UNIVERSITÉ DU QUÉBEC

THÈSE PRÉSENTÉE À
L'UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

COMME EXIGENCE PARTIELLE
DU DOCTORAT EN BIOLOGIE CELLULAIRE ET MOLÉCULAIRE

PAR
APARNA SINGH

DISCOVERY OF NOVEL GENES AND ENZYMES INVOLVED IN AMARYLLIDACEAE ALKALOID BIOSYNTHESIS USING INTEGRATED METABOLOMICS AND TRANSCRIPTOMICS IN *NARCISSUS PSEUDONARCISSUS* 'KING ALFRED'

JUIN 2018

UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

Cette thèse a été dirigée par :

Isabel Desgagné-Penix, Ph. D.	Université du Québec à Trois-Rivières
Directrice de recherche	Institution à laquelle se rattache l'évaluateur

Jury d'évaluation de la thèse :

Isabel Desgagné-Penix, Ph. D.	Université du Québec à Trois-Rivières
Prénom et nom, grade	Institution à laquelle se rattache l'évaluateur
Hugo Germain, Ph. D.	Université du Québec à Trois-Rivières
Prénom et nom, grade	Institution à laquelle se rattache l'évaluateur
Céline Van Themsche, Ph. D.	Université du Québec à Trois-Rivières
Prénom et nom, grade	Institution à laquelle se rattache l'évaluateur
Charles Goulet, Ph. D.	Université Laval
Prénom et nom, grade	Institution à laquelle se rattache l'évaluateur

Thèse soutenue le 28 mai 2018

To my parents and husband who were my
strength in this journey

ACKNOWLEDGEMENTS

I am thankful to my research supervisor Isabel Desgagné-Penix for considering me in her laboratory and acquainted me enthusiastically to the field of plant specialized metabolism. She has been an inspiration as a researcher and mentor throughout my Ph.D. She always supported me with her guidance for a successful implementation of the research project. Working in her laboratory has helped me tremendously in developing scientific skills and knowledge.

I am also equally thankful to Professor Hugo Germain (Université du Québec à Trois-Rivières) for his unconditional support, guidance, and manuscript reviewing. His well-established lab has provided me with instruments and reagents, as per the requirement of the project. I am also grateful to Professor Gervais Berubé (Université du Québec à Trois-Rivières) for standard synthesis and manuscript writing. I would like to acknowledge Professor Céline Van Themsche (Université du Québec à Trois-Rivières) for sharing instruments.

Special thanks to Ariane Garand and Marie-Ange Massicotte for working on experiments related to my project. I am also thankful to Vincent Ouellette for synthesizing required chemical standards. I would like to thank Tarun Hotchandani for manuscript reviewing.

I would like to express my gratitude to all my labmates, Nicolas Dufour, Tarun Hotchandani, Andrew Diamond, Narimene Fradj, Bulbul Ahmed, Saifur Rahman, Annabelle St-Pierre, Dorian Blondeau, Genevieve Laperriere, Claire Letanneur, Sarah Piette, Joelle Rancourt, Mélodie B. Plourde, Teura Barff, Hur Madina, and Karen Cristine Gonçalves dos Santos for their encouragement and good memories we created in this course of time.

RÉSUMÉ EN FRANÇAIS

Au cours des 50 dernières années, des études sur les métabolites secondaires ont été rapportées pour leur rôle majeur dans l'acclimatation de l'environnement végétal. Ils représentent également des molécules d'importance pharmaceutique. Les alcaloïdes sont des métabolites produits par différentes familles de plantes. Les plantes de la famille des Amaryllidaceae produisent des alcaloïdes des Amaryllidaceae (AA) impliqués dans la défense et le développement végétal. Ils possèdent un grand nombre de propriétés bioactives qui en font une cible intéressante pour le développement de médicaments. Les exemples d'AA comprennent la lycorine anti-mitotique, la galanthamine inhibant l'acétylcholinestérase et la crinine cytotoxique. Tous les AA partagent un précurseur intermédiaire commun dans leur voie de biosynthèse, nommé la norbelladine. Comme d'autres métabolites végétaux, les AA sont également produits en très faibles quantités dans les plantes, ce qui les rend inaccessibles pour des applications pharmaceutiques. Selon les études, la galanthamine, qui est extraite de diverses plantes d'Amaryllidaceae, contribue à seulement 0,2% du poids sec d'une plante. Jusqu'à présent, seule la galanthamine est utilisée commercialement pour le traitement de la maladie d'Alzheimer. Son extraction à partir de matériel végétal est un processus coûteux. De plus, la structure complexe des AA est difficile à synthétiser chimiquement. Par conséquent, la découverte de nouveaux gènes codant pour des enzymes biosynthétiques est nécessaire afin de faciliter le développement de méthodes de production biotechnologiques alternatives. Comme les Amaryllidaceae sont des plantes non modèles, nous avons appliqué la biologie des systèmes pour atteindre nos objectifs de recherche. Parmi ceux-ci nous avons utilisé deux outils 'O'miques (métabolomique et transcriptomique) dans cette recherche. Les deux objectifs de ce projet étaient, en premier lieu, d'étudier les gènes et les précurseurs de la biosynthèse des AA par une étude comparative métabolomique et transcriptomique de *Narcissus pseudonarcissus* 'King Alfred'. Le deuxième objectif a impliqué le clonage et la caractérisation d'une *norbelladine synthase* (*Np*NBS), un nouveau gène catalysant la première étape dans la biosynthèse des AA. À partir de la première étude, nous avons obtenu un profil de métabolites par analyse UPLC-QTOF-MS qui présentait un grand nombre d'AA synthétisés dans différents tissus végétaux, principalement dans les feuilles et les bulbes. Les trois catégories d'AA ont été détectées dans différents tissus. Les AA de la catégorie *para-ortho'* (c'est-à-dire de type galanthamine) et *ortho-para'* (c'est-à-dire de type lycorine) étaient abondants dans les tissus souterrains alors que l'inverse était observé dans les feuilles. En outre, l'assemblage et l'annotation du transcriptome *de novo* ont donné 11 708 transcrits annotés. L'annotation du transcriptome a révélé la présence de tous les gènes de biosynthèse des AA proposés, qui soutenaient leur implication dans la

biosynthèse des AA susmentionnés dans les tissus de *N. pseudonarcissus* King Alfred. Nous avons également constaté que les gènes les plus exprimés appartenaient à des processus tels que le métabolisme, la survie et la défense. Notamment, *Np*NBS a été détecté parmi les gènes les plus exprimés. Nous avons identifié avec succès la séquence complète du gène *Np*NBS à partir du transciptome annoté de *Narcissus pseudonarcissus* 'King Alfred', qui a été plus tard cloné, exprimé et caratériser par un dosage enzymatique. Nous avons trouvé que *Np*NBS était responsable de la formation de la norbelladine. En outre, phylogénétiquement, elle est liée à une norcoclaurine synthase de *Papaver somniferum* (*Ps*NCS) et est une protéine de la famille des Bet v1/PR-10. *Np*NBS catalyse la première étape engagée dans la voie de biosynthèse des AA et fournit une structure de base commune à divers AA. Ainsi, nous avons pu caractériser un nouveau gène de la voie des AA et contribuer à la découverte de la voie de biosynthèse des AA en construisant une banque de données pour la caractérisation d'autres gènes inconnus dans l'avenir.

Mots-clés: Alcaloïdes des Amaryllidaceae, norbelladine, norcraugsodine, transcriptome, séquençage de nouvelle génération, biosynthèse, métabolisme spécialisé, métabolite secondaire, norbelladine synthase, pathogenesis-related class-10 (PR-10).

ABSTRACT

Over the last 50 years, studies on specialized metabolites have increased for their medicinal properties and major roles in acclimatization of the plant to their environment. They are natural molecules of immense pharmaceutical importance. Alkaloids comprise one of the categories of specialized metabolites that are produced by different plant families. The plants of Amaryllidaceae family produces large number of Amaryllidaceae alkaloids (AAs) thought to be involved in their defense and development. Their enormous bioactive properties make them an interesting target for drug development. Few examples include an anti-mitotic alkaloid-lycorine, an acetylcholinesterase inhibiting alkaloid-galanthamine, and the cytotoxic alkaloid-crinine. All AAs share a common intermediate precursor in their biosynthesis pathway, norbelladine. Like other plant metabolites, AAs are produced in very low quantities in plants which makes them inaccessible for pharmaceutical applications. According to the data, galanthamine extraction from various Amaryllidaceae plants contributes to only 0.2% dry weight of a plant. Until now, only galanthamine is used commercially for the treatment of Alzheimer's disease. Its extraction completely relies on natural sources, which makes it an expensive process. Moreover, the complex structure of AAs is difficult to reproduce chemically. Therefore, a better understanding of their biosynthesis is necessary for their large-scale production through metabolic engineering. As Amaryllidaceae are non-model plants, we applied systems biology to accomplish our research goals. Among several OMICS tools, we utilized metabolomics and transcriptomics in this study. The two objectives of this project were, first, to study the AAs biosynthesis genes and precursors by a comparative metabolomics and transcriptomics of *Narcissus pseudonarcissus* 'King Alfred'. Second, to clone and characterize *Norbelladine Synthase (NpNBS),* a novel gene catalyzing the first committed step in AAs biosynthesis. From the first study, we obtained a metabolite profile by UPLC-QTOF-MS analysis, which displayed an array of AAs synthesized in different plant tissues, predominantly in leaves and bulbs. All three categories of AAs were detected in different tissues. *Para-ortho'* (*i.e.* galanthamine-type) and *para-para'* (*i.e.* crinamine-type) AAs were abundant in all tissues while *ortho-para'* (lycorine-type) were abundant in the leaves. In addition, *de novo* transcriptome assembly and annotation yielded 11,708 annotated transcripts. It revealed the presence of all proposed AAs biosynthesis genes, which supported their involvement in the biosynthesis of above-mentioned AAs in *N. pseudonarcissus* 'King Alfred' tissues. We also found that the top expressed genes belonged to processes like metabolism, survival, and defense. Notably, *NpNBS* was detected among the top expressed genes. We cloned a full-length *NpNBS* gene sequence from *Narcissus pseudonarcissus* 'King Alfred' transcriptome, which was later induced for protein expression to performed enzyme assay. We found that *Np*NBS is responsible for

norbelladine formation. Also, phylogenetically it related to a norcoclaurine synthase from *Papaver somniferum* (*Ps*NCS) and is a Bet v1/PR-10 protein. *Np*NBS enzyme was found to catalyze the first committed step in AAs biosynthesis pathway and provides a common core structure to diverse AAs. Thus, we were able to characterize one novel gene of AAs pathway and contributed to AAs biosynthesis pathway discovery by building a platform for the characterization of other unknown genes in future and synthetic biology.

Keywords: Amaryllidaceae alkaloids, norbelladine, norcraugsodine, transcriptome, next-generation sequencing, biosynthesis, metabolism, specialized metabolite, norbelladine synthase, pathogenesis-related class-10 (PR-10) protein.

TABLE OF CONTENTS

LIST OF TABLES

LISTE OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

AAs	Amaryllidaceae alkaloids
AD	Alzheimer's disease
AKR	Aldo-keto reductase
BBE	Berberine bridge enzyme
BIA	Benzylisoquinoline alkaloids
BLAST	Basic local alignment search tool
BSA	Bovine serum albumin
CAS 9	CRISPR associated protein 9
CID	Collision-induced dissociation
CNMT	Coclaurine *N*-methyltransferase
CODM	Codeine *O*-demethylase
CoR	Codeinone reductase
CRISPR	Clustered regulatory interspaced short palindromic repeat
CYP	Cytochrome P450
C3H	*p*-coumarate hydroxylase
C4H	*Trans*-cinnamate hydroxylase
DBOX	Dihydrobenzophenanthridine oxidase
DFI-MS	Direct flow injection-tandem mass spectroscopy
DODC	DOPA decarboxylase
DMSO	Dimethyl sulfoxide
EST	Expressed sequence tags
FPKM	Fragment per kilobase million

GAL	Galanthamine
GAM-HRP	Goat anti-mouse-horse radish peroxidase
HBS	Hydroxybenzaldehyde synthase
HCT	Hydroxycinnamoyl transferase
HIS	Histone
HPLC-PDA	High performance liquid chromatography-photodiode array detector
IPTG	Isopropyl-β-D-thiogalactopyranoside
LB	Luria Bertani
LC-MS	Liquid chromatography-mass spectrometry
L-DOPA	L-3,4-dihydroxyphenylalanine
LYC	Lycorine
MRM	Multiple reaction monitoring
NaH_2PO_4	Sodium phosphate monobasic
NAR	Narciclasine
NBS	Norbelladine synthase
NCS	Norcoclaurine synthase
NGS	Next generation sequencing
NMCH	*N*-methylcoclaurine 3'-hydroxylase
NMR	Nuclear magnetic resonance
NMT	*N*-methyltransferase
N4OMT	Norbelladine-4'-*O*-methyltransferase
NR	Noroxomaritidine reductase
ODM	*O*-demethylase
OMT	*O*-methyltransferase
ORF	Open reading frame

MIA	Monoterpene indole alkaloids
PAL	Phenylalanine ammonium lyase
PA	Pyrrolizidine alkaloid
PCR	Polymerase chain reaction
PR-10	Pathogenesis related protein family 10
PTZ	Pretazettine hydrochloride
PVDF	Polyvinylidene difluoride
QA	Quinolizidine alkaloids
R	Reductase
RIN	RNA Integrity number
SDRs	Short chain dehydrogenase/reductase
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
STOX	(S)-tetrahydroxyprotoberberine
TBST	Tris-buffered saline (TBS) containing tween 20
TLC	Thin layer chromatography
TyrAT	Tyrosine aminotransferase
TYDC	Tyrosine/DOPA decarboxylase
T6ODM	Thebaine-6-*O*-demethylase
SalSyn	Salutaridine synthase
SalR	Salutaridine reductase
SAM	S-adenosyl methionine
SalAT	Salutaridinol 7-*O*-acetyltransferase
Sh/Q	Shikimate/quinate
TALEN	Transcription activator-like effector nucleases
TyrAT	Tyrosine aminotransferase

T6ODM	Thebaine 6-*O*-demethylase
V11H	Vittatine 11-hydroxylase
VIGS	Virus-induced gene silencing
3OHase	Tyrosine/tyramine 3-hydroxylase
4CL	4-hydroxycinnamoyl CoA ligase
4HPAA	4-hydroxyphenylacetyldehyde
4HPPDC	4-hydroxyphenylpuruvate decarboxylase
4'OMT	3'-hydroxyl-*N*-methylcoclaurine 4'-*O*-methyltransferase
6OMT	Norcoclaurine 6-*O*-methyltransferase

CHAPTER I

INTRODUCTION

1.1 Plant Specialized Metabolites – A Medicinal Resource

Plants produce a variety of compounds referred as 'specialized metabolites' or 'phytochemicals' to adapt well to their ecological habitat (Bourgaud, Gravot et al. 2001, Oksman-Caldentey and Saito 2005). Unlike primary metabolites, these molecules were not found to be involved in the primary biochemical processes of plants such as cell division, growth, respiration, reproduction, and storage of nutrients (Bourgaud, Gravot et al. 2001). Most of the specialized metabolites are found to be associated with defense and aid plants to fight against numerous biotic and abiotic stresses (Croteau, Kutchan et al. 2000, Akula and Ravishankar 2011, Calabrò 2015). Their absence does not lead to the direct death of a plant, but makes them more susceptible to diseases, pest, and predators, which in turn affects its long-term survival. Based on their biosynthetic origin, these structurally diverse compounds are categorized into alkaloids, terpenoids, flavonoids, glycosinolates, carotenoids, lignins and phenolics (Fig.1.1) (Der Marderosian and Beutler 2002). Their biosynthesis is derived from common intermediates such as shikimic acid, acetyl-coenzyme A (acetyl-CoA), mevalonic acid and 1-deoxyxylulose-5-phosphate (Dias, Urban et al. 2012).

Specialized metabolites have been used for centuries in traditional medicines for the treatment of diseases and illnesses. The oldest record of their medicinal use comes from Mesopotamia (2600 B.C.), which reports about *Cupressus sempervirens* (Cypress) and *Commiphora* species (myrrh) oils for cough, cold and inflammation treatment (Cragg and Newman 2005, Dias, Urban et al. 2012). Egyptian pharmaceutical records a collection of more than 700 plant-based drugs such as pills and ointments (Cragg and Newman 2005). Similarly, Chinese and Greek civilizations have documented thousands of such medicinal plants (Dias, Urban et al. 2012).

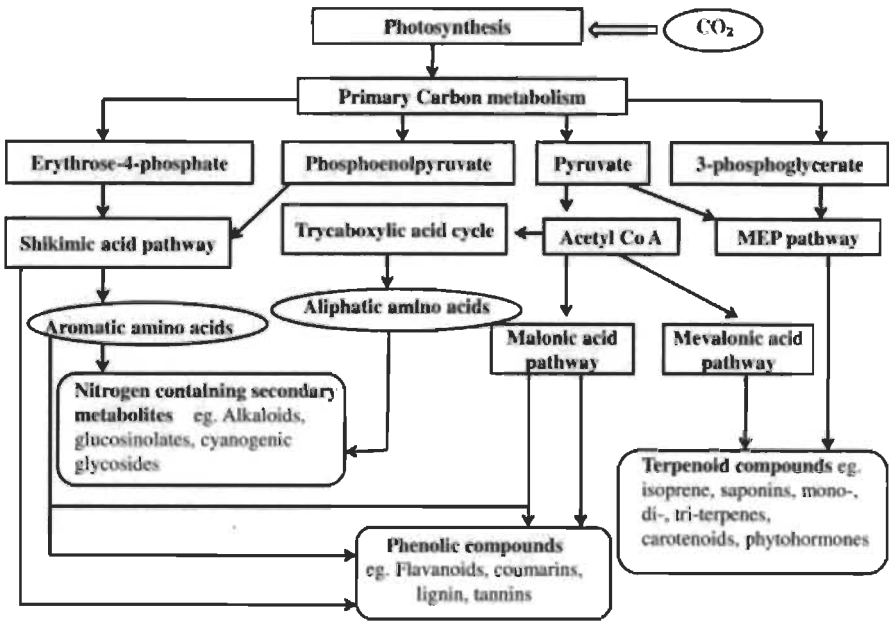


Figure 1.1 Classification of primary and specialized metabolite biosynthesis pathway. Source (Verma and Shukla 2015)

Currently, most of the plant metabolites are extracted from natural sources and used by pharmaceutical industries for their broad range of biological effects (Table 1.1). Famous examples of naturally derived medicine include aspirin, isolated from the bark of willow tree *Salix alba* (DerMarderosian and Beutler 2002), codeine and morphine which were extracted from *Papaver somniferum L.* in 1803 (DerMarderosian and Beutler 2002), vincristine and vinblastine were obtained from *Catharanthus roseus*, tetrahydrocannabinol present in *Cannabis sativa*. Another specialized metabolite, quinine is extracted from the bark of *Cinchona succirubra* Pav. ex Klotsch, and used as an anti-malarial drug (Der Marderosian and Beutler 2002). Also, pilocarpine alkaloid obtained from *Pilocarpus jaborandi* is used as a clinical medicine to treat chronic open-angle glaucoma and acute angle-closure glaucoma (Aniszewski 2015).

Class	Compounds	Sources	Effects and Uses
Alkaloids	Nicotine, theobromine	Tobacco	Interfere with neurotransmission, block enzyme action
Monoterpenes	Menthol	Mint	Interfere with neurotransmission, block ion transport, anesthetic
Sesquiterpenes	Parthenolid	Parthenium (*Asteraceae*)	Contact dermatitis
Diterpenes	Gossypol	Cotton	Blocks phosphorylation
Triterpenes	Digitogenin	Digitalis	Stimulate heart muscle, alter ion transport
Tetraterpenes	Carotene	Many plants	Antioxidant
Sterols	Spinasterol	Spinach	Anti-inflammatory
Coumarins	umbelliferone	Carrots	Cross-link DNA, block cell division
Lignans	podophyllins	Mayapple	Allergic dermatitis
Tannins	Gallotannin, condensed tanin	Oak, legumes	Bind to proteins, enzymes, block digestion, antioxidants
Lignin	Lignin	All land plants	Structure, toughness, fiber

Table 1.1 **List of plant specialized metabolites and their biological effects.** Source: (www.biologyreference.com).

Today, specialized metabolites are majorly known for their biological functions and investigated at an increasing pace. They are expressed by plants under certain conditions and vary among species (Dewick, 2002). Approximately 200,000 compounds have been identified from different species by far, each showcasing a unique bouquet of specialized compounds (Hartmann 2004).

Unfortunately, many of these plant metabolites are produced in a very low quantities in plants because their concentration is influenced by developmental, physiological and environmental factors (Figueiredo, Barroso et al. 2008). Various abiotic factors such as light intensity, temperature, drought, nutrients, salinity, wound, UV radiation and biotic factors like pathogen attack affect their induction and accumulation in plants (Gouvea, Gobbo-Neto et al. 2012, Verma and Shukla 2015). The factors influencing specialized metabolite production in plants are principally divided into four categories; genetic, ontogenic, morphogenetic environmental factors (Fig. 1.2) (Verma and Shukla 2015). There are ongoing clinical trials on more than 100 plant derived drugs and several compounds are at preclinical stage (Katiyar, Gupta et al. 2012). Also, *de novo* chemical synthesis of natural compounds obtained a limited success due to their highly complex structures (Oksman-Caldentey and Inzé 2004). Such challenges often create an inadequate access to these compounds and suggests that the reliance on the natural resources for their extraction is not a sustainable solution. Hence, an implementation of metabolic engineering has gained researchers interest for industrial production of specialized metabolites. It has been already performed for the production of morphine and codeine with engineered yeast to produce key precursors for artemisinin synthesis, which eventually leads to morphine and codeine production (Ro, Paradise et al. 2006, Hawkins and Smolke 2008). Therefore, expansion in the knowledge of specialized metabolism biosynthesis is required to develop tools for synthetic biology.

