCGAAGAGGATCAGGTTTCTTA
TAACTACAGGCTCATGTATAAACACATGGTTTCCCTGGCCAAGACCACCCGGACCTTCCATGGCGCGC
CGCGCAGATCCACCGGCCCGGACAGTTCGACATGCACTTGCGCCGCCGGAACAGTGGAGATTGTGCC
GCACGACACTGTCCATTCTTGGACTGGCAACCTCATTACTGAACAGTGTCCATGCCCGACCAAGACAT
GCTACGGGGAGGACATGGGAGTCTTTTATGCCGCTGCACGGGACCCCATTTTCTACGCTCTCCACGCC
AACGTCGACCGGATGTGGACGATTTGGAAGAAGCTCCCCGGGCGGGAGGACTACACCGACCCGGACC
TGCTCGACACCCTCTTTTACTTCTACGACGAAAACAAACTCCTCGTCAGCGTCCGCGCCGGCGACGCC
TTCGACCCGGCGGAGCTCCGCTATAGCTACGAAGCTTCCCCTACTCCCTGGCTCCAAAAACCAGTGTA
TAACGATATTCGGCCCGCATTGCCTCCCCTGGCCGGCGGCGCAGTTCAGAAATGCAACTCGCAATACA
ATATACTTTTTTCTCAGCCGATCAAGGTGCCCGTACCGAGGCAGGGGGGAAGGAGCAATTTTTTTGGG
CTTATGGCGCGGGAGGAGCTTCTAGTTATCAAAGGGATACGCGTGCAGAGAAACCTTGCGGTTAAGT
TTGATATTTTTGTGAATCTGGATGATCGGCTGCTGGCTACGTCCTACGTGGACAGCGTGCCACGGGAG
TACGCAGGGACGTTTCTTGATCTGCCGCGCGGAGCGCGAGATGACGCGGAGGGCGGGTGTGGGCAAT
CGTCGCTGTGTGTGGGGATTTCCGAGGTGCTGAGGGATTTGAATGTGAAGCAGAATGAGGAGAGTAT
TGTGGTGACTCTGCTTCCTAAAGTGAAGGTGGAAGAAGCCATTAAAATCTCTGCTATCAAGCTCGAAT
ACAGTGAAAGAAATTTGTTTGGAATGAGATGGATGAAGAGCCTCCTCCGCTCAGTACTGATCTCTCCA
TCTAAGTGGGGACTTTGCCCGCCGTTGGGTGTGTTCCCATATCATTAA

>*Ch*PPO2 (MZ833480), Number of Nucleotides: 1557 bp

ATGCCTCCTCCCAATCTCAAACGCTGCCATGAGCCCGACCTTCCAACCAACGCCGACAATGTCAACTG
CTGCCCACCCTACTCCAAAAAGGCGATAGATTTTGAGTTCCCGGAGAACTTGCCCATGCGAGTCCGTC
CTTCTGCGCACAACTTGGACGACACATATGTCGAAAAGTACAAGCGGGCCGTGCAGCTGATGCGCGA
GCTGCCGTCGGACGATCCTCGCAGTTTCTCACAGCAAGCCAACGTTCACTGCGCCTACTGCGACGGGG
CTTACAACGAAGGATCCCTTTCGACCGAGCTCCAAATCCACAACTCCTGGTTTTTCCTGCCGTGGCAT
CGCTGGTATCTCTATTTCCACGAACGCATTCTCGCAAAACTCGTGGGAGACGACAGCTTCGCGCTGCC
CTTCTGGAATTGGGACGCCCCCGCCGGAATGACTTTTCCCCCAGTATACACGGACTCCAGCTCTTCTTT
ATATGATCAGCTCAGAAACGCTGCCCACATGCCGCCGACTATAATTGATCTCAACTTCGACCCAACGG
CGACAACAAAGGTAACGAACGACAATCTGATTGCATCCAACACGAATCTTATGTACAGACAAATCGT
CTCAGGCGCAAAGACGACGATGCTTTTCGAGGGTCAACCCTACAGATATGGCGACGATCCCGACCCA
GGGGCCGGAACGCTGGAGAACGCCCCTCACGGTCCCGTGCACCTCTGGACCGGGTCCTCCCGCCAGC
CCAACCTGGAGGACATGGGAACTTTCTATTCTGCCGCTCGCGACCCCATTTTCTTTGCCCACCATGCA
AACGTGGACCGAATTTGGACCATCTGGAGATCACTGGGAGGGAAGCGGCGAAGAGATTGGGACGAC
AAGGATTTCTTGGGGTCGCGCTTTCTTTTCTACGACGAGAATGCGCGTTTGGTGCGGGTGAAGGTGGC
GGACGCTTTGGACCCCAATAAGCTCATGTTCTCATACCAAAAAGTTGACATTCCATGGATAAATGCGC
GGCCGAAGAAGAGCAAGACAAAAGTACAGACGGCGTCGACCTTCCTCACGCCGCCACCTTCAGCTAT
GGCGGCTGCTCGCGTTGCTGAATTCGGCTCAAAACCCAGACAGCTGGTGCGTCCCATCAGTGCTCGCG

TGAAGAGACCGAAGAAGAAGCGCAGCGCGAAGGAGATCGAGGAGGAAACGGAGGAAGTGCTGGTG
ATTCAGGGGATCGAGGTGAAAAGAGGCTCCACTGCAAAGTTTGATGTGTTCATAAATTTACCGGAGG
CTGATCCCGACACCGTTATTTCTTGTGCGGAGTACGCAGGGACATTCGCAAATGTTCCTCACCATCAG
CACCACCATGAGCAGCGAAGTAGCAAGACCAAGTTATCAAACGTTAAAAGCCCATACAGAAAGTCGT
CGTTTAAGGTGGGAATTACAGAGATATTAGAAGATTTAAATGCAAGCGAGGACGAGGATATCGTGGT
TACTCTTATCCCAAAAGGAGATTTCAAGAACGACCCCATCAAAATCTCATCGATCAAGATCGAATACG
ATTGA

>*Ch*PSS (MZ833481), Number of Nucleotides: 483 bp

ATGGTGGCGAAAAGTATAACAGTAGAAATTGATTCTCCAGTGGAGGCAAAGAGATTTTGGGCTGCAA
TTGTTAAGGACTATGATCTCCTCCCTAGGCAAATGCCAGGGGTATGTTCAGGCGTCACACTTGTTAAA
GGCGACGGAGGTGTCGGCACCATCATACAAATTGACTACACCCCTGTGAACAAGGATTTTAGTTACGT
GAAGGAGCGAGTGGATGAAGTAGATGAGGGAAACTTTGTTTACAGCTTCAGCTATGTGGAAGGGGGA
GAGGTGGGAACGAAATGGGCATCTGCTAAGTTTAAGATGCAATTGACACCGAAAACAGAGGGTGGAT
GCGTTTTGAAGCATACATGCGAGTATGACACCCTGCCCGGTATTCCCCTTGACGAAGCCAAACTGGAA
GAGATGGAGAGGAGCGCCGCTGGTCTCTTAAAATCCATAGAAGCATACCTCGTCTCTAACCCTACCGT
ATATTGTTAA