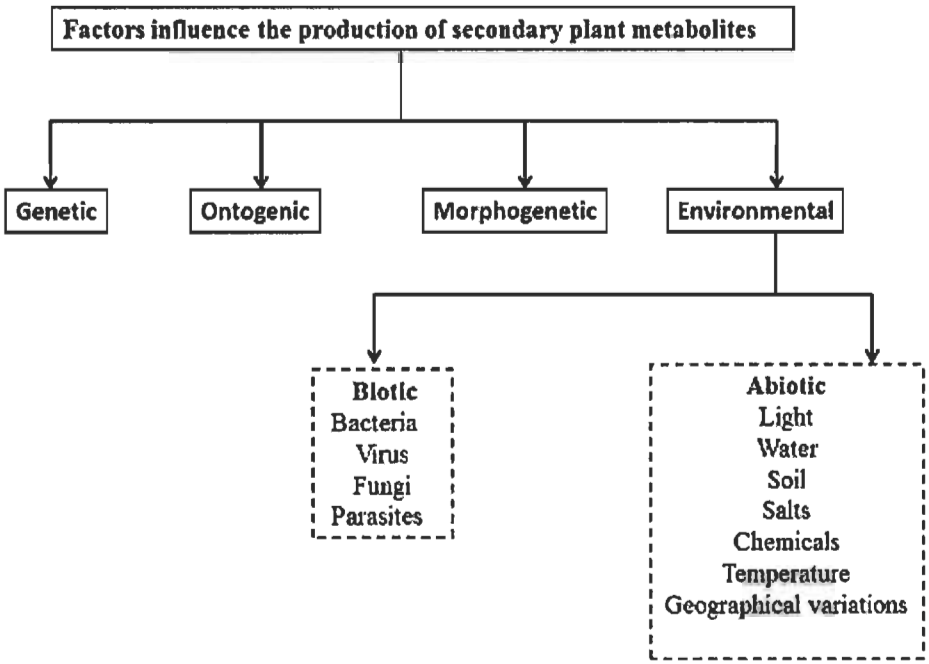


Figure 1.2 Different factors influencing the induction of specialized metabolite biosynthesis in plants. Source (Verma and Shukla 2015).

.

1.2 Classification of Specialized Metabolites

Based on their biosynthetic origin, plant specialized metabolites were classified into three major groups the phenolics, the terpenes and the alkaloids.

1.2.1 Phenolics

These metabolites are characterized by the presence of atleast one aromatic ring, having one or more hydroxyl group (Achakzai, Achakzai et al. 2009). More than a thousand phenolics have been reported by far that are dispersed within plant kingdom (Strack 1997). They are biosynthesized essentially from malonic and shikimate pathway. The malonic acid pathway is less commonly found in higher plants whereas shikimic acid pathway is significantly common in plants. Phenolic compounds have been subdivided into lignins, lignans, coumarins, flavonoids, tannins, stilbenes, styrylpyrones and arylpyrones (Fang, Yang et al. 2011). Phenylpropanoids such as *trans*-cinnamic acid, *p*-coumaric acid and its derivatives are some commonly known simple phenolic compounds and lignin is the examples of complex phenolic compound which are present abundantly in plants (Lincoln and Zeiger 2006). They have a broad range of roles in plant development and growth, reproduction, and defense against various environmental factors (Achakzai, Achakzai et al. 2009). Some also exhibit allelopathic effect by inhibiting the growth of neighboring plants and seedlings(Lincoln and Zeiger 2006). Due to their antioxidant properties, they reduce the risk of different human cancers (Gangwar, Gautam et al. 2014). Plants uses them as a protective agent against nematodes, fungal and bacterial pathogens (Dakora and Phillips 1996, Lattanzio, Lattanzio et al. 2006). They are secreted as phytoalexins during wound or pathogen attack and also provide structural strength to a plant by lignin biosynthesis (Bhattacharya, Sood et al. 2010).

1.2.2 Terpenes

Terpenes constitute the largest class of structurally diverse natural compounds (Cheng, Lou et al. 2007). All terpenoids are derived from the repetitive fusion of five-carbon (isoprene) units (Cheng, Lou et al. 2007). They are classified based on the number of isoprene units such as monoterpenes (consisting of two isoprene units), sesquiterpenes (derived from three isoprene units), diterpenes (four isoprene units), triterpenes, and tetraterpenes (eight isoprene units) (Cheng, Lou et al. 2007). There are more than 23,000 identified terpenes with a significant role in plant defense against insects and other pathogens (Dudareva, Pichersky et al. 2004). Few terpenes which have a crucial role in plant growth include gibberellins (diterpenes), sterols (triterpenes), carotenoids (tetraterpenes), and abscisic acid (sesquiterpenes) (Verma and Shukla 2015). Their synthesis in plants occurs through two different pathways, the mevalonate (MVA) and methylerythritol 4-phosphate (MEP) pathway. The MVA pathway occurs in cytosol and MEP pathway is found in plastid. The MEP pathway provide precursor for mono and diterpenes, isoprene, carotenoids, phytohormones, gibberellins, abscisic acid (ABA), phytol, phylloquinones, tocopherols, chlorophylls biosynthesis and MVA pathway synthesizes isopentenyl diphosphate to allow biosynthesis of sesquiterpenes, sterols, brassinosteroids, polyphenols and their moieties (Pateraki and Kanellis 2010).

1.2.3 Alkaloids

Alkaloids are a large group of naturally occurring plant specialized metabolites. They have been used long before their discovery, in a herbal form for curing diseases. The term 'alkaloid' was coined by W. Meissner in 1819 (Pelletier 1983). According to his observations, these compounds behave like alkali. They are a low molecular weight, heterocyclic-nitrogen containing compounds (Aniszewski 1994). Alkaloids may also possess primary, secondary, tertiary or quaternary ammonium groups (Aniszewski 1994). Many of these alkaloids are also used by humans since early times as therapeutic agents. Although, they do not have a vital role in plant growth many of them are known for their pharmacological properties (Table 1.2). Mostly, these alkaloids are derivatives of amino

acid precursors such as tyrosine, lysine, tryptophan and phenylalanine (Facchini 2001, Fattorusso and Taglialatela-Scafati 2008). They provide protection to the plant by accumulating in those parts which are commonly attacked by herbivores (Bate-Smith 1972, Bernays 1983). Also to prevent the plant from toxicity, alkaloids accumulate in specialized cells inside plants, known as laticifers (De Luca and St Pierre 2000). Studies conducted on morphine shows that only 0.3-1.5% morphine accumulates in whole poppy capsule and more than 25% morphine accumulates in the specialized cells of dried Opium poppy latex (Roberts, McCarthy et al. 1983, Ziegler and Facchini 2008). Likewise, quinolizidine and pyrrolizidine alkaloids accumulate in aerial tissues which are more prone to pathogen attack (Lee, Pate et al. 2006).

Plant species	Metabolites	Properties
Camptotheca acuminate	Camptothecin	Anti-cancer, antiviral
Gingko biloba	Ginkgolides	Aging disorders
Gmelina arborea	Verbascoside	Stomach disorders, fevers, skin problems
Gynostemma pentaphyllum	Gypenoside	Detergent
Linum flavum	Coniferin	Anti-cancer
Papaver somniferum	Morphine, sanguinarine, codeine	Sedative, analgesic
Pueraria phaseoloides	Puerarin	Hypothermic, spasmolytic, hypotensive, anti-arrhythmic
Rauvolfia micrantha	Ajmalicine, ajmaline	Anti-hypertensive
Saussurea medusa	Jaceosidin	Anti-tumor
Solidago altissima	Polyacetylene (*cis* dehydromatricaria ester)	Unknown

Table 1.2 Example of a few alkaloids with their source and applications. Source (Guillon, Trémouillaux-Guiller et al. 2006)

1.3 Transportation and storage of specialized metabolites in the plants

Specialized metabolites are often secreted or stored by plants in special compartments which involves a complex organization of multiple cells showing an exhaustive inter-cellular and intra-cellular transport process of metabolites and enzymes (Ziegler and Facchini 2008). Structures like glandular trichomes, idioblasts, laticifers, resin canals, epidermis, endodermis, pericycle and cortex are known to be associated with their accumulation inside plants (Facchini, 2005). Fundamental reason behind the evolution of these structures is to prevent autotoxicity, provide defense and maintain

active production and delivery of these products in order to improve the interaction with other organisms (Bidlack, Omaye et al. 2000, Ziegler and Facchini 2008).

Various studies on sequestration of specialized metabolites have been reported. Some of them include accumulation of benzylisoquinoline alkaloids in *Papaver sominferum* is confined to laticifers near to sieve elements of the phloem (Alcantara, Bird et al. 2005). The cytoplasm of laticifers, also known as latex, contains alkaloids in discrete vesicles. In addition, a red color polycyclic napthodianthrone, hypericin well-known for its anticancerous, anti-viral and antidepressant activities was found to be located in trichomes of *Hypericum punctatum* forming a dot pattern throughout the flowers, leaves, stems, sepals, and stamens (Fig. 1.3A). A proposed schematic representation of tropane alkaloids and biosynthesis enzymes trafficking shown in figure 1.3B suggests localization of hyoscyamine and scopolamine and their pathway intermediates in the root endodermis and nearby cortical cells (Ziegler and Facchini 2008).

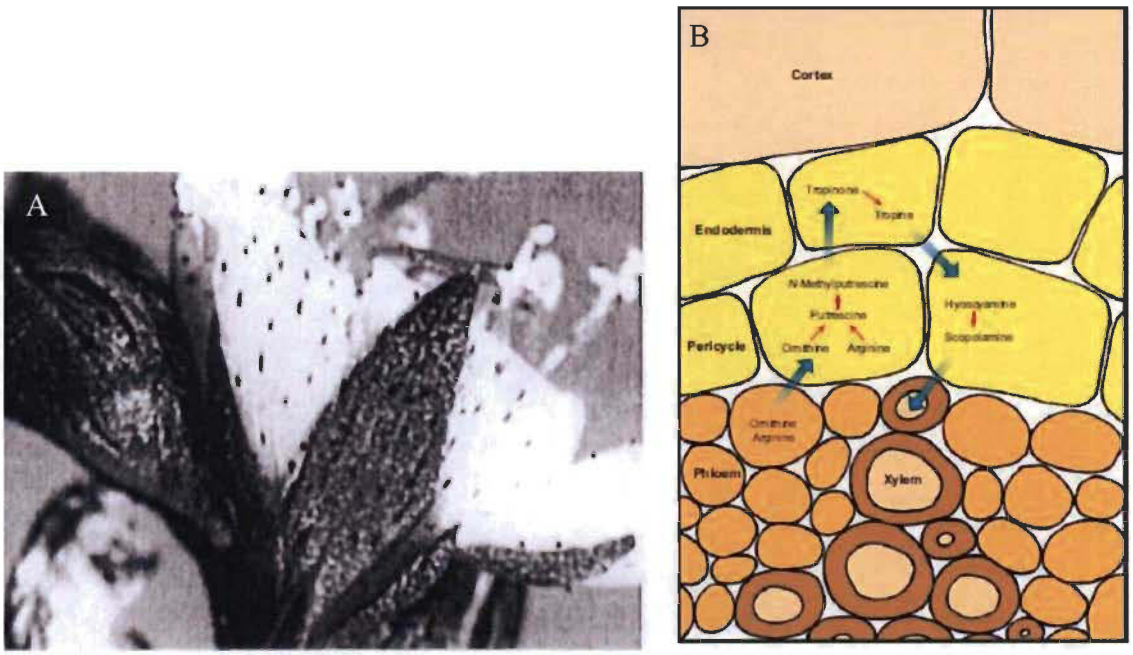


Figure 1.3 Transport and localization of specialized metabolites. (A) Trichomes (dotted structures) containing hypericin present on leaves and sepals of *Hypericum punctatum.* (B) Representation of putative tropane alkaloids and intermediates localization (blue arrows). Source (Bidlack, Omaye et al. 2000, Ziegler and Facchini 2008)

1.4 Classification of alkaloids

To deal with these large number of alkaloids, they are principally divided into different groups (Purine alkaloids, Tropane alkaloids, Pyrrolizidine alkaloids, Quinolizidine alkaloids, Benzylisoquinoline , alkaloids Monoterpene indole alkaloids and Amaryllidaceae alkaloids) based on their structures (Pelletier 1983, Dewick 2002, Hesse 2002).

1.4.1 Purine alkaloids

Purine alkaloids are mainly derived from xanthosine, produced by degradation of purine neuclotides such as adenine and guanine (Zulak, Liscombe et al. 2006). These alkaloids are produced in a variety of taxonomically unrelated species, e.g., *Coffea arabica*, *Camellia sinensis*, *Theobroma cacao*, *Ilex paraguariensis*, *Paullinia cupana*, and *Colanitida* (Springob and Kutchan 2009). Purine alkaloids have numerous roles in plants including protection of vegetative parts against herbivores (Ashihara, Sano et al. 2008). They exhibit an allelopathic effect by inhibiting the germination of other seedlings and also act as a deterrent for slugs and snails (Hollingsworth, Armstrong et al. 2002, Ashihara and Suzuki 2004). Few examples of these alkaloids with their structures are shown in the figure 1.4. The most abundant purine alkaloid is caffeine followed by theobromine and theophylline (Springob and Kutchan 2009). The caffeine content of young dry leaves of *Camellia sinensis* variety *sinensis* is 2-3% whereas it varies from 0.4% to 2.4% in different species of coffea plant (Ashihara and Kubota 1987). Given its stimulating effect, caffeine is widely used in tea, coffee and also as cold and analgesic medicine (Benowitz 1990). More recently, the attention towards caffeine has increased because coffee drinkers have shown a reduced risk for Parkinson's disease (Qi and Li 2014). Purine alkaloids are extensively studied due to their ability to stimulate the central nervous system.

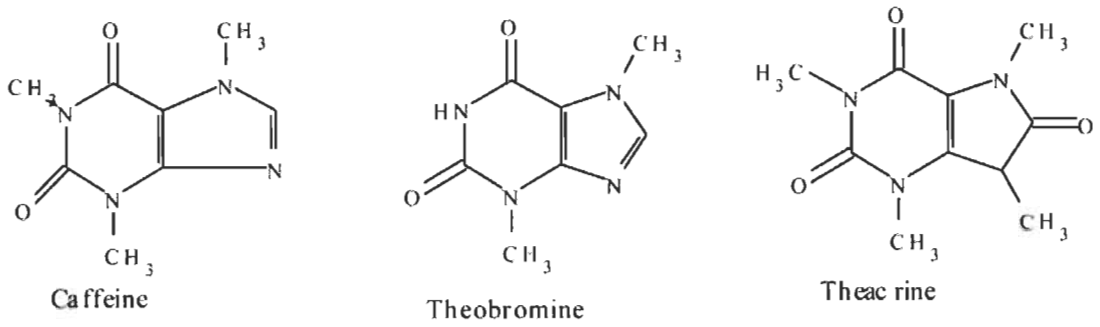


Figure 1.4 Few examples of plant-derived purine alkaloids with their structures.
Source (Ncube and Van Staden 2015)

1.4.2 Tropane alkaloids

Tropane alkaloids possess a bicyclic tropane skeleton with seven-membered ring containing an *N*–bridge between C-1 and C-5. Two amino acids precursors, ornithine and/or arginine are involved in their biosynthesis (Springob and Kutchan 2009). They were first acquired from the nightshade family (Solanaceae). Most of these alkaloids are esters of the alcohols tropine or pseudotropine possessing aliphatic or aromatic acids (Fig. 1.5) (Springob and Kutchan 2009). Tropanes with different structures have been discovered in the *Solanaceae* family, and also from few disparate plant families such as *Brassicaceae, Euphorbiaceae, Erythroxylaceae, Proteaeceae*, and *Rhizophoraceae.* They were used among folklores for their hallucinogenic properties (Griffin and Lin 2000).

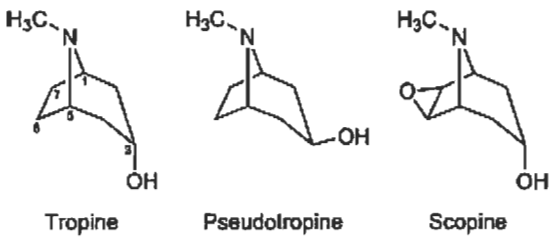


Figure 1.5 Tropane amino alcohols.
Source (Kutchan 2009)

1.4.3 Pyrrolizidine alkaloids

Pyrrolizidine Alkaloids (PAs) are also known as necine, based on the structure of hydroxymethylpyrrolizidine. They are produced by many plants to defend against insects and herbivores (Fig. 1.6) (Hartmann, 1999). Necine is a basic structure biosynthesized from arginine derivatives, spermidine and putrescine (Hartmann, Sander et al. 1988, Ziegler and Facchini 2008). PAs have been discovered from Asteraceae, Boraginaceae, Fabaceae, and Orchidaceae families of a plant (Hartmann, 2000).

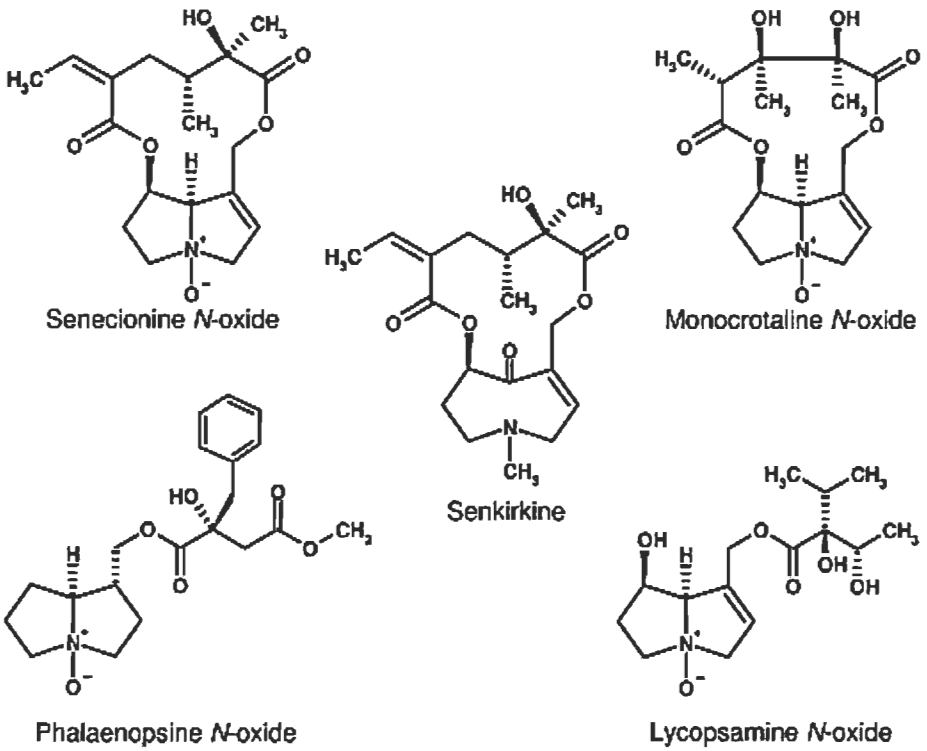


Figure 1.6 Structural types of different pyrrolizidine alkaloids. Some PAs exist in *N*-oxides form in plants. Source (Wang, Beuerle et al. 2012)

By far, more than 660 pyrrolizidine alkaloids have been identified and a majority of them show hepatotoxic effect (Li, 2011). Unsaturated pyrrolizidine alkaloids are well-known for hepatotoxicity and lead to liver cancer (Schoental and Magee 1959, Schoental

1968). These alkaloids were detected in shoots, inflorescence tissues and roots at all stages of plant development, with different concentrations in each plant tissue (Hartmann and Zimmer 1986). In *Senecio vulgaris*, these alkaloids are exported from roots to shoots via phloem and subsequently allocated to their storage site (Hartmann, Ehmke et al. 1989).

1.4.4 Quinolizidine alkaloids

Quinolizidine Alkaloids (QAs) are synthesized from lysine to form lupanine or sparteine which mostly occurs in the aerial parts of leguminous plants (Osbourn and Lanzotti 2009, Frick, Kamphuis et al. 2017). However, it also occurs in other families, such as Berberidaceae, Chenopodiacae, Ranunculaceae, Rubiaceae, and Solanaceae (Wink 2002). Primarily, QAs act as defense compounds and help in nitrogen transport in the phloem. They help plants to deter herbivore feeding, inhibit microbial growth and also display allelochemical effect (Roberts 2013). Examples of quinolizidine alkaloids include sparteine, lupanine, cytisine and lupinine (Fig. 1.7) (Boschin and Resta 2013).

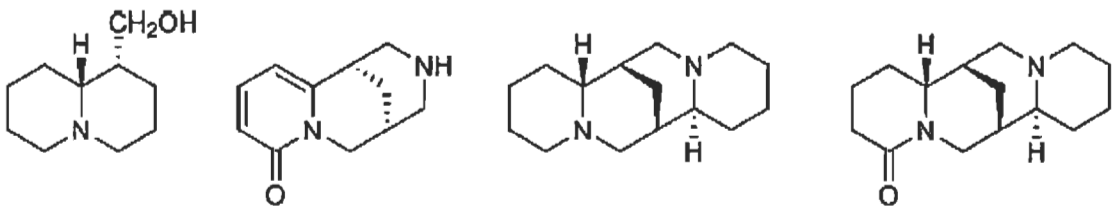


Figure 1.7 Four representative structures of quinolizidine alkaloids.
Source (Osbourn and Lanzotti 2009)

1.4.5 Monoterpene indole alkaloids

Monoterpene indole alkaloids (MIAs) are diverse in terms of structure and function and comprise of over 2000 members (O'Connor and Maresh 2006, Akhtar,

Ahmad et al. 2017). They belong to a special category of terpenoids with two isoprene units (Akhtar, Ahmad et al. 2017). These alkaloids are of major pharmaceutical interest as they possess potent anti-cancer, anti-malarial, and anti-arrhythmic activity (Cordell 2008). Two MIAs, vincristine and vinblastine are known for their powerful anti-cancerous property and are majorly obtained from Madagascar periwinkle (Mukherjee, Basu et al. 2001). Many MIAs have gained attraction for synthetic methodologies due to their complex structures and promising bioactivity (Pfaffenbach and Gaich 2017). One of the most studied MIA pathway is the biosynthesis of Ajamaline from *Raowolfia serpentina*. An integrated root transcripotome and metabolome analysis led to the functional characterization of two candidate genes, orthologous to previously characterized γ-tocopherol-like N-methyltransferase (γ-TLMT) and involved in indole and side chain N-methylation of ajamaline. In addition, *N*-methyltransferase (γ-TLMT) transcript expression and its enzymatic activity was consisted with ajamaline alkaloid accumulation (Cázares-Flores, Levac et al. 2016). Major subgroups of monoterpene alkaloids according to their carbon structure, are shown in figure 1.8.

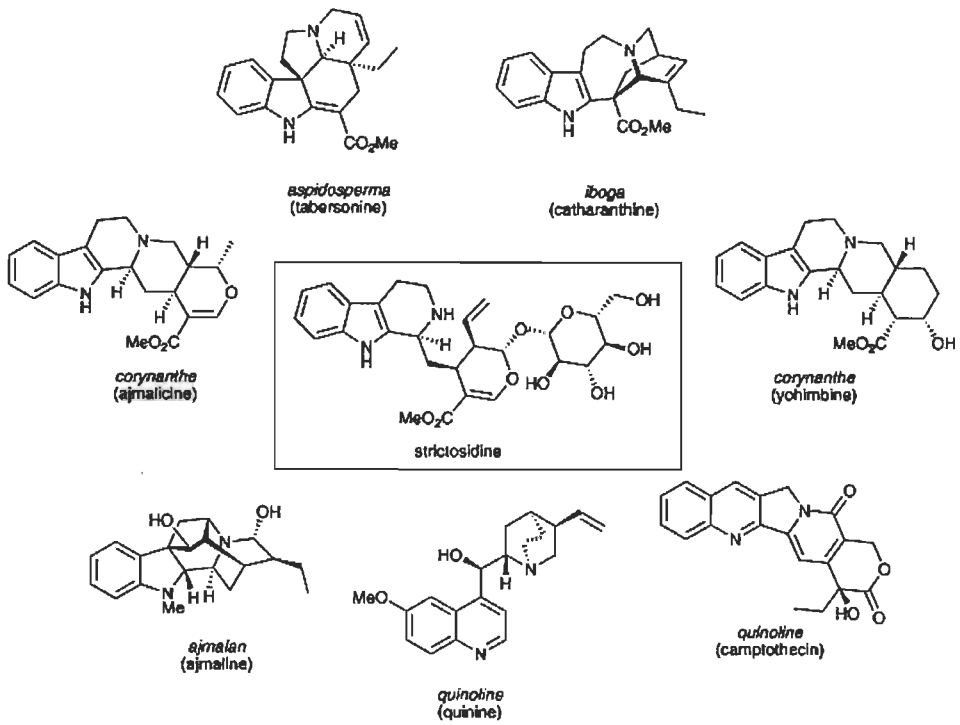


Figure 1.8 Different type of Monoterpene indole alkaloids.
Source (O'Connor and Maresh 2006).

1.4.6 Isoquinoline alkaloids

Isoquinoline Alkaloids (IAs) comprise one of the largest group of specialized metabolites. These alkaloids are synthesized from amino acid precursors, phenylalanine and tyrosine and possess isoquinoline or tetrahydroisoquinoline as basic ring structure (Grycová, Dostál et al. 2007). IAs are largely distributed among plants of Papaveraceae, Fumariaceae, Menispermaceae and Ranunculaceae family. This special category of alkaloids includes various structurally homogenous group, which arises due to different degree of oxygenation, intramolecular rearrangements, their distribution and presence of extra rings. The eight main subgroups are benzylisoquinoline, aporphine, protoberberine, benzo[c]phenanthridine, protopine, phthalideisoquinoline, phenethylisoquinoline (amayllidaceae), morphinan, and emetine alkaloids (Kukula-Koch and Widelski 2017). Among several groups benzylisoquinoline alkaloids and phenethylisoquinoline (amaryllidaceae alkaloids) are discussed below in detail.

1.4.6.1 Benzylisoquinoline alkaloids

Benzylisoquinoline alkaloids (BIAs) are a large group of around 2,500 structurally diverse nitrogen-containing compounds (Fig. 1.8) (Desgagné-Penix and Facchini 2011). They are synthesized from two L-tyrosine derivatives, dopamine and 4-hydroxyphenylacetaldehyde (4-HPAA) and are produced by 90% of the plants of several families including Papaveraceae, Ranunculaceae, Berberidaceae, Fumariaceae, and Menispermaceae (Facchini and De Luca 1994, Ziegler and Facchini 2008). Some important BIAs possess a potent pharmacological properties such as analgesics morphine and codeine, a muscle relaxant tubocurarine, an anti-microbial agents sanguinarine and berberine (Desgagné-Penix and Facchini 2011). The dried latex of of *Papaver somniferum*, also known as 'opium' was among the initial drugs used by humans as it contained analgesic morphine and codeine (Desgagné-Penix and Facchini 2011).

BIAs have also been extensively studied for their synthesis and storage in plants. Different studies on BIA producing plants have demonstrated that biosynthetic genes and enzymes locate in companion cells and sieve elements and synthesize BIAs. Once synthesized, toxic BIAs are finally transported to laticifers for the storage (Bird, Franceschi et al. 2003, Samanani, Alcantara et al. 2006, Lee and Facchini 2010, Beaudoin and Facchini 2014).

Recently, with the aid of OMICS tools, Virus induced gene silencing of *RNMT* transcript encoding reticuline *N*-methyltransferase revealed changes in alkaloid profile in roots compared to control plants and its role in magniflorine biosynthesis. Magnoflorine abundance was significantly decreased and corytuberine (intermediate precursor for magnoflorine) concentration was high in *RNMT*-silenced plants (Morris and Facchini 2016). Similiarly, correlation of glaucine alkaloid abundance and *glaucium flavum O-methyltransferase* (*GFLOMTs)* transcripts expression in aerial organs of *Glaucium flavum* revealed the role of *GFLOMT1, GFLOMT2* and *GFLOMT6* in glaucine formation (Chang, Hagel et al. 2015).

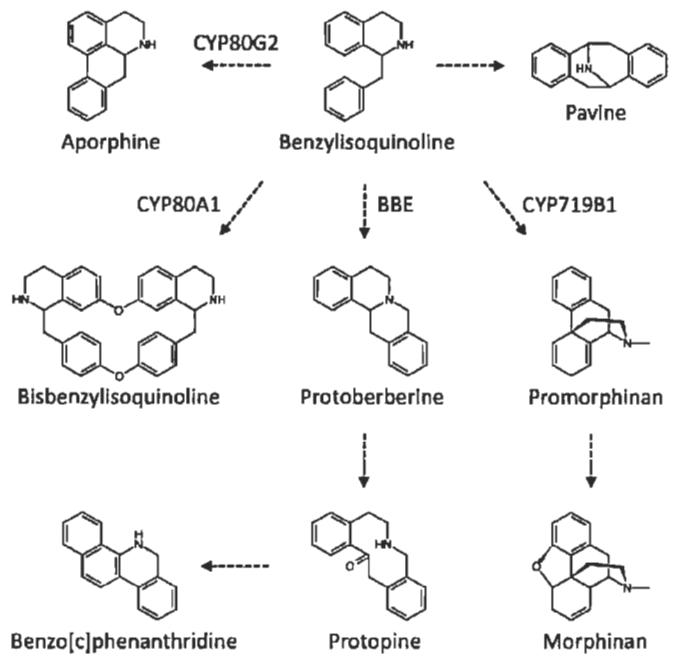


Figure 1.9 Example of structurally diverse BIAs derived from basic benzylisoquinoline skeleton. Source (Farrow, Hagel et al. 2012).

1.4.6.2 Amaryllidaceae alkaloids

Amaryllidaceae family plants are bulbous, monocots, perennial flowering plants of order Asparagales. They are distributed throughout the tropical and subtropical areas of the world (Chase, Reveal et al. 2009). This family consists of more than 1,100 species and 75 generas and approximately 500 AAs with different structures have been identified until now (Jin 2013). Four main regions of the world showing Amaryllidacaeae distribution are Eurasia, Australia, America, and Africa (Fig. 1.10) (Jin 2013).

Figure 1.10 Geographical distribution of Amaryllidaceae family plants (shaded black). Source (Jin and Xu 2013).

Many species belonging to genus *Amaryllis*, *Narcissus* (Daffodils) and *Galanthus* (snowdrops) are well known for their ornamental and pharmacological properties because of their large and actinomorphic flowers. The flowers are usually star-like or trumpet-shaped with a wide range of colors from red and orange to white (Meerow and Snijman 1998). Different developmental stages of *Narcissus* species plant are shown in figure 1.11.

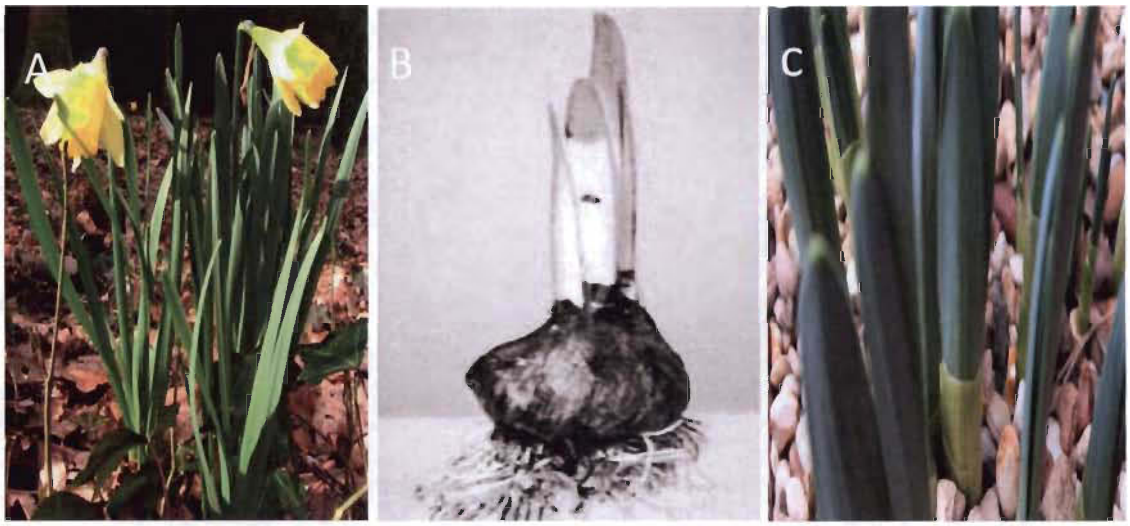


Figure 1.11 Images of *Narcissus spp. pseudonarcissus* (A) *Narcissus* plants flowering in late-march (B*) Narcissus* bulb with roots and shoot (C) *Narcissus* shoot emerging with sheathed leaves. Source: (wpnature.com, www.britannica/plant/amaryllidaceae).

The ancient evidence of their remedial usage is recorded from 4th century B.C. when Hippocrates used *Narcissus* oils for treating uterine tumors (Kornienko and Evidente 2008). Moreover, *Narcissus poeticus* L. was used by Greeks as a folk medicine which is now discovered to contain around 0.12 g of narciclasine per kg in bulbs (Evidente and Kornienko 2009).

1.4.6.2.1 Biological activities

Amaryllidaceae plants are a promising source of bioactive natural AAs with diverse structures and functions including anti-cancerous and cytotoxic properties (Table 1.3). A study conducted on narciclasine shows that it promotes apoptosis mediated cytotoxic effect only on human cancer cell lines by triggering the activation of caspases-8 and caspases-10 of death receptor pathway in human MCF-7 breast and PC-3 prostate carcinoma cells. Normal fibroblast cells were 250 fold less affected by narciclasine and

therefore, no apoptosis was observed in these cells (Dumont, Ingrassia et al. 2007). Pretazettine hydrochloride (PTZ) acts against *Corynebacterium fascians* and also inhibits *Cryptococcus neoformans,* the pathogenic yeast (Pettit, Melody et al. 2002). Evaluation of narciclasine, lycoricidine, pancraristatin, 7-deoxypancratistatin, isonarciclasine, *cis* and *trans*-dihydronarciclasine, lycorine and pretazettine shows their anti-viral activity against RNA-containing flaviviruses, bunyaviruses, Punta Toro and Rift Valley fever virus (Gabrielsen, Monath et al. 1992). Narciclasine displayed melanoma cells growth inhibition in 30-100 nM concentrations and potentially binds to elongation factor, eEF1A a target to comat melanomas and causes remarkable disorganization of actin cytoskeleton and cytokinesis impairment (Van Goietsenoven, Hutton et al. 2010).

Another well-known AAs, galanthamine was extracted from *Galanthus* (snowdrops) and was used for treating *Poliomyelitis* and *Myasthenia gravis* during initial years of its development in the 1950's (Plaitakis and Duvoisin 1983). It was during 1980's that the researchers examined the use of galanthamine in symptomatic treatment of Alzheimer's disease (AD) (Heinrich and Teoh 2004). The clinical development of galanthamine as an Alzheimer's treatment drug occurred mainly in the period of 1990's (Heinrich and Teoh 2004). As a result in 1996, it was first launched as 'NIVALIN' by Sanochemia Pharmazeutika as AD treatment drug (Heinrich and Teoh 2004). At present, galanthamine is obtained from few species of *Narcissus, Galanthus,* and *Leucojum* for the treatment of AD symptoms as it is a long-acting, reversible, selective and competitive acetylcholinesterase enzyme inhibitor. It has a potential of crossing a blood-brain barrier which is a crucial requirement for treating neurological conditions using drugs (Maelicke 2017). Galanthamine could also stimulate pre- and post-synaptic nicotinic receptors, thereby increasing the release of neurotransmitters and neuronal function (Schilström, 2007).

An AAs, narwedine, has shown respiration stimulating activity. It increases the amplitude and decreases the frequency of cardiac contractions and might play an important role in reducing blood loss during surgery (Refaat, Kamel et al. 2013).

AA	Group	Plant source	Biological activity
Galanthamine	*para-ortho'*	*Narcissus pseudonarcissus* *Leucojum aestivum* *Galanthus niveali*	acetylcholinesterase inhibitor, treatment for Alzheimer's disease
Narciclasine	*para-para'*	*Narcissus spp.* *Leucojum aestivum* *Lycoris radiata,* *Pancratium maritimum* *Hymenocallis expansa*	Cytotoxic
Montanine	*papa-para'*	*Hippeastrum vittatum*	anxiolytic, antidepressant,
Pretazettine	*para-para'*	*Narcissus spp.*	Anti-viral, cytotoxic, Analgesic
Crinine	*para-para'*	*Crinum spp.* *Galanthus reginae–olgae* *Brunsvigia gregaria* *Pancratinum maritimum*	cytotoxic, hallucinogenic
Haemanthamine	*para-para'*	*Narcissus pseudonarcissus* *Narcissus primigenius* *Pancratinum maritimum*	cytotoxic, analgesic, anti-malarial, anti-retroviral, fungicide hypotensive
Vittatine	*para-para'*	*Narcissus pseudonarcissus* *Crinum bulbispermum*	analgesic, anti-microbial, cytotoxic
Galanthine	*ortho-para'*	*Galanthus elewesii* *Lycoris sunguinea*	Cytotoxic
Narcissidine	*ortho-para'*	*Leucojium autumnale*	Cytotoxic
Lycorine	*ortho-para'*	*Narcissus spp.* *Leucojum aestivum* *Leucojium autumnale* *Galanthus reginae–olgae* *Galanthus elewesii* *Hippeastrum equestre*	Anti-viral, analgesic, anti-microbial, fungicide, anti-malarial, anti-inflammatory, cytotoxic
Homolycorine	*ortho-para'*	*Narcissus primigenius* *Narcissus pseudonarcissus*	cytotoxic, anti-viral anti-malarial

Table 1.3 **Examples of few Amaryllidaceae alkaloid, their plant source and biological activities.** Source (Singh and Desgagne, 2015).

A study conducted on lycorine and its related compounds displayed different levels of anti-tumor activity against four cancer cell lines. Lycorine has shown fifteen times more anti-proliferative activity against cancer cells compared to a normal cell, when applied in a micromolar concentration (Lamoral-Theys, Andolfi et al. 2009). Lycorine has

also been reported to display promising anti-cancer properties against various cancer cells including lymphoma, carcinoma, melanoma, multiple myloma, leukemia, and human Hs683 anaplastic oligodendroglioma cell lines (Wang, Yuan et al. 2014). Lycorine has also been reported in decreasing proliferation, migration, invasion and survival of prostate cancer cell lines (Hu, Peng et al. 2015). A recent study clinical trial of lycorine conducted on Renal cell carcinoma (RCC) has demonstrated a promising growth inhibiting property by arresting cell-cycle. Moreover, a synergistic treatment of Lycorine hydrochloride and anti-CTLA-4 led to loss of tumor weight and lung metastasis (Li, Xu et al. 2017). Along with lycorine, ungeremine, galanthine, zephgrabetaine, haemanthamine, hamayne and tortuosine were isolated from *Zephyranthus grandiflora* for cytoxicity investigation against C-6 rat glioma and CHO-K1 Chinese hamster ovary cells, which revealed lycorine as the most powerful at concentration lower than 10ug/ml (Katoch, Kumar et al. 2013). Apart from anti-tumor activity, lycorine also acts as fungicide against *Candida albicans* (Sundarasekar, Sahgal et al. 2012). Homoloycorine and hippeastrine display antiretroviral and anti-viral activities respectively (Szlávik, Gyuris et al. 2004). An AAs, bulbispermine has shown cytotoxic effect against HL-60 cells and also reported to inhibit the proliferation of glioblastoma cells (Cimmino, Masi et al. 2017).

Haemanthamine and haemanthidine, belonging to the crinine group of AAs, have been studied extensively for their cytotoxic properties against cancer cell lines (Havelek, Seifrtova et al. 2014). A dose-dependent cytotoxicity assay performed using haemanthamine, haemanthidine and lycorine against Caco-2, FHs 74 Int and HT-29 cells demonstrates a significant decrease in their growth (Fig. 1.12). An AA, vittatine also exhibits cytotoxic effects against HT29 colon adenocarcinoma and lung carcinoma (Silva, De Andrade et al. 2008). Another AA, ungeremine was studied for its potential in pest management and was detected with strong bactericidal against few pathogenic bacteria species (Schrader, Andolfi et al. 2010). Montanine type AAs were found to act against *Staphyloccocus aureus* and *Pseudomonas aeruginosa* (Evidente, Andolfi et al. 2004). In addition, clivatine and nobilisine possess anti-bacterial activity against gram-positive bacteria *S. aureus* and anti-fungal activity against *C. albicans* (Cimmino, Masi et al. 2017). A novel AA named *N*-methylhemeanthidine chloride shows increased cytotoxicity

more compared to gemcitabine (chemotherapeutic agent) in pancreatic cancer cells where it inhibits AKT signaling pathway (Guo, Yao et al. 2014).

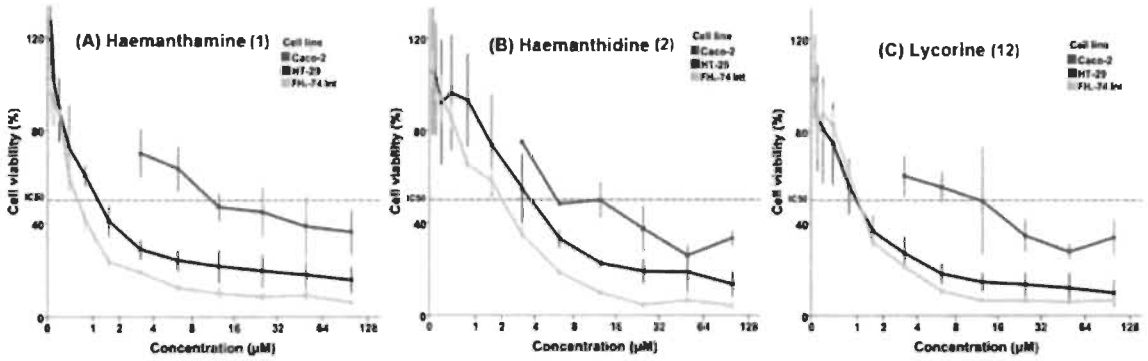


Figure 1.12 Dose–response cytotoxicity curves determined for most active alkaloids. (A) haemanthamine (1), (B) haemanthidine (2) and (C) lycorine (12) using MTT assays against Caco-2 and HT-29 cells at concentrations ranging from 0.01 to 100mM/mL and against FHs 74 Int cells at concentrations ranging from 3.12 to 100mM/mL. Source (Doskočil, Hošťálková et al. 2015).

1.4.6.2.2 Amaryllidaceae alkaloids biosynthesis

Though an extensive study on AAs pharmacological role has been performed, the knowledge of their biosynthesis pathway is still incomplete (Singh and Desgagné-Penix 2014). Amaryllidaceae alkaloids have a complex biosynthesis pathway, starting from two amino acid precursors, L-tyrosine and L-phenylalanine. These precursors undergo a sequence of chemical reactions to form intermediate precursors, 3,4-dihydroxybenzaldehyde (3,4-DHBA) and tyramine. Surprisingly, the reactions involved in the formation of 3,4-dihydroxybenzaldehyde is shared from phenylpropanoid biosynthesis pathway whereas synthesis of tyramine is analogous to an initial step of isoquinoline pathway (Singh and Desgagné-Penix 2014). Further, condensation and reduction of these intermediate results in norbelladine synthesis, a central intermediate of AAs biosynthesis pathway (Figure 1.13). Initial reaction step includes phenylalanine ammonia lyase (PAL) enzyme that catalyzes the removal of ammonia group to produce *trans*-cinnamic acid. Recently, two *PAL* isoforms, *PAL1* and *PAL2* were isolated and

characterized from *Lycoris radiata* and the expression of *PAL1* correlated with galanthamine accumulation, suggesting its involvement in AAs biosynthesis (Jiang, Xia et al. 2011, Jiang, Xia et al. 2013). The other amino acid precursor, L-tyrosine is decarboxylated by tyrosine decarboxylase (TYDC) to form tyramine. PAL and TYDC are two unique regulatory enzymes, allowing the conversion of primary metabolites into specialized metabolites (Singh and Desgagné-Penix 2014). In the biosynthesis of isoquinoline alkaloids such as analgesics morphine and codeine, TYDC was found to maintain an adequate amount of tyramine (Facchini, Huber-Allanach et al. 2000, Desgagné-Penix and Facchini 2011). Cinnamate-4- hydroxylase (C4H) and coumarate-3-hydroxylase (C3H) are two CYP450 enzymes, catalyzing two hydroxylation reaction and removal of two carbon atoms to form 3,4-dihydroxybenzaldehyde. The enzymes involved in the biosynthesis of intermediate precursors still requires a deep study but previous feeding experiment in *N. pseudonarcissus* and *Lycoris aurea* transcriptome data mining shows the existence of AAs intermediates and *PAL, CYP450, OMT, NMT* and *TYDC* transcripts (Wang, Xu et al. 2013).

Further, a reaction between 3,4-DHBA and tyramine is an example of a condensation reaction between an aldehyde and an amine resulting in the formation of Schiff's base to yield norbelladine (Dewick 2009) (Ghosal, Shanthy et al. 1988). This reaction allows entry of primary metabolites into AAs pathway and is commonly observed in plant alkaloid biosynthesis (Sobarzo-Sanchez 2015, Singh and Desgagné-Penix 2017). It is also known as the first committed step involving a novel enzyme which catalyzes the coupling of two intermediate precursors. For example, condensation of two intermediates, 4-hydroxyphenylacetaldehyde (4-HPAA) and dopamine in benzylisoquinoline alkaloid (BIA) biosynthesis by a Pictet-Spengler enzyme, catalyzing the first committed step of BIA biosynthesis pathway, is commonly known as norcoclaurine synthase (Samanani, Liscombe et al. 2004, Minami, Dubouzet et al. 2007, Lee and Facchini 2010). Downstream in the biosynthesis pathway, norbelladine is methylated by *norbelladine 4'-O-methyltransferase* (*NpN4OMT*). The expression of *NpN4OMT* was in coordination with galanthamine accumulation in leaf, inflorescence and bulb tissues (Kilgore, Augustin et al. 2014). 4'-*O*-methylnorbelladine is further cyclized via three modes of C-C oxidative

phenol coupling known as *para-para'*, *ortho-para'* and *para-ortho'*. The representative alkaloid of each category includes galanthamine (*para-ortho'*), haemanthamine (*para-para'*), and lycorine (*ortho-para'*) (Figure 1.13). *CYP96T1*, was the first phenol coupling enzyme characterized from monocot (*Narcissus* sp. *aff. pseudonarcissus, Galanthus* sp., and *Galanthus elwesii*) by comparative transcriptomics and co-expression analysis. It utilizes 4'-*O*-methylnorbelladine to produce (10b*R*,4a*S*)-noroxomaritidine and (10b*S*,4a*R*)-noroxomaritidine (Kilgore, Augustin et al. 2016). The three different kinds of C-C phenol coupling reactions results in a big diversity of AAs. Each plant variety and cultivar represent a different AAs profile. Nonetheless, *noroxomaritidine reductase (NR)* was characterized whose function is to reduce C-C double bond downstream in the pathway to form oxomaritinamine from noroxomaritidine (Kilgore, Holland et al. 2016). In conclusion, only a few genes have been identified in the biosynthetic pathway leading to AA. Further studies are required to expand the knowledge on AAs biosynthetic genes/enzymes and to provide tools for biotechnological and synthetic biology applications.

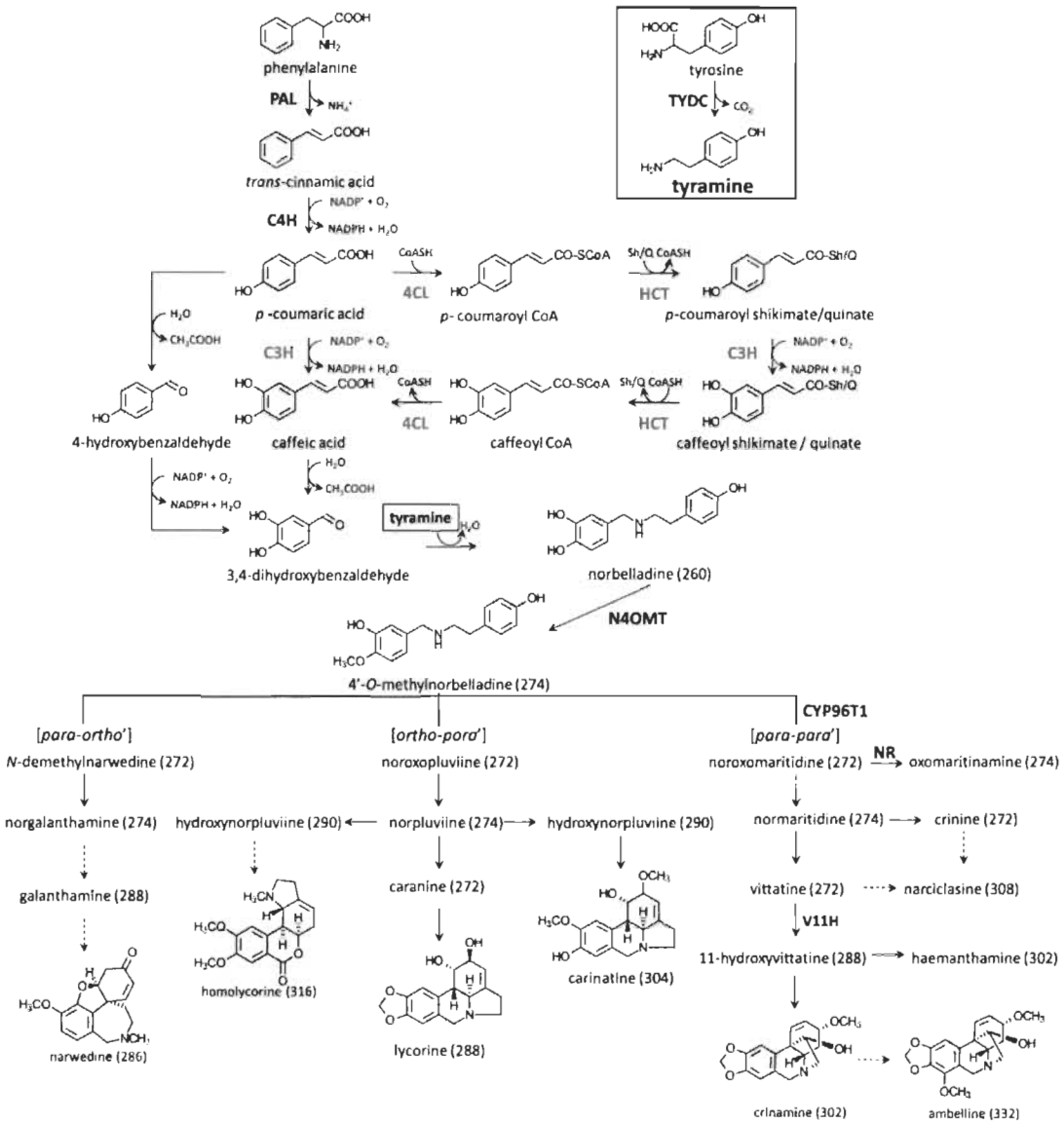


Figure 1.13 **Proposed pathway leading to multiple Amaryllidaceae Alkaloids.** Enzymes for which corresponding genes have been isolated from Amaryllidaceae are shown in bold black whereas from other plants are shown in bold grey. Broken arrow represents more than one biochemical reaction. Number in parenthesis corresponds to molecular ion m/z of [M +H] of alkaloid in positive MS mode detection. Abbreviations: *PAL*, phenylalanine ammonia lyase; *C4H*, *trans*-cinnamate hydroxylase (CYP73A1); *4CL*, 4-hydroxycinnamoyl CoA ligase; *HCT*, hydroxycinnamoyl transferase; *C3H*, p-coumarate hydroxylase (CYP98A3); *TYDC*, tyramine decarboxylase; *N4OMT*, norbelladine 4-*O*-methyltransferase; *NR*, noroxomaritidine reductase; *V11H*, vittatine 11-hydroxylase; *Sh/Q*, shikimate/quinate. Source (Singh A and Desgagne-Penix I, 2017)

1.5 Discovery of unknown genes using systems biology

Unlike a single linear pathway generating a single product, specialized metabolites originate from a long, complex biosynthesis network synthesizing several compounds and thus, the discovery of all the biosynthetic enzymes becomes a challenging (Hagel, Mandal et al. 2015). Moreover, the limited knowledge on alkaloid biosynthesis regulatory mechanisms is due to the lack of genomic resources, which could be improved by using powerful systems biology tools. This interdisciplinary approach has become popular recently for characterizing targeted genes to understand multiple components working in synergy to make a functional biological system (Riaño-Pachón, Nagel et al. 2008, Muranaka and Saito 2013). This data-mining framework combines metabolomics, bioinformatics, transcriptomics, genomics and high throughput experiments to provide new information about the biological process (Fig. 1.14). One of the early studies that attempted the combination of omics datasets for interpreting plants biological systems, used transcriptome and metabolome data of potato tubers. They revealed high transcripts-metabolite correlation and proposed this approach for functional characterization of candidate genes (Urbanczyk-Wochniak, Luedemann et al. 2003). Another study conducted on *Arabidopsis thaliana* also used similar transcriptome-metabolome data integration to establish a gene-to-metabolite network for understanding a response of a plant subjected to a nutritional stress (Hirai, Yano et al. 2004). Also, a comprehensive metabolome and transcriptome analysis of *Arabidopsis thaliana* over-expressing the *PAP1*gene which encodes MYB transcription factor. This study identified eight novel anthocynins and thirty-eight induced genes in *PAP1* overexpressed lines. They also identified two glycosyltransferases encoding as flavonoid 3-*O*-glucosyltransferase and anthocyanin 5-*O*-glucosyltransferase(Tohge, Nishiyama et al. 2005).

Transcriptome study gives the initial representation of plausible active processes undergoing in a biological system under a given condition. Next-generation sequencing (NGS) technologies, such as SOLiD, Illumina platform with MiSeq, HiSeq and HiSeq X provides large EST database compared to Roche 454 pyrosequencing, for identification of gene transcripts (Kilgore and Kutchan 2016). It is considered as a vital method for gene

identification of non-model pants as it doesn't require a prior genomic knowledge (Han, Rai et al. 2016). These new sequencing technologies enable sequencing at an unprecedented speed because they generate thousands to millions of sequencing reaction in parallel, enabling a full sequence coverage (Van Dijk, Auger et al. 2014). Currently, highest throughput per run is achieved by lllumina along with the cheapest per-base cost of $0.02/million base (Liu, Li et al. 2012). The data generated provide many useful information such as low and high abundant transcripts, differential expression pattern, unique transcripts, exon-intron boundaries, simple sequence repeats and alternate splicing (Martin, Fei et al. 2013).

In addition, metabolomics has also been an intrinsic part of systems biology. This tool has served an essential role in the identification of several medicinal compounds of a plant such as paclitaxel, artemisinin, vincristine, and vinblastine, camptothecin and morphine (Rai, Saito et al. 2017). It helped in gathering information about the spatial and temporal distribution of phytochemicals and their intermediate (Rai and Saito 2016). For example, metabolic analysis of twenty non-model benzylisoquinoline alkaloids (BIA) producing species was performed by applying multi-platform approach using NMR, direct flow injection (DFI)-MS/MS, HPLC-UV and LC-MS/MS to portray an entire view of different metabolic processes and their distribution across the species. This study enabled the discovery of metabolites including 110 lipid derivates and 15 phenolic compounds (Hagel, Mandal et al. 2015).

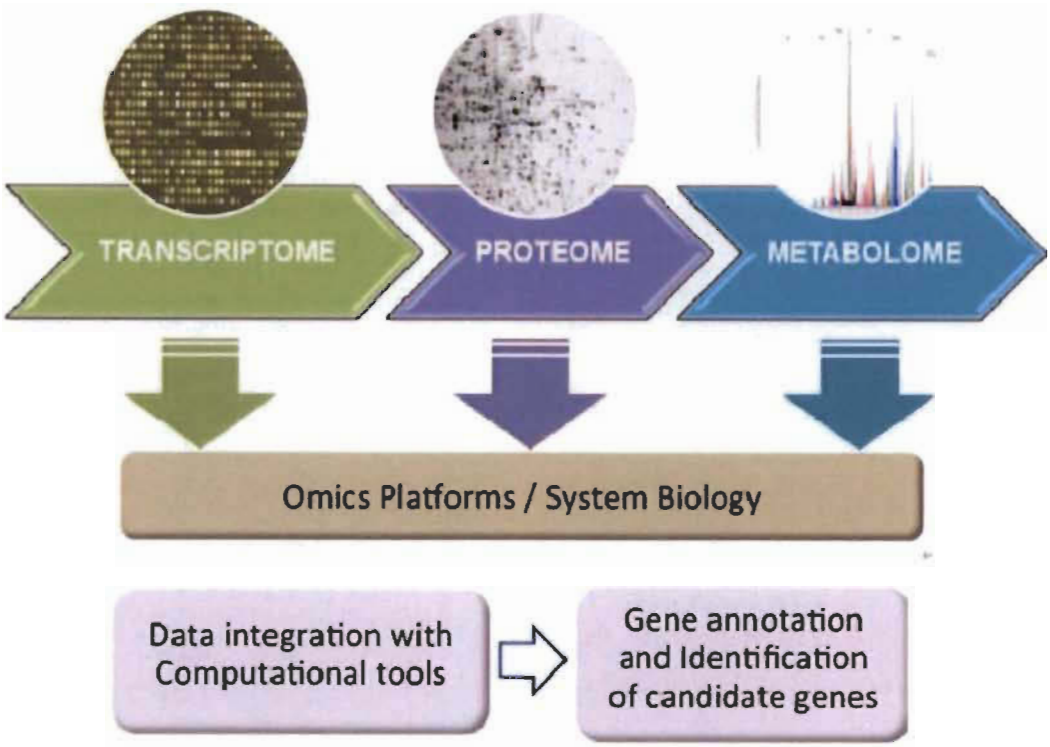


Figure 1.14 **Application of omics platform by an integrated study of different databases (transcriptomics, proteomics, and metabolomics) for functional characterization of candidate genes.** Source (Ben Amar, Daldoul et al. 2016).

Therefore, systems biology is an original approach for exploring genetic information behind the biosynthesis of valuable medicinal AAs in *N. pseudonarcissus* 'King Alfred'. It will bring in the knowledge that would allow the development of transgenic plants with enhanced production of interesting AAs and pave a way for synthetic biology to scale-up their productiion. As in the case of BIAs, identification of a large number of genes through omics tools facilitated the reconstruction of a pathway leading to the production of (S)-reticuline in *E. coli* (Diamond and Desgagné-Penix 2016).

1.6 Importance, hypotheses, and objectives

In comparison to the extensive research conducted on AAs for the study of their therapeutic properties, very little is known about their metabolism, which always attracted

researchers for their gene identification and characterization. Moreover, these alkaloids have limited availability because of their low level of production in plants. At present, among several AAs, galanthamine is the only AAs that is used clinically and usually extracted from *Leucojum* and *Narcissus* cultivated plant species (Takos and Rook 2013). Most AAs available in the market are obtained from their natural sources. In addition, it is difficult to implement molecular biology protocols on non-model plants including Amaryllidaceaes for gene identification and characterization because of lack of availability of genomic information.

To enable high-scale AAs production using synthetic biology, identification and characterization of AAs biosynthesis genes by integrated transcriptomics and metabolomics analysis is prerequisite. In this work, we selected *N.pseudonarcissus* 'King Alfred' based on its medicinal AAs history. Previous studies on *N. pseudonarcissus* 'King Alfred' have shown the presence of therapeutic AAs, narcidine, crinine, lycorenine and galanthamine as an essential alkaloid constituent in bulbs (Tojo 1991). A radiotracer experiments have shown the conversion of norpluviine into lycorenine–type alkaloid, in *N. pseudonarcissus* 'King Alfred' (Harken, Christensen et al. 1976). GC-MS analysis conducted on plants of *Narcissus* genus detected 134 alkaloids with lycorine and homo-lycorine type as the predominant alkaloids (Berkov, Martínez-Francés et al. 2014). Another study on alkaloid profile of *Narcissus broussonetii* using GC-MS and LC-ESI-LTQ-Orbitrap-MS identified 23 alkaloids including lycorine, tazettine, homolycorine, 8-O-demethylhomolycorine, plicamine, pretazettine and galanthindole (de Andrade, Pigni et al. 2012). Quantitative ^{1}H NMR analysis was performed on bulb, leaf and root tissues of *N. pseudonarcissus* cv. Carlton during the growth season shows high alkaloid concentration in leaves before flowering. Low alkaloid content but high total alkaloid yield was reported from bulb due to its bigger biomass (Lubbe, Gude et al. 2013). GC-MS and HPTLC based investigation of alkaloid profile of *N.tazetta* and *N.papyraceus* tissues identified 58 alkaloids including tazettamide, sternbergine, 1-*O*-acetyllycorine, 2,11-didehydro-2-dehydroxylycorine. These alklaoids were reported first time in the plant of genus *Narcissus (Shawky, Abou-Donia et al. 2015)*. Also, GC-MS analysis of 101 ornamental varieties of *Narcissus* reported total 61 AAs and 2 protoalkaloids. Lycorine-

type were predominant in leaves and galanthamine-type in bulbs (Torras-Claveria, Berkov et al. 2014). All these studies confirm the presence of medicinal AAs in genus Narcissus. For our study, we performed a preliminary AAs quantitative analysis with different Amaryllidaceae species and cultivars (*N. pseudonarcissus* 'Dutch Master', *N. pseudonarcissus* 'King Alfred', *N. papyraceus*, *Galanthus elwesii, Leucojum aestivum, Crinum powellii* and *Galanthus nivalis*) using TLC and HPLC (data not shown). Many known (galanthamine, lycorine and narciclasine) and unknown compounds were detected in these different species. Based on the concentration and variety of AAs represented by each species, *N. pseudonarcissus* 'King Alfred' was selected for the further studies. Therefore, these results make it an interesting target for further investigation of AAs metabolism.

*Np*NBS was selected for gene identification and characterization because this gene was thought to be involved in the synthesis of norbelladine and catalyze the first committed step of AAs pathway. This intermediate provides a core structure to all the structurally diverse AAs. The proposed AAs pathway also shows the special condensation reaction between tyramine and 3,4-DHBA resulting in the formation of norbelladine. (Singh and Desgagné-Penix 2014). This study is required to develop a workflow that will be useful for new biosynthesis pathway genes identification from non-model plants.

1.6.1 Objective I: Understanding AAs metabolism using integrated transcriptomics and metabolomics in *Narcissus pseudonarcissus* 'King Alfred'.

For this research, we conducted our experiments on *Narcissus pseudonarcissus* 'King Alfred', a non-model plant belonging to *Narcissus* genus of Amaryllidaceae family. Previous integrated transcriptomics and metabolomics study on eight *Papaver sominferum* cultivars had revealed the genetic and metabolic factors responsible for differential BIA accumulation pattern among cultivars (Desgagné-Penix, Farrow et al. 2012). Similarly, to understand the plausible roles of proposed AAs biosynthesis genes in AAs biosynthesis, we hypothesize a correlation study between transcriptomic and metabolomic profile of *Narcissus pseudonarcissus* 'King Alfred' tissues. To test this hypothesis, we performed AAs profiling from different plant tissues (bulbs, roots, stems,

leaves, and flowers) using HPLC and UPLC-MS-qTOF and were able to confirm the presence of 14 known AAs. For transcriptomics, Illumina RNA sequencing and assembly was performed which resulted in 11,708 transcripts, including the complete sequences of all proposed AAs biosynthesis genes and their isoforms. Overall, we conducted a thorough comparative study of these datasets by integrated analysis, which provided a complete image of gene-to-metabolite networks in *Narcissus pseudonarcissus* 'King Alfred'.

1.6.2 Objective II: Cloning and characterization of *Norbelladine synthase*, a novel gene involved in norbelladine synthesis in *Narcissus pseudonarcissus* 'King Alfred'.

The proposed AAs pathway shows the condensation reaction between tyramine and 3,4-DHBA to form norbelladine, an entry compound of AAs biosynthesis pathway (Singh and Desgagné-Penix 2014). This reaction shares similiarity with a reaction between an amine and aldehyde in benzylisoquinoline alkaloid (BIA) biosynthesis pathway, catalyzed by an enzyme norcoclurine synthase (Samanani, Liscombe et al. 2004). Therefore, we conducted a *norcoclaurine synthase (NCS)* ortholog BLAST search in *N. pseudonarsissus* 'King Alfred' transcriptome to identify and select candidate genes. We investigated enzyme's phylogenetic relationship and in order to carry out its functional characterization, the selected *NpNBS* cDNA was isolated, cloned into a gateway pET301/CT-DEST destination vector. *Np*NBS protein expression was carried out in *E. coli* Rosetta™ (DE3) pLysS host strain and purified. *Np*NBS lacked N-terminal signal peptide and phylogenetically it is closer to *NCS* from other related species as it shares 41% homology to *Papaver somniferum Ps*NCS1and *Ps*NCS2. Its expression was found to be high in bulb compared to other tissues of the plant which was consistent with high AAs diversity in bulbs. LC-MS/MS analysis of enzyme assay confirmed that *Np*NBS is capable of producing norbelladine and plays a crucial role in AAs metabolism. The discovery of this enzyme supports our hypothesis of gene-metabolite correlation. This discovery will eventully aid the development of a synthetic biology platform for the high scale production of medicinally potent AAs.

Chapter II contains a study that presents a comprehensive transcriptome and metabolome analysis of *Narcissus pseudonarcissus* 'King Alfred' to understand AAs metabolism.

CHAPTER II

TRANSCRIPTOME AND METABOLOME PROFILING OF *NARCISSUS PSEUDONARCISSUS* 'KING ALFRED' REVEAL COMPONENTS OF AMARYLLIDACEAE ALKALOID METABOLISM

APARNA SINGH AND ISABEL DESGAGNE-PENIX

Published on 11th of December 2017 in *Scientific Reports*

2.1 Contribution

A.S. and I.D.P. conceived and designed the experiments. A.S. performed the experiments. A.S. and I.D.P. analyzed the data, prepared the figures, wrote and reviewed the manuscript.

2.2 Abstract

Amaryllidaceae alkaloids represent a diverse class of plant specialized metabolites and many display potent pharmacological activities. The AA metabolic pathway is poorly understood and resources are minimal. To enable AA pathway elucidation and novel biosynthetic enzymes discovery, we generated comprehensive metabolomic and corresponding transcriptomic datasets from different tissues of *Narcissus pseudonarcissus* 'King Alfred'. In this study, we performed untargeted UPLC-QTOF-MS metabolite analysis from different tissues, which generated exhaustive list of compounds, including several AAs, most predominant and diverse in bulbs. RNA sequencing of *N. pseudonarcissus* 'King Alfred' bulbs yielded 195,347 transcripts, after assembly. Top expressed genes belong to process like metabolism, survival, and defense including alkaloid biosynthetic genes. The transcriptome contained complete sequences for all proposed genes encoding AA-biosynthetic enzymes such as tyrosine decarboxylase

(*TYDC1* and *TYDC2*), phenylalanine ammonia-lyase (*PAL1* and *PAL2*) and phenolic acids hydroxylases (*C4H* and *C3H*) to name a few. Furthermore, transcriptome data were validated using RT-qPCR analysis and expression study in different tissues of *N. pseudonarcissus* 'King Alfred' was performed. Here, we present the first comprehensive metabolome and transcriptome study from *N. pseudonarcissus* 'King Alfred' providing invaluable resources for metabolic engineering and biotechnological applications.

2.3 Introduction

Plants produce a plethora of chemicals important for growth, development, and defense through the primary and specialized (*aka* secondary) metabolic pathways. The specialized metabolites, also named phytochemicals or natural products, are specific and restricted to some taxonomic groups and play pivotal roles in plant defense, protection, and survival (Dewick 2009). The medicinal properties of the Amaryllidaceae are owed to the presence of specialized metabolites, the Amaryllidaceae alkaloids (Grabherr, Haas et al.), which are specific to this plant family. Several AAs display pharmaceutical activities such as the anti-cancer narciclasine, the anti-parasitic haemanthamine, the anti-viral lycorine and the anti-acetylcholine esterase galanthamine (Thomsen and Kewitz 1990, López, Bastida et al. 2002, Sener, Orhan et al. 2003, Szlavik, Gyuris et al. 2004, Dumont, Ingrassia et al. 2007, Kornienko and Evidente 2008, He, Qu et al. 2015, Hotchandani and Desgagne-Penix 2017). Despite AAs being a huge therapeutic reserve, galanthamine is the only one used medically to treat the symptoms of Alzheimer's disease through its reversible acetylcholine esterase inhibition and its nicotinic receptor binding activity (Heinrich and Lee Teoh 2004, Barik, Dajas-Bailador et al. 2005). Currently, the knowledge of AA metabolism and regulation is limited and only a few genes encoding biosynthetic enzymes have been identified (Kilgore, Augustin et al. 2014, Kilgore, Augustin et al. 2016, Kilgore, Holland et al. 2016). A better understanding of AA biosynthesis will allow for the development of new cultivars or biotechnologies to help produce these valued phytochemicals. Additionally, over 600 AAs have been identified (Jin and Xu 2013), adding to the diversity of AA chemical structures and also demonstrating the complexity of their biosynthetic pathway.

Previous studies using radiolabeled precursors led to the biochemical elucidation of the initial steps in AA biosynthesis (Battersby, Fales et al. 1961, Barton, Kirby et al. 1963, Eichhorn, Takada et al. 1998). Despite the vast structural diversity, all AAs are derived from a common central intermediate norbelladine, which is formed through the condensation of 3,4-dihydroxybenzaldehyde (3,4-DHBA) and tyramine (Fig. 2.1). Similarities to several plant specialized metabolic pathways can be observed. For example, tyrosine is decarboxylated to tyramine (Fig. 2.1). In *Papaver sominferum*, this reaction is catalyzed by tyrosine decarboxylase (TYDC), a key regulating enzyme in the formation of isoquinoline alkaloids such as morphine and codeine (Facchini, Huber-Allanach et al. 2000, Hagel and Facchini 2005, Desgagné-Penix and Facchini 2011). The synthesis of 3,4-DHBA features steps similar to the phenylpropanoid pathway starting with the deamination of phenylalanine, a reaction catalyzed by phenylalanine ammonia-lyase *(PAL)* to generate *trans*-cinnamic acid (Fig. 2.1). *PAL* genes have been characterized from various species including Amaryllidaceae species(Jiang, Xia et al. 2011, Jiang, Xia et al. 2013). Next, cytochrome P450-dependent monooxygenases *(CYPs)* namely cinnamate-4-hydroxylase *(C4H)* and coumarate-3-hydroxylase *(C3H)* may be involved to hydroxylate *trans*-cinnamic acid and *p*-coumaric acid, respectively. From there, several routes have been hypothesized to yield 3,4-DHBA (Fig. 2.1). However, it is still not clear whether the chain shortening reaction occurs on *p*-coumaric acid or caffeic acid (Podstolski, Havkin-Frenkel et al. 2002, Gallage, Hansen et al. 2014, Yang, Barros-Rios et al. 2017).

The first committed step in AA biosynthesis starts with the coupling of the two precursors, defining the entry point of primary metabolites into AA biosynthetic pathway (Bastida, Berkov et al. 2011, Takos and Rook 2013, Singh and Desgagne-Penix 2014, Singh and Desgagné-Penix 2015). The combination of 3,4-DHBA and tyramine results in the formation of a Schiff base intermediate, which following reduction, yields norbelladine. Similar condensation reactions are described in other plant alkaloid biosynthetic pathways. For example, the first committed step in the formation of benzylisoquinoline alkaloids is the combination of 4-hydroxyphehylacetaldehyde and

dopamine catalyzed by norcoclaurine synthase *(NCS)* to produce norcoclaurine (Samanani, Liscombe et al. 2004, Minami, Dubouzet et al. 2007, Lee and Facchini 2010). Next, norbelladine is methylated to 4'-*O*-methylnorbelladine by the norbelladine 4'-*O*-methyltransferase (N4OMT) which was identified in *Narcissus* sp. *aff. pseudonarcissus*, and enzymatically characterized by heterologous expression in *Escherichia coli*(Kilgore, Augustin et al. 2014). The cyclization of 4'-*O*-normethylbelladine can occur by three different ways of intramolecular C-C oxidative phenol coupling named *para-ortho'*, *ortho-para'*, and *para-para*' that generates three backbone structures (Fig. 2.1). The three AA backbones form the basis of further alkaloid diversity. A complex network of enzymatic reactions exists to produce a spectrum of compounds that differs between species, varieties, and cultivars and even between the different tissues of the same plant. These biochemical modifications are achieved by a multitude of enzymes catalyzing various types of reactions, such as C-C and C-O bond formations, *O*- and *N*-methylations, demethylations, hydroxylations, oxidations and reductions (Fig. 2.1).

Although a few genes involved in AA metabolism in *Narcissus* have been isolated, several genes have not yet been identified. Transcriptomic analysis coupled with metabolite profiling has become a popular tool for the discovery of novel genes encoding proteins involved in alkaloid biosynthesis (Desgagne-Penix, Farrow et al. 2012, Xiao, Zhang et al. 2013, Hagel, Mandal et al. 2015, Hagel, Morris et al. 2015). Specifically, RNA sequencing using Illumina technology has been performed to analyze transcriptomes of various medicinal or nutritional plants, and the coupling with metabolomic analysis has greatly helped with the elucidation of several specialized metabolic pathways (Ge, Xiao et al. 2011, Gahlan, Singh et al. 2012, Hao, Ma et al. 2012, Hyun, Rim et al. 2012). For example, metabolite and genome-wide transcript profiling collectively yielded known and previously undescribed gene transcripts and metabolites associated with the terpenoid-indole alkaloids in *Cathanranthus roseus* (Rischer, Orešič et al. 2006). In addition, comparing Illumina sequencing based transcriptome profile with metabolic data led to a better understanding of the biosynthesis of anti-cancerous alkaloids, camptothecin and anthraquinones (Yamazaki, Mochida et al. 2013).

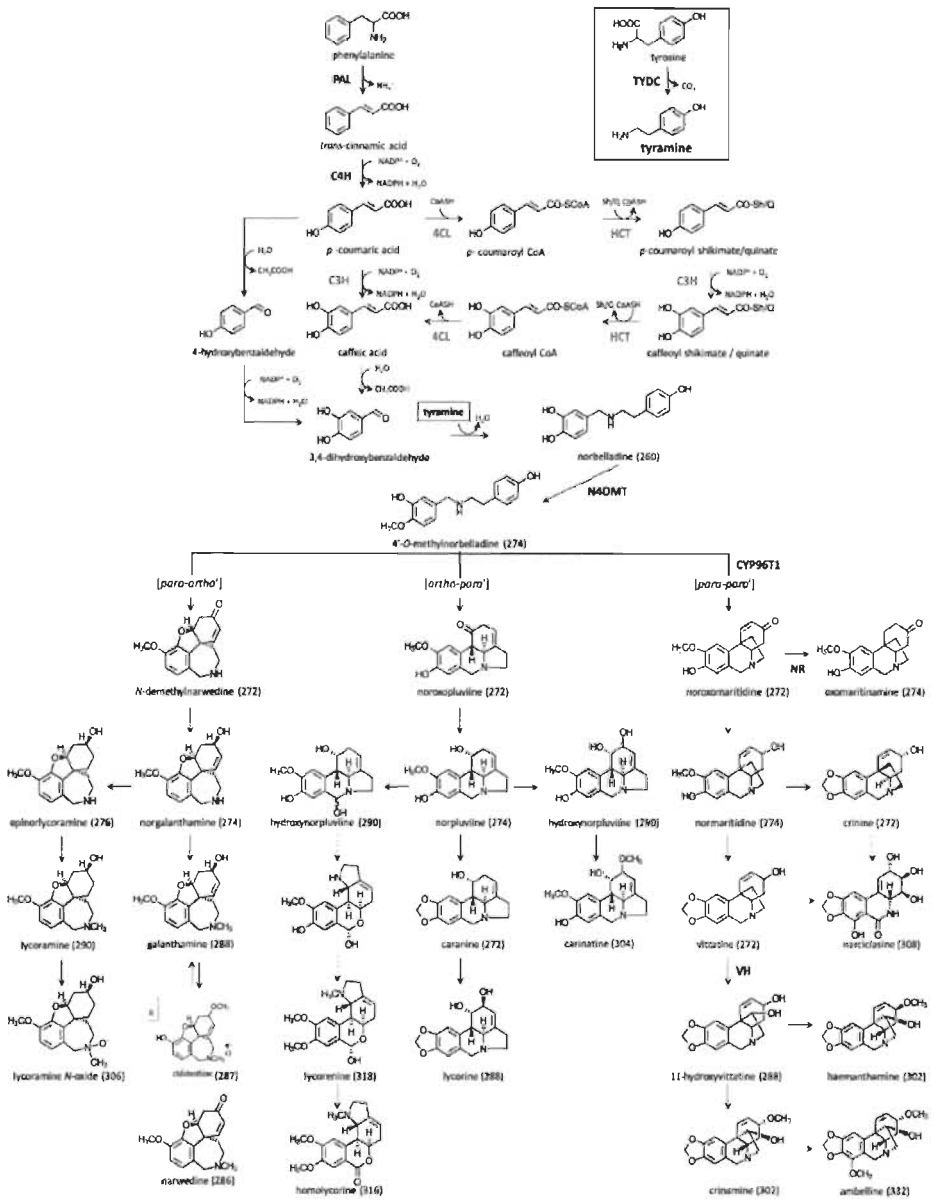


Figure 2.1 **Proposed biosynthetic pathway leading to multiple Amaryllidaceae alkaloids**. Enzymes for which corresponding genes have been isolated from Amaryllidaceae are shown in bold black whereas from other plants are shown in bold grey. Broken arrow represents more than one biochemical reaction. Number in parenthesis corresponds to molecular ion *m/z* of [M + H] of alkaloid in positive MS mode detection. Abbreviations: PAL, phenylalanine ammonia lyase; C4H, *trans*-cinnamate hydroxylase (CYP73A1); 4CL,4-hydroxycinnamoyl CoA ligase; HCT, hydroxycinnamoyl transferase; C3H, *p*-coumarate hydroxylase(CYP98A3); TYDC, tyramine decarboxylase; N4OMT, norbelladine 4-*O*-methyltransferase; NR,noroxomaritidine reductase; V11H, vittatine 11-hydroxylase; Sh/Q, shikimate/quinate.

Although RNA sequencing has been performed on few species of Amaryllidaceae including *Lycoris aurea*, *Narcissus spp.*, *Galanthus spp.*, and *Leucojum aestivum* (De Felice, Manfellotto et al. 2013, Wang, Xu et al. 2013, Kilgore, Augustin et al. 2014), AA metabolism still remains poorly understood. To enable pathway elucidation and novel enzyme discovery, we first performed metabolite profiling from different tissues of *N. pseudonarcissus* 'King Alfred' to identify the best tissues for RNA-seq. Leaves and bulbs showed high levels of AAs but since bulbs displayed more diversity, they were selected for transcriptomic analysis. Next, we created an extensive dataset for *N. pseudonarcissus* 'King Alfred'. Using Illumina HiSeq2000 sequencing, we obtained a deeply sequenced transcriptome with 73,081,603 reads that were assembled into 195,347 transcripts. Comprehensive analysis of the transcriptome indicated the presence of previously characterized (*e.g. N4OMT, CYP96T1, NR*) and proposed (*e.g. C4H, C3H, HCT, 4CL*) AA biosynthetic genes. We further validated our findings by performing expression analysis of each AA biosynthetic gene in different tissues of *N. pseudonarcissus* 'King Alfred' using RT-qPCR analysis. Used together, these unprecedented resources allow the assembly of a biochemical snapshot representing AA metabolism, guiding pathway elucidation and search efforts for new biosynthetic enzymes. Moreover, the availability of enzymes variants will expand the 'toolbox' essential to synthetic biology efforts (Diamond and Desgagne-Penix 2016).

2.4 Results

Alkaloid profiling of *Narcissus pseudonarcissus* 'King Alfred'. An optimized acid-base extraction was used to separate alkaloids from other organic compounds based on their acid-base properties. Alkaloid profiles (composition and concentration) from different extracted tissues (bulbs, roots, stems, leaves, and flowers) of *N. pseudonarcissus* 'King Alfred' were initially screened using chromatographic methods (Fig. 2.2). Thin-layer chromatography (TLC) showed qualitative differences among alkaloids from different tissues (Fig. 2.2A-B). We found that all plant parts contain various known and unknown compounds. Bulbs possess a larger number of diverse alkaloids, including large

spots on TLC with R_f corresponding to galanthamine (R_f 0.02) and lycorine (R_f 0.13) standards but not to narciclasine (R_f 0.08) (Fig. 2.2A-B). Stem extracts appear to contain fewer alkaloids than the other tissues. Roots, leaves and flowers also contain spots with R_f corresponding to galanthamine and lycorine (Fig. 2.2A-B). Spots with R_f corresponding to narciclasine were not detected in any plant samples. Thus, TLC analyses showed strong differences in the alkaloid profiles of the extracts from different tissues of *N. pseudonarcissus* 'King Alfred'.

To quantify the differences in alkaloid profiles, high-performance liquid chromatography (HPLC) with a photodiode array (PDA) detector analyses were performed. HPLC-PDA chromatograms confirmed differences among the extracts from different tissues. Peaks were detected at similar retention times (R_t) as those of AA standards lycorine (R_t 3.92 min), galanthamine (R_t 4.62 min) and narciclasine (R_t 7.75 min). For example, peaks for lycorine and galanthamine were detected in bulbs, roots and leaves (Fig. 2.2C), although in some case the absorption spectra were slightly different than those of standards. Also, peaks corresponding to narciclasine (R_t and absorption spectra) were not detected in any tissue tested, which correlates with TLC data. Interestingly, an unknown compound (R_t 7.85 min) was present in all tissue tested and was most abundant in leaves whereas compounds at R_t 6.31 and 16.92 min were two times more abundant in bulbs compared to other tissues (Fig. 2.2C-D). In addition, compounds at R_t 3.28, 9.22 and 13.95 min were only detected in bulbs, suggesting specificity for this tissue. Relative quantification of the alkaloid content of extracts from different tissues was done after normalization using papaverine as an the internal standard. Total AA levels were highest in leaves and were normalized to 100 %, followed by bulbs (84 %), roots (81 %) and stems (55 %) (Fig. 2.2D). Flowers (29 %) contained low alkaloid levels. Altogether, the results show different alkaloid profiles for the different tissues of *N. pseudonarcissus* 'King Alfred' with leaves containing the highest level of compounds followed by bulbs, each having high levels of numerous specific compounds.

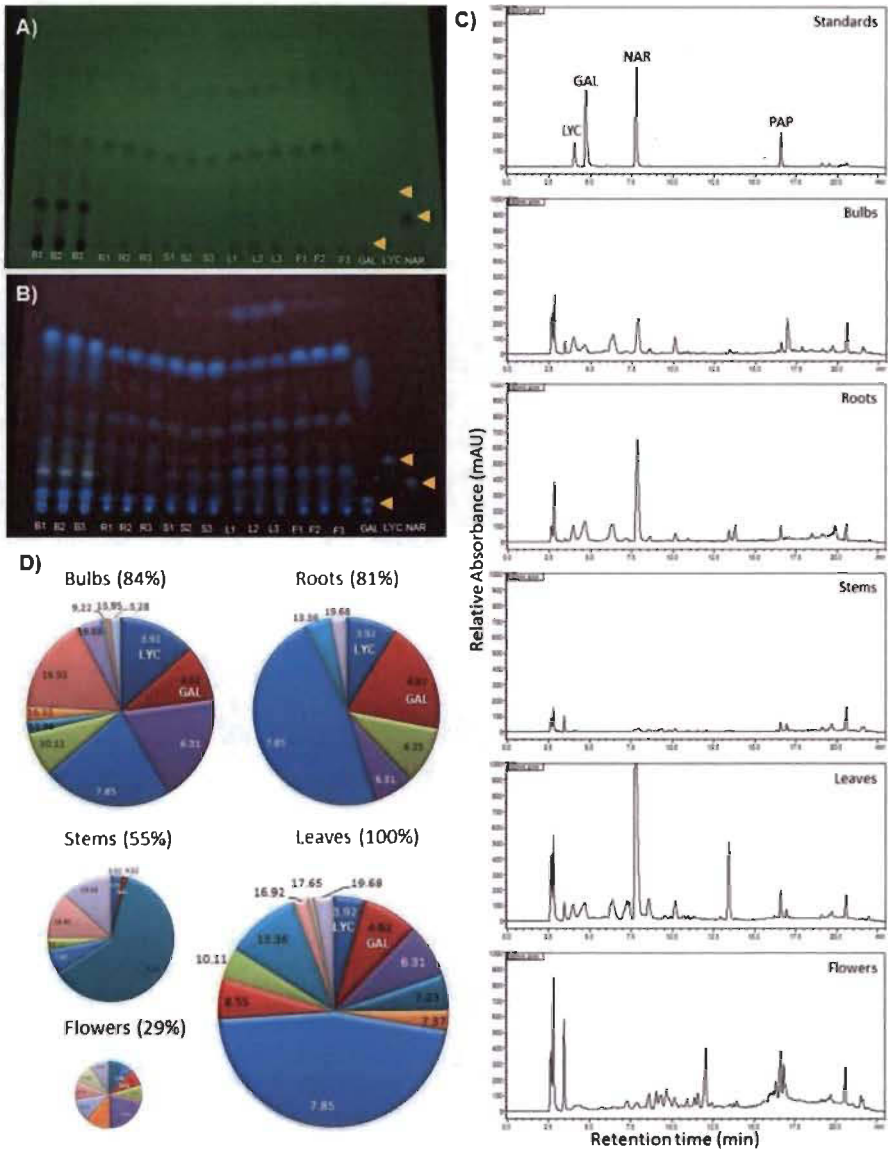


Figure 2.2 **Chromatographic analysis of compounds obtained from an acid-base extraction of different tissues of *N. pseudonarcissus* King Alfred.** (**A**,**B**) Thin-layer chromatography of extracts developed at 254 nm (**A**) and at 365 nm (**B**). (**C**) Representative HPLC chromatogram for known standards (top) such as lycorine (LYC), galanthamine (GAL), narciclasine (NAR) and papaverine (PAP) and from extracts of five different tissues of *N.pseudonarcissus* King Alfred. (**D**) Pie charts representing the relative abundance of major compounds in extracts of *N. pseudonarcissus* and normalized using the papaverine internal standard, determined by HPLC analysis. The number in each portion of the pie chart represents the retention time (Rt) of each compound.

Metabolite profiling by LC-MS. Untargeted UPLC-QTOF-MS profiling was performed to generate extensive mass lists corresponding to a wide variety of metabolites, including alkaloids present in our various extracts of *N. pseudonarcissus* 'King Alfred' tissues. Less overall number of ion masses were detected in analysis performed in negative ion mode analysis, between 4 and 12 (Table 2.1), with 4-hydroxybenzyldehyde was the most abundant compound present in bulbs and roots. Analysis performed in positive ion mode detected between 55 and 71 compounds including several alkaloids, fatty acids, organic acids, and sugars (Table 2.1). AAs detected using UPLC-QTOF-MS analyses in positive mode were extracted and compiled in Table 2.2 to identify and expand our targeted AA profile of our different extracts of *N. pseudonarcissus* 'King Alfred'. It must be noted that some metabolites are being listed twice with slightly different retention times or ion masses. There are several reasons explaining this. First, the identification provided is a proposed identification. For the same AA compound, different structural isomers may be present and separate on the column to yield different retention times. For slightly different Rt it may be epimers whereas larger Rt differences may indicate enantiomers. For example, crinamine [M+H] is a C-3 methoxy epimer of haemanthamine, thus sharing the same [M+H]. Another example related to the epimers vittatine and crinine also with the same masses. Secondly, the MZmine software identifies peaks via searches through custom and online database. Alkaloids are present but only few AAs are available. Thus, the potential identification of AA using these databases is lower than for well-known metabolites.

The most abundant AA detected was crinamine (R_t 6.07; [M+H] 302.1341) (Table 2.2), followed by lycorenine (R_t 5.36; [M+H] 318.1866) in leaves. The most abundant AA in flower, stem, root and bulb tissues was crinamine (R_t 6.12, [M+H] 302.1425). Lycorine and homolycorine were only detected in bulbs whereas epinorlycoramine was detected in all tissues (Table 2.2). Interestingly, 4'-*O*-methylnorbelladine was only detected in flower tissues.

N. pseudonarcissus 'King Alfred' tissue extracted	Number of ion masses	
	Positive ion mode (M+H)	Negative ion mode (M-H)
Bulbs	71	12
Roots	61	10
Stems	58	7
Leaves	71	5
Flowers	55	4

Table 2.1 Total number of ion masses identified by UPLC-QTOF-MS analysis of extracts of *N. pseudonarcissus* 'King Alfred' tissues using positive and negative ionization mode.

Exact ion mass (M+H)	Retention time (min)	Formula	Proposed identity	Area under the curve (AUC)					Phenol coupling type
				Bulbs	Roots	Stems	Leaves	Flowers	
276,1747	1,28	C16H21NO3	Epinorlycoramine						*para-ortho'*
290,1862	1,31	C17H23NO3	Lycoramine						*para-ortho'*
274,1602	1,34	C16H19NO3	Norgalanthamine						*para-ortho'*
288,1785	1,35	C17H21NO3	Galanthamine						*para-ortho'*
290,1944	1,35	C17H23NO3	Lycoramine						*para-ortho'*
274,1602	3,89	C16H19NO3	Norgalanthamine						*para-ortho'*
276,1747	3,90	C16H21NO3	Epinorlycoramine						*para-ortho'*
288,1785	4,38	C17H21NO3	Galanthamine						*para-ortho'*
290,1862	4,41	C17H23NO3	Lycoramine						*para-ortho'*
290,1944	4,42	C17H23NO3	Lycoramine						*para-ortho'*
318,1866	4,68	C18H23NO4	Lycorenine						*para-ortho'*
306,1930	4,89	C17H23NO4	Lycoramine *N*-oxide						*para-ortho'*
286,1615	5,71	C17H19NO3	Narwedine						*para-ortho'*
288,1703	1,35	C16H17NO4	Lycorine						*ortho-para'*
318,1866	4,12	C18H23NO4	Lycorenine						*ortho-para'*
288,1703	4,37	C16H17NO4	Lycorine						*ortho-para'*
318,1866	4,37	C18H23NO4	Lycorenine						*ortho-para'*
304,1729	4,86	C17H21NO4	Carinatine						*ortho-para'*
318,1866	5,05	C18H23NO4	Lycorenine						*ortho-para'*
318,2125	5,32	C18H23NO4	Lycorenine						*ortho-para'*
318,1866	5,36	C18H23NO4	Lycorenine						*ortho-para'*
318,1952	5,39	C18H23NO4	Lycorenine						*ortho-para'*
316,1702	6,32	C18H21NO4	Homolycorine						*ortho-para'*
272,1451	5,46	C16H17NO3	Vittatine						*para-para'*
302,1342	6,07	C17H19NO4	Crinamine						*para-para'*
302,1425	6,12	C17H19NO4	Crinamine						*para-para'*
302,1846	6,13	C17H19NO4	Crinamine						*para-para'*
302,1762	6,16	C17H19NO4	Crinamine						*para-para'*
308,0923	7,01	C14H13NO7	Narciclasine						*para-para'*
332,1983	7,36	C18H21NO5	Ambelline						*para-para'*
274,1602	3,92	C16H19NO3	4'-*O*-Methylnorbelladine						
296,1425	8,81	C18H17NO3	Mecambrine						
			TOTAL	**38247**	**35437**	**12314**	**74625**	**17864**	

Table 2.2 Amaryllidaceae alkaloids detected by UPLC-QTOF-MS and the relative abundance detected through the area under the curve.

cDNA library, Illumina sequencing, *de novo* assembly, and annotation. To have a better understanding of the AA biosynthetic pathway and to enable the discovery of genes involved in their formation, we performed RNA sequencing of *N. pseudonarcissus* 'King Alfred'. Bulb tissues were chosen for transcriptomic analysis primarily based on alkaloid accumulation profiles determined above. RNA was extracted and screened for sufficient quality and quantity prior to cDNA library creation and deep sequencing using Illumina. A total of 73,081,603 raw paired reads were obtained, which were trimmed to give 66,054,792 surviving paired reads that corresponds to 90.4% (Table 2.3). For non-model plants such as Amaryllidaceae, whose genomic information is lacking, a *de novo* assembly is necessary. Detailed overview of the sequencing output, trimming, normalization, *de novo* assembly, and transcript length distribution graphs is provided in Table C.1. A total of 10,523,399 pair reads were obtained after the normalization step corresponding to 15.93%. Normalized reads were used to assemble the transcriptome generating 195,347 transcripts (or unigenes).

***N. pseudonarcissus* 'King Alfred'**	
Total number of raw paired reads[a]	73 081 603
Total number of surviving paired reads trimmed[b]	66 054 792
Total number of surviving paired reads normalized[c]	10 523 999
Number of transcripts[d]	195 347
Number of components[e]	98 332
Number of annotated transcripts[f]	11 708

Table 2.3 Summary of the transcriptome database generated from one lane of Illumina sequencing of *N. pseudonarcissus* 'King Alfred' bulbs. a. Number of Paired Reads obtained from the sequencer. b. Number of Remaining Paired Reads after the trimming step. c. Number of Remaining Paired Reads after the normalization step. d. Number of transcripts:Trinity has created a list of transcripts representing the transcriptome isoforms. e. Te transcripts are grouped in components loosely representing genes. Transcript names are prefixed by the component/gene name e.g. transcripts c115_g5_i1 and c115_g5_i2 are derived from the same isolated de Bruijn graph and therefore share the same component/gene number c115_g5. f. Each transcript has been aligned against the uniprot_sprot.trinotate_v2.0.pep protein database using blastx program from the NCBI BLAST family.

For identification, the assembled transcripts were aligned against the uniprot_sprot.trinotate_v2.0.pep protein database. BLAST annotation yielded 11,708 transcripts of an average length of 1,463 bp (Table C.1) Thus, a large number of transcripts showed no similarity to known genes and appears to represent transcripts of uncharacterized genes or sequences specific to *N. pseudonarcissus* 'King Alfred'. Altogether, we concluded that a good quality transcriptome was developed with a high number of surviving pair reads and assembled transcripts.

The 125 most abundant transcripts present in the transcriptome of *N. pseudonarcissus* 'King Alfred' bulbs were identified and compiled in Table C.2 and represent approximately 16 % of the transcriptome. First, it is interesting to note that bulb tissues contain viral RNA and the top four expressed transcripts belong to the *Narcissus* mosaic virus although no symptoms were detected. The top expressed plant genes belong to a group of plant survival and defense genes such as metallothionein-like protein, catalase, and peroxidases that help plants to defend against pathogens such as viruses. The second most highly represented category was primary metabolism including sucrose synthase, fructosyltransferase, glyceraldehyde 3-phosphate dehydrogenase, enolase, starch synthase, mannose-1-phosphate guanylyl transferase, UDP glucoronate 4-epimerase, *S*-adenosylhomocysteinase, and ATP synthase. Other highly represented categories were structural and transport proteins such as tubulin, aquaporin, transmembrane protein, BAC transporter, GRP2A, cathepsin, ADP ribosylation, and ADP/ATP carrier protein. Altogether, most abundant transcripts represent important sets of genes involved in cellular and metabolic process such as plant defense, primary metabolism, and structural and transport genes. Three transcripts encoding alkaloid biosynthetic genes, *S-norcoclaurine synthase1-like* (*NCS*; rank 62), *vittatine 11-hydroxylase* (rank 98) and *N4OMT3* (rank 125), were detected among the most abundantly expressed transcripts (Table C.2). NCS catalyzes the first committed step in benzylisoquinoline alkaloid biosynthesis which involves a stereoselective Pictet-Spengler condensation of dopamine and 4-hydroxyphenylacetaldehyde to form (*S*)-norcoclaurine (Samanani, Liscombe et al. 2004, Minami, Dubouzet et al. 2007, Lee and Facchini 2010).

Local BLASTx analyses were performed to identify specific gene transcripts encoding enzymes putatively involved in alkaloid biosynthesis (Fig. 2.3A). From the precursor pathway leading to norbelladine (Fig. 2.1), several transcript isoforms of orthologous genes were identified. For example, *PAL* genes have been cloned and characterized from various plant species including the Amaryllidaceae *Lycoris radiata* where two isoforms (*LrPAL1* and *LrPAL2*) were identified (Jiang, Xia et al. 2011, Jiang, Xia et al. 2013). BLASTx searches for *PAL* genes in the transcriptome led to the identification of 19 transcripts (5 full-length) from 6 different components. A component is a family of similar transcripts. For example, component 'contig1' may include several similar transcript sequences (contig1A, contig1B, contig1C, etc.) most likely RNA variants or isoforms of the same family. Of those, only two transcripts from 2 different components were >98% similar to *PAL* sequence. Thus, we identified two full-length *PAL* isoforms identified as *NpPAL1* and *NpPAL2*. *NpPAL1* was most similar to *LrPAL1* and *NpPAL2* isoforms annotated closer to *PAL* from *Ornithogalum longebracteatum* (sea onion) with an E-value of zero.

A)

Name	Reads (FPKM)	Length (bp)	Top Annotation	Species	e value	Accession number
TYDC1	1881	1641	tyrosine decarboxylase	*Narcissus aff. pseudonarcissus MK-2014*	0	KT378599.1
TYDC2	1111	2004	tyrosine decarboxylase 1-like	*Phoenix dactylifera*	0	XM_008812594.2
PAL1	1911	2277	phenylalanine ammonia-lyase (LrPAL1)	*Lycoris radiata*	0	FJ603650.2
PAL2	1790	2395	phenylalanine ammonia lyase 2	*Ornithogalum longebracteatum*	0	KF741222.1
C4H	3309	1800	cinnamic acid 4-hydroxylase	*Ornithogalum longebracteatum*	0	KF741224.1
C3H	2018	1830	p-coumarate 3-hydroxylase	*Narcissus tazetta*	0	KC455938.1
4CL1	1645	1955	4-coumarate:CoA ligase	*Dendrobium catenatum*	7E-143	KT894020.1
4CL2	1173	2010	4-coumarate:CoA ligase-like protein	*Ornithogalum longebracteatum*	0	KM393181.1
HCT	2572	1529	hydroxycinnamoyltransferase 1	*Elaeis guineensis*	3E-140	XM_019845714.1
N4OMT	21722	1062	norbelladine 4'-O-methyltransferase (N4OMT3)	*Narcissus aff. pseudonarcissus MK-2014*	0	KJ584563.1
CYP96T1	16956	1866	CYP96T1	*Narcissus aff. pseudonarcissus MK-2014*	0	KT693311.1
CYP96T2	16956	1839	CYP96T2	*Narcissus aff. pseudonarcissus MK-2014*	0	KT693312.1
NorRed	6876	1049	noroxomaritidine/norcraugsodine reductase	*Narcissus aff. pseudonarcissus MK-2014*	0	KU295569.1
Histone	9658	656	histone H3	*Lycoris longituba*	0	HQ403596.1

B)

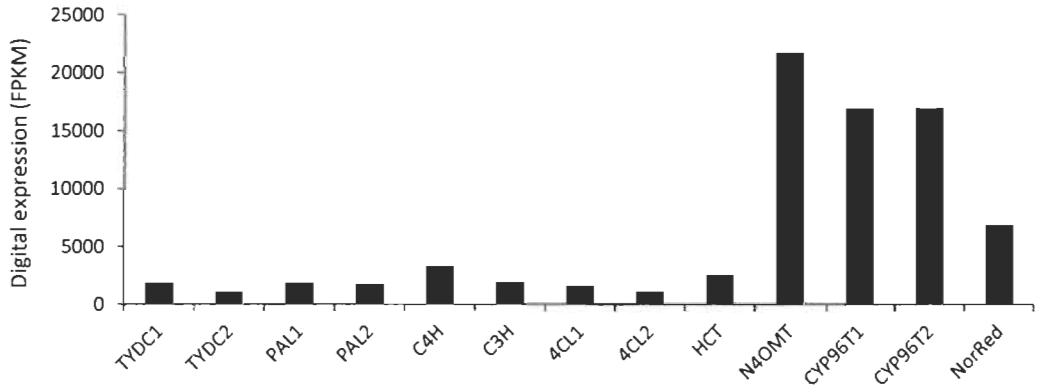


Figure 2.3 Summary of the biosynthetic genes transcripts corresponding to enzymes involved in Amaryllidaceae alkaloids metabolism in the bulbs transcriptome of *N. pseudonarcissus* King Alfred. (A) Table listing AA biosynthetic gene transcripts identified with top BLASTx annotation, corresponding e-value and GenBank accession number. All genes reported in this table correspond to full-length transcript. (B) Digital expression of AA biosynthetic genes expressed in fragments per kilobase million (FPKM).

(Fig. 2.3A). Similarly, 14 gene transcripts (2 components) were found for *TYDC* but only two full-length isoforms were identified. For gene transcripts specific to AA biosynthesis, we were able to identify all three genes with one isoform of *N4OMT*, two isoforms of *CYP96Ts* (*CYP96T1* and *CYP96T2*) and one isoform of *NorRed* (Fig 2.3A). Thus, we were able to find full-length transcripts and several isoforms with E-values corresponding to genes encoding enzymes involved in AA biosynthesis.

In the central AA biosynthetic pathway, from norbelladine to oxomaritidine (Fig. 2.1), sequence reads corresponding to *N4OMT3*, cytochrome P450s (*CYP96T1* and *CYP96T2*), and *NorRed* were more abundant than those operating in precursor pathways such as *PAL*, *TYDC,* and others. This was observed using comparative fragments per kilobase million (FPKM) digital expression (Fig. 2.3B). FPKM normalizes for sequencing depth and gene length. From the precursor pathway leading to norbelladine (Fig. 2.1), most transcripts were expressed at similar levels suggesting coordinated regulation for the formation of AA precursors.

Gene expression validation with RT-qPCR. To validate the gene expression data from the bulb transcriptome of *N. pseudonarcissus* 'King Alfred' and to assess transcript expression profiles across tissues, we performed reverse transcription-quantitative real-time PCR (RT-qPCR) analysis. All of the proposed AA biosynthetic gene transcripts were analyzed for expression profiles in different plant tissues (bulbs, roots, stems, leaves, and flowers) (Fig. 2.4). Only *N4OMT*, *TYDC,* and *PAL2* transcripts showed overexpression in bulbs compared to other tissues. Most phenylpropanoid precursor pathway genes (*PAL1*, *C4H*, *C3H*, *4CL,* and *HCT*), displayed a similar expression pattern; relatively high in stems followed by bulbs and low in other tissues (Fig. 2.4). Interestingly, *PAL1* and *PAL2* showed similar levels of expression in stems but *PAL2* expression was higher in bulbs than *PAL1*. From the central pathway, *N4OMT* was relatively the most highly expressed gene in bulb tissues (Fig. 2.4). AA specific genes *CYP96T* and *NorRed* were highly expressed in stems compared to other tissues, which did not correlate with *N4OMT* expression (Fig. 2.4). For most of the genes studied, leaves and flowers displayed the lowest expression compared to all other tissues. Taken together, the RT-qPCR results reflect relative differential gene expression profiles in different tissues.

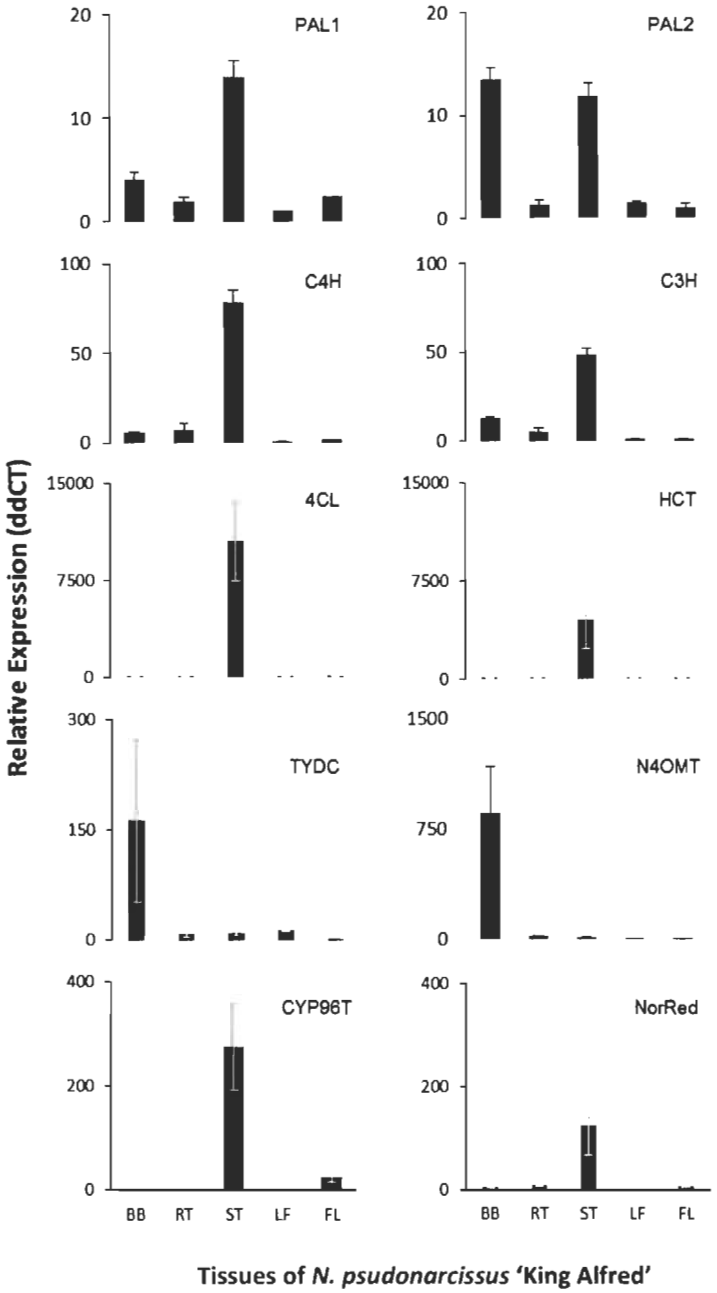


Figure 2.4 **Quantitative real-time PCR results of ten AAs biosynthesis pathway genes in different tissues of *N. pseudonarcissus* 'King Alfred'.** Graphs are plotted using normalized ddCT values scaled to lowest. Histone was used for internal reference. Expression fold change and error bars were calculated using the comparative $2^{-\Delta\Delta Ct}$ method from three independent experiments. Bars represent the mean standard deviation of three independent replicates. Abbreviations: Tissues of *N. pseudonarcissus* 'King Alfred' are BB, bulb; RT, root; ST, stem; LF, leaf; FL, flower.

Integration of transcriptome data. A broad survey of cellular metabolism involved in the conversion of sucrose to AAs resulted in the identification of transcripts and metabolites corresponding to a substantial number of metabolic steps (Fig. 2.5). With the exception of ribulose 5-phosphate epimerase (Ru5P epimerase), all enzymes required for the formation of AA precursors, tyramine and 3,4-DHBA were represented in the transcriptome database, whereas 14 AAs were found in the UPLC-QTOF-MS generated metabolite database. Many of the transcripts encoding primary metabolic enzymes were also among the most abundant transcripts (Table C.2). Notably, three independent gene transcripts in the top 125 encoded enzymes involved in the metabolism of *S*-adenosyl methionine (SAM). SAM is a common cosubstrate providing the methyl donor for the various reactions of *O*- and *N*-methylation in alkaloid biosynthesis. For example, the enzymes involved in the *N*-methylation of epinorlycoramine to lycoramine or the *O*-methylation of 11-hydroxyvittatine to crinamine could use SAM as the methyl donor (Fig. 2.1). Transcripts encoding enzymes proposed to be involved in AA precursor (tyramine and 3,4-DHBA) biosynthesis were identified along with the 4-hydroxybenzaldehyde metabolite. In addition, the three known AA specific gene transcripts *N4OMT*, *CYP96T* and *NorRed* were all identified in the transcriptome. However, AA metabolites from the oxomaritinamine branch pathway were not present in the QTOF-MS dataset.

The remaining AA biosynthetic enzymes, for which cognate cDNAs have not been isolated, catalyze the missing reactions. Those include the 1) chain-shortening steps involved in the formation of 3,4-DHBA, 2) the condensation of precursors to form norbelladine, and 3) the conversions to multiple AA intermediates (Figs 2.1 and 2.5). Some of these enzymes likely belong to known protein families including the CYPs, *O*- and *N*-methyltransferases, reductases, norcoclaurine synthase, dioxygenases, and oxidoreductases. Candidate transcripts with substantial identity to these protein families and other enzyme categories potentially involved in AA metabolism were found in the Illumina-generated transcript database. Thus, through comparative transcriptomics and

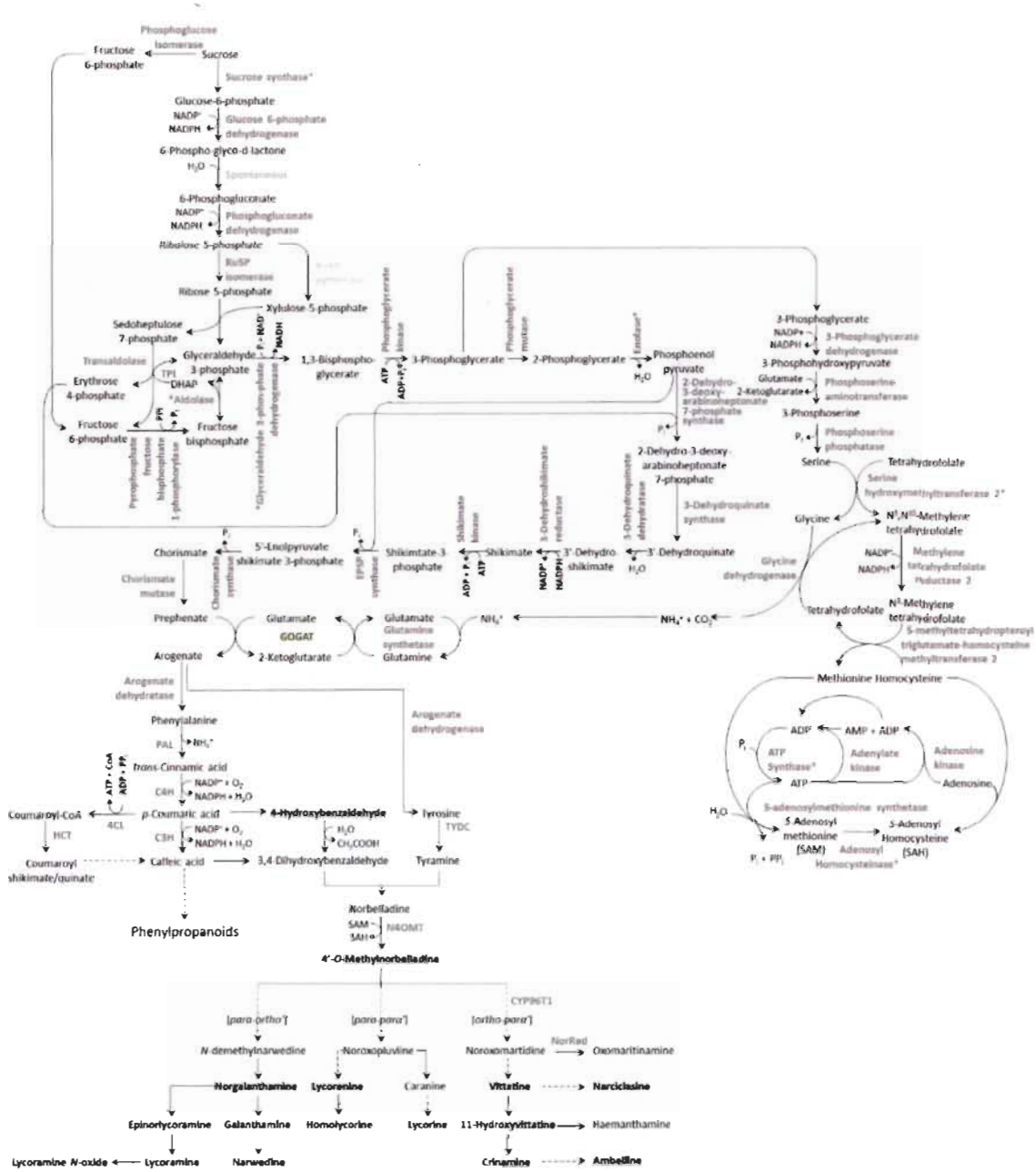


Figure 2.5 **Metabolic networks from sucrose to Amaryllidaceae alkaloids. Gene transcripts corresponding to enzymes shown in dark grey next to arrow were identified in the transcriptome database.** [Only Ru5P epimerase written in pale grey was not]. The (*) next to the enzyme name indicates gene transcripts present among the top 125 most abundant genes. Metabolites identified by UPLC-QTOF-MS are highlighted in bold black.

metabolomics approach, we were able to detect several AAs and biosynthetic genes pointing towards high quality and thorough transcriptome that could be very useful for further research on missing genes of AAs pathway.

2.5 Discussion

Integration of UPLC-QTOF-MS metabolite profiling and Illumina RNA sequencing was used to survey the metabolome and the transcriptome of *N. pseudonarcissus* 'King Alfred' tissues. The depth of each database provides new insights into AA metabolism and establishes valuable resources for the discovery of new genes involved in alkaloid metabolism.

Different alkaloid profiles were observed within the different tissues of *N. pseudonarcissus* 'King Alfred' with higher concentration of AA in leaves but higher diversity in bulbs (Fig. 2.2, Table 2.1). These results relate to previously published data for on *N. pseudonarcissus* 'Carlton' and *N. confusus* where highest amount of AAs occurred in bulbs (López, Bastida et al. 2003, Lubbe, Pomahacova et al. 2010, Berkov, Bastida et al. 2011, Lubbe, Verpoorte et al. 2012, Lubbe, Gude et al. 2013). Our QTOF-MS data showed predominantly *ortho-para'* (*i.e.* lycorine-type) AAs in leaves whereas *para-ortho'* (*i.e.* galanthamine-type) and *papa-para'* (*i.e.* crinamine-type) were abundant in all tissues. These results correlate with GC-MS analysis of six ornamental species of *Narcissus* (Torras-Claveria, Berkov et al. 2014). This suggests that different biosynthetic routes might be active within tissues. Transport of alkaloid precursors, intermediates or end-products might also be possible between tissues.

RNA-Seq is a useful tool to generate resources for non-model plants with restricted genomic information. Transcriptome assembly yielded 195,347 transcripts with a mean length of 761 base pairs. Similar numbers, 106,450 transcripts of mean length of 551 bp, were found in the transcriptome of *Narcissus spp.* (Kilgore, Augustin et al. 2014), supporting the quality and the depth of the transcriptome. The four most abundant transcripts from the *N. pseudonarcissus* 'King Alfred' transcriptome belong to viruses

(Table C.2), suggesting a contamination of plant material. In horticulture, ornamental plants such as *Narcissus* species are usually propagated vegetatively which makes them more susceptible to an accumulation of viruses (Wylie, Li et al. 2014). Virus infections are fairly common in Amaryllidaceae and up to 25 viruses have been reported (Brunt, Crabtree et al. 1996, Tijssen, Agbandje-McKenna et al. 2011). Mosaic viruses are the most common and several methods to detect them have been developed (Jin, Shen et al. 2017).

The 'plants' most abundant transcripts encoded proteins associated with metabolism, defense, transport and cellular structure similar to other plants (Delano-Frier, Aviles-Arnaut et al. 2011). For example, *c*omparison of the most abundant transcripts of Amaryllidaceae bulbs with *Papaver sominferum* revealed a similar pattern of expression, demonstrating that the genes involved in the common function of plant growth, development and defense are expressed at relatively consistent levels (Desgagne-Penix, Farrow et al. 2012). Also, transcripts encoding biosynthetic enzymes involved in the regeneration of (*S*)-adenosyl-methionine (Fig. 2.5; Table C.2) were among the most abundant in the database, which is in agreement with abundant genes found in the transcriptome of other alkaloid-producing plants (Desgagné-Penix, Khan et al. 2010, Desgagne-Penix, Farrow et al. 2012). A *NCS-like* sequence was found among the top most expressed genes, suggesting its plausible role in AAs biosynthesis at the norbelladine formation step. Similar results were observed in *Papaver sominferum* cell cultures (Desgagné-Penix, Khan et al. 2010). Moreover, at rank 115 we detected a putative cysteine proteinase similar to the C3 chain-shortening enzyme from *V. planifolia* (Podstolski, Havkin-Frenkel et al. 2002, Gallage, Hansen et al. 2014, Yang, Barros-Rios et al. 2017) capable of generating precursors 4-HBS or 3,4-DHBA. Interestingly, the presence of *vittatine 11-hydroxylase* among the top expressed genes correlates with elevated levels of crinamine (Table C.2; Table 2.3).

FPKM is directly proportional to the abundance of specific transcript in the transcriptome, thus quantification of the data provided an accurate measure of relative expression through sequencing depth and gene length. From the precursor pathway leading to norbelladine (Fig. 2.1), most gene transcripts were expressed at similar levels

suggesting coordinated regulation of the formation of AA precursors in bulbs (Fig. 2.3). Among the AA specific transcripts, *N4OMT3* and cytochrome P450s (*CYP96T1* and *CYP96T2*) were the most highly expressed ones, suggesting that they play crucial roles in bulb metabolic processes and that the AA central biosynthetic pathway expression is important in this tissue. Alkaloids are well known for their metabolic effects in animals (*e.g.* caffeine, nicotine, morphine and cocaine), and have probably evolved as defense compounds against herbivores and pathogens (Fürstenberg-Hägg, Zagrobelny et al. 2013). It was reported that the high alkaloid content of bulbs and leaves of Amaryllidaceae, primarily serves to protect the plant's carbohydrate resources from herbivores and microbial organisms (Ruiz, Ward et al. 2002, Gomez, Azorin et al. 2003). *PAL1* and *PAL2* expression was similar in bulbs, which correlates with previously reported *LrPAL1* and *LrPAL2* expression (Jiang, Xia et al. 2013). The presence of all proposed AA biosynthetic genes and the high levels of AA-specific genes, from the central pathway, correlates with the presence and high levels of AA in bulbs.

Comparison between digital expression (FPKM) and qPCR analysis from bulbs showed similar expression patterns for most genes (Figs 2.3B and 2.4). This was the case for *N4OMT,* which was highly expressed in both cases, and present among the most abundant transcripts (Table C.2). *N4OMT* expression was followed by *CYP96T1* and *NR*. Also RT-qPCR analysis confirmed the expression of all AA biosynthetic genes found in the transcriptome and revealed several novel gene transcripts in the Illumina-generated transcriptome. *PAL1* and *PAL2* showed similar levels of expression in stems but *PAL2* expression was higher in bulbs than *PAL1* suggesting different regulations/roles for each isoform/enzyme variant.

Altogether, qPCR results reflect relative differential gene expression profiles in different tissues. Since phenylpropanoids are widely distributed throughout plants(Kurkin 2003), it is expected that the expression of genes encoding phenylpropanoid enzymes (*e.g. PAL, C4H, C3H, 4CL,* and *HCT*) do not necessarily correlate with alkaloid content. The higher expression of *CYP96T* and *NorRed* in tissues containing the less amount of AAs (*e.g.* stems) is puzzling. However, AAs derived from these reactions were not detected in

any of the *N. pseudonarcissus* 'King Alfred' tissues. Only *TYDC* and *N4OMT*'s high expression levels, correlated with high alkaloid content in bulbs, support a relation between high alkaloid content and AA biosynthetic gene expression. It is also possible that transport of enzymes and AAs occurs within plants. Such transport has previously been observed and described for benzylisoquinoline alkaloids (Bird, Franceschi et al. 2003, Ziegler and Facchini 2008, Onoyovwe, Hagel et al. 2013).

In conclusion, we generated the first comprehensive metabolome and transcriptome databases from *N. pseudonarcissus* 'King Alfred' using UPLC-QTOF-MS and next-generation Illumina sequencing. We identified diverse AA profiles in different tissues and confirmed the expression of proposed AA biosynthetic genes. Since the integration of such databases, metabolome and transcriptome, allows for the discovery of novel genes, this study not only increases the knowledge of AA metabolism but also provides tools for metabolic engineering. Indeed, a deeper understanding of the molecular mechanism involved in AA biosynthesis is crucial to take advantage of new biotechnologies for the production of a class of health-relevant alkaloids.

2.6 Materials and methods

Plant material and chemicals. Ornamental bulbs of *N. pseudonarcissus* cultivar 'King Alfred' were purchased from Fraser's Thimble farms (BC, Canada). Bulbs were planted in September in well-drained soil to grow until the flowering stage in the spring. Different plant tissues such as bulbs, roots, stems, leaves and flowers were harvested in May for metabolic and transcriptomic analyses. HPLC grade acetonitrile and methanol were purchased from Fisher Scientific (https://www.fishersci.com). Standards narciclasine and galantamine were purchased from Tocris Bioscience (Bristol, U.K.) whereas lycorine and papaverine was purchased from Sigma-Aldrich (ON, Canada).

Alkaloid extraction and HPLC analysis. *N. pseudonarcissus* 'King Alfred' plant tissues (2 g/each in triplicate; bulbs, roots, stems, leaves and flowers) were crushed in

liquid nitrogen and extracted with methanol (10 ml) for 24 hrs at RT on a shaker (200 rpm). Extracts were centrifuged at 12,000 rpm for 5 minutes to collect supernatant without debris and left for complete evaporation. The raw alkaloids obtained were further subjected to a modified acid-base extraction method (specific for alkaloids), where they were resuspended in methanol 100% pH-8 (adjusted with NH3) and H_2SO_4 (2% v/v). Organic impurities were removed by washing twice with chloroform. Alkalization was done using NH3. The purified alkaloid extracts obtained were dried under N2 gas and finally solubilized in 300 µL methanol. Qualitative analysis of alkaloids was performed on TLC silica gel 60 F_{254} aluminum sheets 20 x 20 cm, (Merck, Darmstadt, Germany). Ten µl of 100 ppm for each standard (GAL, LYC, NAR) and 15 µl of samples were loaded in triplicates on TLC and visualized under 280 nm and 365 nm. For qualitative and quantitative analyses of alkaloids, each sample extract was diluted in 1/10 ratio using 1% ammonium acetate buffer. 15 ul of each sample was injected and analyzed on Shimadzu Prominence-i LC-2030C with diode array detector (PDA). HPLC oven temperature was set at 40°C. Chromatography assay was performed on Kinetex C18 column (150 x 4.6 mm, 5 µm particle size; Phenomenex). Elution was carried out in gradient mode with 1% ammonium acetate buffer (solvent A) and 100 % acetonitrile (solvent B). Initially 90:10 gradient ratio of solvent B and solvent A was maintained for 10 min. Eventually, gradient was shifted to 69:31 over 5 min, 10:90 over 2 min, and then 90:10 over 3 min. Eventually, after 18 min ammonium acetate was reduced to 10% and acetonitrile was increased to 90%, continuing till 30 min. HPLC Chromatograms of all standards and plant samples were extracted at 280 nm.

UPLC-QTOF-MS analysis. The UPLC analysis was performed using a Waters Acquity Ultra-Performance LC system (Waters), equipped with a binary pump system (Waters). An Acquity Ethylene Bridged Hybrid (BEH) C18 column (100 mm_2.1 mm id, 1.7 mm particle size) from Waters was used. The molecules were separated with a mobile phase that consisted of 0.2% acetic acid (solvent A) and 100 % acetonitrile (solvent B). The flow-rate was 0.2 mL/min and the gradient elution was initial, 2% B ; 0-1 min, 2–100% B; 1-30 min, isocratic 100% B; 30-33 min, 100-2% B ; 33-33.5 min, isocratic 2% B; 33-40 min. The MS analyses were carried out on a QTOF Micro mass spectrometer

(Waters) equipped with a Z-spray electrospray interface. The analysis was performed in both positive and negative mode and the data were acquired through a mass scan from 100 to 1250 *m/z* without collision. The ionization source parameters were source temperature, 120°C; cone gas flow rate, 50 L/h and desolvation gas flow rate, 350 L/h; desolvation temperature, 200°C. The cone and capillary voltages were respectively set at 30 V and 1150 V for the negative mode and 70 V and 1800 V for the positive mode. Nitrogen (99% purity) was used as nebulizing gas. Data acquisition was carried out with the MassLynx 4.1 software. Masses extraction, deconvolution, isotopes and library search was performed using MZMine 2 according to Pluskal et al. (2010) (Pluskal, Castillo et al. 2010).

RNA extraction, next-generation Illumina sequencing, and *de novo* assembly. Five grams of triplicate bulbs, stems, roots, leaves or flowers of *N. pseudonarcissus* 'King Alfred' were crushed using a pestle and a mortar with liquid nitrogen and transferred to pre-chilled 50 ml tubes to proceed with CTAB (cetrimonium bromide) method for total RNA extraction(Desgagne-Penix, Farrow et al. 2012). After RNA extraction, bulb RNA was selected for Illumina sequencing, and stem, root, leaf and flower RNA were used for RT-qPCR analysis. For transcriptome analysis, integrity of the bulb RNA was checked on a nanodrop and bioanalyzer. Nanodrop quantification yielded 805.68 ng/µl of total RNA with the ratio 260 nm/230 nm of 2.12 and 260 nm/280 nm of 1.91. Bioanalysis of RNA gave a RNA integrity number (RIN) of 8.3 with 28S/18S of 1.546656 confirming the RNA quality as pure and intact *i.e.* not degraded.

The mRNA were converted into cDNA library and sequenced through Illumina HiSeq 2000, PE 100 paired ends, at McGill University and Genome Quebec Innovation Centre (Montreal, Canada). Raw paired reads were trimmed from 3'-end and Illumina sequencing adapters were removed, maintaining 50 bps of minimum read length to obtain surviving paired reads. Trimming and clipping were done using Trimmomatic (http://www.usadellab.org/cms/index.php?page=trimmomatic) for quality filtering to have clean reads (Bolger, Lohse et al. 2014). *De novo* assembly of clean reads was done using trinity assembler (https://github.com/trinityrnaseq/trinityrnaseq/wiki). Once

surviving pair data were generated, normalization was performed to eliminate redundant reads in datasets without affecting its *Kmer* content. Final obtained unigenes were functionally annotated using trinotate (http://trinotate.github.io/), using the Trinity normalization utility inspired by the diginorm algorithm (Brown, Howe et al. 2012). Normalized reads were used to assemble the transcriptome using the Trinity assembler (Grabherr, Haas et al. 2011). To quantify the gene transcript abundance, the raw RNA-Seq reads were mapped to assembled transcripts applying Bowtie (Langmead, Trapnell et al. 2009) using default parameters. The gene transcript abundance was calculated as 'fragments per million mapped reads per kilobase' (FPKM) using the RSEM package(Li and Dewey 2011). The genes with extremely low FPKM values *i.e.* with a maximum FPKM of less than 1 across two samples, were filtered out before subsequent analysis.

RT-qPCR analysis. Two micrograms of RNA from different tissues (bulb, root, stem, leaf, and flower) of triplicate plants of *N. pseudonarcissus* 'King Alfred' were reverse transcribed to form cDNA using oligo dT primers and the Omniscript Reverse Transcription Kit (Qiagen) according to given manufacturer's protocol. Gene specific primers were designed using Integrated DNA Technology (www.idtdna.com) and Tm Calculator, New England Biolabs (tmcalculator.neb.com), to select suitable annealing temperatures. The sequences for all of the primers used in this study are listed in table C.3. SensiFAST SYBER Lo-ROX mix (Bioline) was used to prepare a 20 µl reaction containing 1x SensiFAST SYBER Lo-ROX mix, 200 µM of each forward and reverse primer and 3 µl of cDNA sample. Each experiment was performed in triplicate and histone was used as an internal reference gene. Real-time quantitative PCR was performed on CFX Connect Real-Time PCR System (BioRad). PCR conditions for amplification were 95°C for 3 min, 95°C for 10 sec and annealing temperature (varies with gene primers, see Table C.3) for 30 sec for 40 cycles. This was followed by a dissociation step (as provided by software) - 95°C for 10 sec, 65°C for 5 sec and 95°C for 5 sec. The comparative $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen 2001) was used for relative quantification of the gene expression levels.

Accession numbers. The sequences described in this paper have been deposited in the National Center for Biotechnology Information Sequence Read Archive (https://www.ncbi.nlm.nih.gov/sra/) under the accession number SRR5788585. Gene transcript sequences were deposited in Genbank with the following accession numbers for nucleotide sequences: *tyrosine decarboxylase 1* (MF405171), *tyrosine decarboxylase 2* (MF405172), *phenylalanine ammonia lyse 1* (MF405173), *phenylalanine ammonia lyse 2* (MF405174), *cinnamate 4-hydroxylase* (MF416091), *coumarate 3-hydroxylase* (MF416092), *4-coumarate-CoA ligase 1* (MF416093), *4-coumarate-CoA ligase 2* (MF416094), *hydroxycinnamoyltransferase* (MF416095), *norbelladine 4'-O-methyltransferase* (MF416096), *noroxomaritidine synthase 1* (MF416097), *noroxomaritidine synthase 2* (MF416098), *noroxomaritidine/norcraugsodine reductase* (MF416099) and *histone* (MF405170).

2.7 Acknowledgments

We kindly thank Professor Hugo Germain and Tarun Hotchandani for the revision and helpful comments on a previous version of this manuscript. This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) award number RGPIN 05294-2014 (Discovery) to I.D-P. This work was also supported by the NSERC award number EQPEQ 472990-2015 (Research tools and instruments) for the acquisition of the HPLC- PDA and the qPCR.

2.8 References

Barik, J., et al. (2005). "Cellular responses to nicotinic receptor activation are decreased after prolonged exposure to galantamine in human neuroblastoma cells." Br J Pharmacol **145**(8): 1084-1092.

Barton, D. H. R., et al. (1963). "Phenol oxidation and biosynthesis. Part VI. The biogenesis of Amaryllidaceae alkaloids." J. Chem. Soc.: 4545-4558.

Bastida, J., et al. (2011). Chapter 3: Chemical and biological aspects of Amaryllidaceae alkaloids. E. D. Muñoz-Torrero, T. R. Network, F. P. O. 37/661 (2), T.-. 023 and I. Kerala, Recent Advances in Pharmaceutical Sciences.

Battersby, A. R., et al. (1961). "Biosynthesis in the Amaryllidaceae. Tyrosine and norbelladine as precursors of haemanthamine." J. Amer. Chem. Soc. **83**: 4098-4099.

Berkov, S., et al. (2011). "Development and validation of a GC-MS method for rapid determination of galanthamine in Leucojum aestivum and Narcissus ssp.: a metabolomic approach." Talanta **83**(5): 1455-1465.

Bird, D. A., et al. (2003). "A tale of three cell types: alkaloid biosynthesis is localized to sieve elements in opium poppy." Plant Cell **15**(11): 2626-2635.

Bolger, A. M., et al. (2014). "Trimmomatic: a flexible trimmer for Illumina sequence data." Bioinformatics **30**(15): 2114-2120.

Brown, C. T., et al. (2012). "A reference-free algorithm for computational normalization of shotgun sequencing data." arXiv preprint arXiv:1203.4802.

Brunt, A. A., et al. (1996). "Plant Viruses Online: Descriptions and Lists from the VIDE Database. Version: 16th January 1997." URL http://biology.anu.edu.au/Groups/MES/vide/.

De Felice, B., et al. (2013). "Comparative transcriptional analysis reveals differential gene expression between Sand Daffodil tissues." Genetica **141**(10-12): 443-452.

Delano-Frier, J. P., et al. (2011). "Transcriptomic analysis of grain amaranth (Amaranthus hypochondriacus) using 454 pyrosequencing: comparison with A. tuberculatus, expression profiling in stems and in response to biotic and abiotic stress." BMC Genomics **12**: 363.

Desgagné-Penix, I. and P. J. Facchini (2011). Benzylisoquinoline Alkaloid Biosynthesis. Plant Metabolism and Biotechnology, John Wiley & Sons, Ltd: 241-261.

Desgagne-Penix, I., et al. (2012). "Integration of deep transcript and targeted metabolite profiles for eight cultivars of opium poppy." Plant Molecular Biology **79**(3): 295-313.

Desgagné-Penix, I., et al. (2010). "Integration of deep transcriptome and proteome analyses reveals the components of alkaloid metabolism in opium poppy cell cultures." BMC Plant Biology **10**(1): 252.

Dewick, P. M. (2009). Medicinal natural products: a biosynthetic approach., John Wiley and Sons Ltd.

Diamond, A. and I. Desgagne-Penix (2016). "Metabolic engineering for the production of plant isoquinoline alkaloids." Plant Biotechnol J **14**(6): 1319-1328.

Dumont, P., et al. (2007). "The Amaryllidaceae isocarbostyril narciclasine induces apoptosis by activation of the death receptor and/or mitochondrial pathways in cancer cells but not in normal fibroblasts." Neoplasia **9**(9): 766-776.

Eichhorn, J., et al. (1998). "Biosynthesis of the Amaryllidaceae alkaloid galanthamine." Phytochemistry **49**(4): 1037-1047.

Facchini, P. J., et al. (2000). "Plant aromatic L-amino acid decarboxylases: evolution, biochemistry, regulation, and metabolic engineering applications." Phytochemistry **54**(2): 121-138.

Fürstenberg-Hägg, J., et al. (2013). "Plant Defense against Insect Herbivores." Int J Mol Sci **14**(5): 10242-10297.

Gahlan, P., et al. (2012). "De novo sequencing and characterization of Picrorhiza kurrooa transcriptome at two temperatures showed major transcriptome adjustments." BMC genomics **13**(1): 126.

Gallage, N. J., et al. (2014). "Vanillin formation from ferulic acid in Vanilla planifolia is catalysed by a single enzyme." Nat Commun **5**: 4037.

Ge, G., et al. (2011). "The first insight into the tissue specific taxus transcriptome via Illumina second generation sequencing." PloS one **6**(6): e21220.

Gomez, D., et al. (2003). "Seasonal and spatial variations of alkaloids in Merendera montana in relation to chemical defense and phenology." J Chem Ecol **29**(5): 1117-1126.

Grabherr, M. G., et al. (2011). "Full-length transcriptome assembly from RNA-Seq data without a reference genome." Nat Biotechnol **29**(7): 644-652.

Hagel, J. M. and P. J. Facchini (2005). "Elevated tyrosine decarboxylase and tyramine hydroxycinnamoyltransferase levels increase wound-induced tyramine-derived hydroxycinnamic acid amide accumulation in transgenic tobacco leaves." Planta **221**(6): 904-914.

Hagel, J. M., et al. (2015). "Metabolome analysis of 20 taxonomically related benzylisoquinoline alkaloid-producing plants." BMC Plant Biol **15**(1): 1-17.

Hagel, J. M., et al. (2015). "Transcriptome analysis of 20 taxonomically related benzylisoquinoline alkaloid-producing plants." BMC Plant Biol **15**.

Hao, D., et al. (2012). "De novo characterization of the root transcriptome of a traditional Chinese medicinal plant Polygonum cuspidatum." Science China Life Sciences **55**(5): 452-466.

He, M., et al. (2015). "Biological and pharmacological activities of amaryllidaceae alkaloids." RSC Advances **5**(21): 16562-16574.

Heinrich, M. and H. Lee Teoh (2004). "Galanthamine from snowdrop--the development of a modern drug against Alzheimer's disease from local Caucasian knowledge." J Ethnopharmacol **92**(2-3): 147-162.

Hotchandani, T. and I. Desgagne-Penix (2017). "Heterocyclic Amaryllidaceae Alkaloids: Biosynthesis and Pharmacological Applications." Curr Top Med Chem **17**(4): 418-427.

Hyun, T. K., et al. (2012). "De novo transcriptome sequencing of Momordica cochinchinensis to identify genes involved in the carotenoid biosynthesis." Plant Molecular Biology **79**(4): 413-427.

Jiang, Y., et al. (2013). "Molecular and analysis of a phenylalanine ammonia-lyase gene (LrPAL2) from Lycoris radiata." Mol Biol Rep **40**(3): 2293-2300.

Jiang, Y., et al. (2011). "Molecular cloning and characterization of a phenylalanine ammonia-lyase gene (LrPAL) from Lycoris radiata." Mol Biol Rep **38**(3): 1935-1940.

Jin, J., et al. (2017). "Narcissus yellow stripe virus and Narcissus mosaic virus detection in Narcissus via multiplex TaqMan-based reverse transcription-PCR assay." J Appl Microbiol **122**(5): 1299-1309.

Jin, Z. and X.-H. Xu (2013). Amaryllidaceae Alkaloids. Natural Products. K. G. Ramawat and J.-M. Mérillon, Springer Berlin Heidelberg: 479-522.

Kilgore, M., et al. (2016). "CYP96T1 of Narcissus sp. aff. pseudonarcissus Catalyzes Formation of the Para-Para' C-C Phenol Couple in the Amaryllidaceae Alkaloids." Frontiers in plant science **7**.

Kilgore, M., et al. (2016). "Identification of a Noroxomaritidine Reductase with Amaryllidaceae Alkaloid Biosynthesis Related Activities." Journal of Biological Chemistry.

Kilgore, M. B., et al. (2014). "Cloning and characterization of a norbelladine 4′-O-methyltransferase involved in the biosynthesis of the Alzheimer's drug galanthamine in Narcissus sp. aff. pseudonarcissus." PloS one **9**(7): e103223.

Kornienko, A. and A. Evidente (2008). "Chemistry, biology, and medicinal potential of narciclasine and its congeners." Chem Rev **108**(6): 1982-2014.

Kurkin, V. A. (2003). "Phenylpropanoids from Medicinal Plants: Distribution, Classification, Structural Analysis, and Biological Activity." Chemistry of Natural Compounds **39**(2): 123-153.

Langmead, B., et al. (2009). "Ultrafast and memory-efficient alignment of short DNA sequences to the human genome." Genome biology **10**(3): R25.

Lee, E. J. and P. Facchini (2010). "Norcoclaurine synthase is a member of the pathogenesis-related 10/Bet v1 protein family." Plant Cell **22**(10): 3489-3503.

Li, B. and C. N. Dewey (2011). "RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome." BMC bioinformatics **12**(1): 323.

Livak, K. J. and T. D. Schmittgen (2001). "Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method." Methods **25**(4): 402-408.

López, S., et al. (2002). "Acetylcholinesterase inhibitory activity of some Amaryllidaceae alkaloids and Narcissus extracts." Life Sciences **71**(21): 2521-2529.

López, S., et al. (2003). "Galanthamine pattern in Narcissus confusus plants." Planta medica **69**(12): 1166-1168.

Lubbe, A., et al. (2013). "Seasonal accumulation of major alkaloids in organs of pharmaceutical crop Narcissus Carlton." Phytochemistry **88**: 43-53.

Lubbe, A., et al. (2010). "Analysis of metabolic variation and galanthamine content in Narcissus bulbs by 1H NMR." Phytochem Anal **21**(1): 66-72.

Lubbe, A., et al. (2012). "Effects of fungicides on galanthamine and metabolite profiles in Narcissus bulbs." Plant Physiol Biochem **58**: 116-123.

Minami, H., et al. (2007). "Functional analysis of norcoclaurine synthase in Coptis japonica." J Biol Chem **282**(9): 6274-6282.

Onoyovwe, A., et al. (2013). "Morphine biosynthesis in opium poppy involves two cell types: sieve elements and laticifers." Plant Cell **25**(10): 4110-4122.

Pluskal, T., et al. (2010). "MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data." BMC Bioinformatics **11**(1): 395.

Podstolski, A., et al. (2002). "Unusual 4-hydroxybenzaldehyde synthase activity from tissue cultures of the vanilla orchid Vanilla planifolia." Phytochemistry **61**(6): 611-620.

Rischer, H., et al. (2006). "Gene-to-metabolite networks for terpenoid indole alkaloid biosynthesis in Catharanthus roseus cells." Proceedings of the National Academy of Sciences **103**(14): 5614-5619.

Ruiz, N., et al. (2002). "Responses of Pancratium sickenbergeri to simulated bulb herbivory: combining defence and tolerance strategies." Journal of Ecology **90**(3): 472-479.

Samanani, N., et al. (2004). "Molecular cloning and characterization of norcoclaurine synthase, an enzyme catalyzing the first committed step in benzylisoquinoline alkaloid biosynthesis." Plant J **40**(2): 302-313.

Sener, B., et al. (2003). "Antimalarial activity screening of some alkaloids and the plant extracts from Amaryllidaceae." Phytother Res **17**(10): 1220-1223.

Singh, A. and I. Desgagne-Penix (2014). "Biosynthesis of the Amaryllidaceae alkaloids." Plant Science Today **1**(2): 114-120.

Singh, A. and I. Desgagné-Penix (2015). Chapter 3 : Biosynthesis of Amaryllidaceae alkaloids: A biochemical outlook. In: Alkaloids: Biosynthesis, Biological Roles and Health benefits., Nova Science Publishers.

Szlavik, L., et al. (2004). "Alkaloids from Leucojum vernum and antiretroviral activity of Amaryllidaceae alkaloids." Planta Med **70**(9): 871-873.

Takos, A. M. and F. Rook (2013). "Towards a molecular understanding of the biosynthesis of amaryllidaceae alkaloids in support of their expanding medical use." Int J Mol Sci **14**(6): 11713-11741.

Thomsen, T. and H. Kewitz (1990). "Selective inhibition of human acetylcholinesterase by galanthamine in vitro and in vivo." Life Sci **46**(21): 1553-1558.

Tijssen, P., et al. (2011). "Virus taxonomy 9th Report of the International Committee on Taxonomy of Viruses. ." 405–425.

Torras-Claveria, L., et al. (2014). "Metabolomic analysis of bioactive Amaryllidaceae alkaloids of ornamental varieties of Narcissus by GC–MS combined with k-means cluster analysis." Industrial Crops and Products **56**: 211-222.

Wang, R., et al. (2013). "De novo sequence assembly and characterization of Lycoris aurea transcriptome using GS FLX titanium platform of 454 pyrosequencing." PLoS One **8**(4): e60449.

Wylie, S. J., et al. (2014). "Complete genome analysis of three isolates of narcissus late season yellows virus and two of narcissus yellow stripe virus: three species or one?" Arch Virol.

Xiao, M., et al. (2013). "Transcriptome analysis based on next-generation sequencing of non-model plants producing specialized metabolites of biotechnological interest." Journal of biotechnology **166**(3): 122-134.

Yamazaki, M., et al. (2013). "Coupling deep transcriptome analysis with untargeted metabolic profiling in Ophiorrhiza pumila to further the understanding of the biosynthesis of the anti-cancer alkaloid camptothecin and anthraquinones." Plant and cell physiology **54**(5): 686-696.

Yang, H., et al. (2017). "A re-evaluation of the final step of vanillin biosynthesis in the orchid Vanilla planifolia." Phytochemistry **139**: 33-46.

Ziegler, J. and P. J. Facchini (2008). "Alkaloid biosynthesis: metabolism and trafficking." Annu Rev Plant Biol **59**: 735-769.

Chapter III contains a study demonstrating that norbelladine synthse (*Np*NBS), a Bet v1/PR-10 protein, is responsible for catalysing the first committed step in Amaryllidaceae alkaloid biosynthesis pathway.

CHAPTER III

CLONING AND CHARACTERIZATION OF NORBELLADINE SYNTHASE CATALYZING THE FIRST COMMITTED REACTION IN AMARYLLIDACEAE ALKALOID BIOSYNTHESIS

APARNA SINGH[1], MARIE-ANGE MASSICOTTE[1], LAURENCE TOUSIGNANT[1], VINCENT OUELLET[1], GERVAIS BERUBE[1], ISABEL-DESGAGNE PENIX[1,2]

Submitted on 28th of February 2018 in *The Plant Cell*

3.1 Contributions

IDP and AS conceived and designed the experiments. AS performed the experiments. MAM and LT performed LC-MS/MS analysis. VO and GB chemically synthesized the norbelladine and norcraugsodine standards and validate them using NMR analysis. IDP and AS interpreted results, prepared the figures, wrote and reviewed the manuscript.

3.2 Abstract

Amaryllidaceae alkaloids (AAs) are a large group of plant-specialized metabolites displaying an array of biological and pharmacological properties. Previous investigations on AA biosynthesis have revealed that all AAs share a common precursor, norbelladine, presumably synthesized by an enzyme catalyzing a Mannich reaction involving the condensation of tyramine and 3,4-dihydroxybenzaldehyde. Similar reactions have been reported. Specifically, norcoclaurine synthase (NCS) which catalyzes the condensation of dopamine and 4-hydroxyphenylacetaldehyde as the first step in benzylisoquinoline alkaloid biosynthesis. With the availability of wild daffodil (*Narcissus pseudonarcissus*) database, a transcriptome-mining search was performed for *NCS* orthologs. A candidate

gene sequence was identified and named *norbelladine synthase* (*NBS*). *NpNBS* encodes for a small protein of 19 KDa with an anticipated pI of 5.5. Phylogenetic analysis showed that *Np*NBS belongs to a unique clade of PR10/Bet v1 proteins and shared 41% amino acid identity to *Papaver sominferum* NCS1. Expression of *NpNBS* cDNA in *Escherichia coli* produced a recombinant enzyme able to condense tyramine and 3,4-DHBA into norbelladine as determined by high-resolution tandem mass spectrometry. Here, we introduce a novel enzyme catalyzing the first committed step of AA biosynthesis and will facilitate the establishment of metabolic engineering and synthetic biology platforms for the production of AAs.

3.3 Introduction

The Amaryllidaceae alkaloids (AAs) are a group of naturally synthesized molecules with more than 500 renowned complex structures (Jin and Xu 2013). They are pharmacologically active compounds that are classified under three different groups of C-C phenol coupling namely *para-para'*, *ortho-para'* and *para-ortho'* (Singh and Desgagné-Penix 2014). An outsized variety of pharmacologically active AAs have been identified with the bioactive properties including the acetylcholine esterase inhibitor galanthamine, anti-tumor activity of lycorine and the cytotoxic haemanthamine (Li, Dai et al. 2012, He, Qu et al. 2015, Hotchandani and Desgagne-Penix 2017). AAs are obtained chiefly from the extracts of plants from *Galanthus*, *Leucojum* and *Narcissus* species, as their complicated structures does not enable cost-effective high-yield organic synthesis (Saliba, Ptak et al. 2015). Though AAs display a large range of pharmaceutical applications, solely galanthamine is accessible in markets as an Alzheimer's treatment drug because of its ability to stabilize behavioral symptoms in the course of six months treatment in comparison to chemically synthesized acetylcholine esterase inhibiting drugs, donepezil and rivastigmine (Prvulovic, Hampel et al. 2010).

Previous investigations on the biosynthesis of AAs *in planta* has revealed that all AAs are made from the common metabolic intermediate, norbelladine (Figure 3.1) (Barton and Cohen 1957, Battersby, Fales et al. 1961, Barton, Kirby et al. 1963, Grisebach

1973, Eichhorn, Takada et al. 1998, Tahchy, Boisbrun et al. 2010, El Tahchy, Ptak et al. 2011). For example, radiolabeled studies showed that deuterium labelled *4'-O-methylnorbelladine* was incorporated in all three different groups of AAs (Tahchy, Boisbrun et al. 2010). To date, only three AAs biosynthetic genes have been identified including the norbelladine *4'-O-methyltransferases (N4OMT)*, encoding N4OMT enzyme catalyzing norbelladine methylation at 4' position to form 4'-O-methylnorbelladine (Kilgore, Augustin et al. 2014), *CYP96T1,* encoding a cytochrome P450 enzyme which catalyzes the synthesis of (S,R)-noroxomaritidine (Kilgore, Augustin et al. 2016) and *noroxomaritidine reductase (NR)* encoding the enzyme forming oxomaritinamine (Kilgore, Holland et al. 2016). However, to this date, the enzyme catalyzing the first committed step leading to norbelladine synthesis has not been identified.

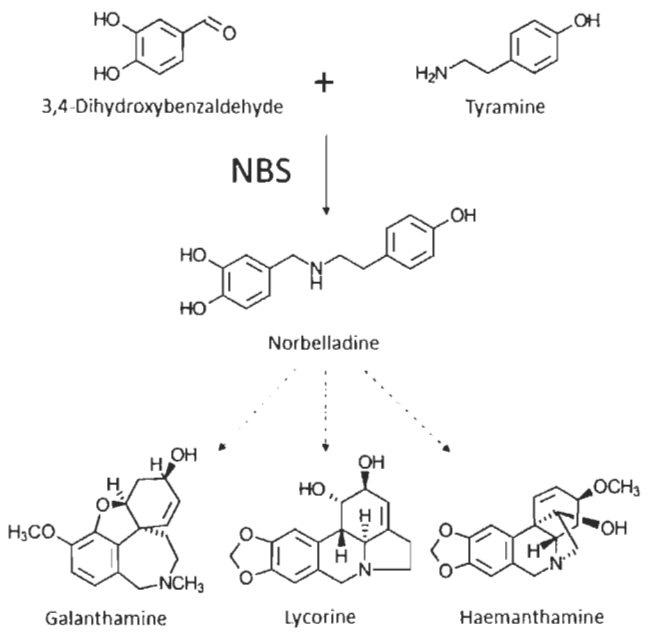


Figure 3.1 Norbelladine synthase (NBS) catalyzes the condensation of tyramine and 3,4-dihydroxybenzaldehyde (3,4-DHBA) to form norbelladine, the common precursor to all Amaryllidaceae alkaloids produced in plants including galanthamine, lycorine and haemanthamine.

It has been proposed that norbelladine is formed by the condensation of tyramine and 3,4-dihydroxybenzaldehyde (3,4-DHBA) using the Mannich reaction (Sobaro-Sanchez, 2015). The involvement of the Mannich reaction has been proposed in many

biosynthetic pathways, especially for alkaloids. For example, norcoclaurine synthase (NCS), catalyzing the first committed step in benzylisoquinoline alkaloids formation, condenses the two precursors molecules; dopamine and 4-hydroxyphenylacetaldehyde to form norcoclaurine (Samanani, Liscombe et al. 2004). Since the AA biosynthetic pathway shares analogy with the benzylisoquinoline alkaloid biosynthesis, we proposed that a *NCS* ortholog in Amaryllidaceae would be an interesting candidate potentially able to catalyze the formation of norbelladine.

Hence for this study, we have applied BLAST searches for candidate sequences ortholog to *NCS* in our previously generated de novo transcriptome from in wild daffodil (*Narcissus pseudonarcissus*) (Singh and Desgagné-Penix 2017). The NCS orthologous sequence identified named *norbelladine synthase* (NBS) was isolated, cloned in an expression vector and transformed into *E. coli*. Following heterologous expression and purification, NBS recombinant protein was used for enzyme assay and LC-MS/MS analysis of the product showed that NBS have ability to convert 3,4-DHBA and tyramine to norbelladine. Thus, we report on the first characterization of norbelladine synthase, the enzyme which catalyzes the first committed step in AAs biosynthesis.

3.4 Materials and Methods

Plant tissue and chemicals. Wild daffodil (*Narcissus pseudonarcissus* 'King Alfred') bulbs were purchased from Fraser's Thimble farms (BC, Canada) and were planted in September. Plants were harvested during peak flowering and were separated into bulbs, roots, stems, leaves, and flowers. Ampicillin, HPLC grade acetonitrile, agarose, methanol, 6x His-tag epitope tag antibody and monosodium phosphate were purchased from Fisher Scientific (Janssen Pharmaceuticalaan, Geel, Belgium). 3, 4-dihydroxybenzaldehyde (3,4-DHBA) were obtained from Acros organic (New Jersey, USA).Tyramine and chloramphenicol were bought from Sigma-Aldrich (MO, USA). Taq DNA polymerase was purchased from gene direx. Sodium chloride (NaCl), isopropyl β-D-1-thiogalactopyranoside (IPTG), and imidazole were bought from Fisher Bioreagents/scientific (New Jersey, USA). sensiFAST SYBER Lo-Rox kit for

quantitative Real-time PCR was obtained from Bioline (London, U.K). Gateway cloning kit was purchased from Invitrogen (CA, USA). Ni-NTA his-tag affinity columns were purchased from Qiagen (Germany), plasmid miniprep kit acquired from Geneaid (Taipei, Taiwan). Mini-protean TGX stain-free precast gels, 4x-laemmli buffer and clarity western ECL substrate were was obtained from Bio-rad (USA) and kanamycin was purchased from Bioshop (Burlington, ON, Canada).

Candidate gene identification. The sequence for Norbelladine Synthase was deposited to GenBank and the GenBank accession number for the nucleotide sequence is MG948545.

Phylogenetic tree and protein alignment. Sequences listed in table D.1 were aligned using CLUSTAL W in MEGA 6 software with default parameters. The evolutionary history was inferred using the Neighbor-Joining method using the branch lengths contained in the inferred tree (Saitou and Nei 1987). Divergence times for all branching points in the topology were calculated with the RelTime method (Tamura, Battistuzzi et al. 2012). Phylogenetic analysis was conducted on MEGA 6 (Tamura, Stecher et al. 2013). Protein sequence of *Np*NBS was aligned with norcoclaurine synthase sequences from *Coptis japonica* (*Cj*NCS2), *Papaver somniferum* (*Ps*NCS1), *Eschscholzia californica* (*Ec*NCS1) and *Thalictrum flavum* (*Tf*NCS1) using Clustal omega. *Np*NBS sequence was analyzed for signal peptide using Signal-BLAST (http://sigpep.services.came.sbg.ac.at/signalblast) and PSLpred (http://crdd.osdd.net/raghava/pslpred). Motif scan search for *Np*NBS, *Ps*NCS1, *Bt*PR10 and *Cj*NCS sequences was performed using myhits (http://myhits.isb-sib.ch/cgi-bin/motif_scan).

Quantitative Real-Time PCR. *N. pseudonarcissus* cDNA for bulbs, roots, stems, leaves, and flowers was generated from 1-2 μg of RNA using the Qiagen omniscript RT kit according to manufacturer's protocol (QIAGEN, Germany). The experiment was performed in triplicates. Total 20 μL reaction containing 1x SensiFAST SYBER Lo-ROX mix, 200 μM of each forward and reverse primers (table D.2) and cDNA sample was used

for qRT-PCR analysis. Histone was used as internal reference gene. Real-time quantitative PCR was performed on CFX Connect Real-Time PCR System from Bio-rad (USA). PCR conditions for amplification were 95°C for 3 mins, 95°C for 10 seconds, annealing temperature 52°C for 30 seconds for 40 cycles. This was followed by dissociation step (as provided by software) - 95°C for 10 seconds, 65°C for 5 seconds and 95°C for 5 seconds. Norbelladine synthase relative expression values were determined by comparative 2-ΔΔCt method and were scaled to lowest (Livak, 2001).

PCR and Cloning. Open reading frame (ORF) of full length *NpNBS* was amplified from *N. pseudonarcissus* bulbs cDNA with 200 μM dNTPs, 1.25 unit *taq* DNA polymerase in recommended amount of buffer in 50 μL reaction and 0.2 μM forward and reverse gateway primers (table D.2). PCR program parameters: 3 mins 94° 1 cycle, 30 s 94°, 45 s 52°, 1 min 72° for 30 cycles, 5 min 72°C 1 cycle. A Gateway cloning technology was used to clone *Np*NBS according to manufacturer's protocol. The gateway-adapted *att*P-flanked pDONR 221 vector with kanamycin resistance gene was used for BP recombinase reaction (catalyzed by BP clonase enzyme) to generate an *att*L-flanked entry clones with *att*B-flanked *Np*NBS DNA fragment. These entry clones were transformed into *E. coli* DH10β competent cells and positive clones were obtained on a kanamycin selection plate. Chloramphenicol was used for counterselection of positive clones. These clones were further used to perform a LR recombination reaction between an *att*L-containing entry clone and *att*R-containing pET301/CT-DEST destination vector (150 ng/μL) with a histidine tags and T7 promoter, lac operator and *att*B recombination site while C-terminal contain 6x His-tag, T7 reverse priming site and T7 terminator. Positively transformed E. coli DH10β competent cells were selected from ampicillin-Luria-Bertani (LB) media plates incubated overnight at 37°C. The positive colonies were obtained using 100 μg/mL ampicillin and confirmed by sequencing.

Protein expression. Above extracted plasmid were transformed into *E. coli* RosettaTM (DE3) pLysS host strain for protein expression using the heat shock transformation protocol (Chan, Verma et al. 2013). Transformed cells were placed onto Luria-Bertani (LB) media with ampicillin and chloramphenicol selection plates overnight

at 37°C. A single colony was picked and grown overnight at 37°C at 200 rpm in 7 mL LB broth containing ampicillin (100 μg/mL) and chloramphenicol (34 μg/mL). The overnight grown pre-culture was diluted 1:100 in fresh LB broth containing ampicillin (100 μg/mL) and chloramphenicol (34 μg/mL) and grown at 200 rpm at 37°C to an A_{600} between 0.5-0.8. Isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.9 mM to induce protein expression. The culture was incubated for 8 hrs at 37°C at 200 rpm. Supernatant and pellet of IPTG induced bacterial culture were separated by centrifugation at 9032X*g* for 15 minutes and stored at -80°C. *E. coli* crude cell pellet were used to purify NpNBS protein. Non-induced protein (not treated with IPTG) was also collected from bacterial pellet.

Protein purification and Western blotting. Protein purification was performed by resuspending cell pellet, obtained from IPTG induced *E. coli* Rosetta (DE3)pLys cell cultures, in 10 mg/mL lysozyme and lysis buffer containing 50 mM NaH_2PO_4, 300 mM NaCl and 10 mM imidazole pH 8 and incubated on ice for 30 minutes. After sonication and centrifugation at 10,000x*g* for 20 minutes at 4°C, supernatant containing the 6x His-tagged protein was collected (Lysate 1), which was loaded on pre-equilibrated nickel affinity spin columns from Qiagen (Germany) and centrifuged at 890x*g* for 5 minutes. Clear lysate (Lysate 2) was collected and saved for SDS-PAGE analysis. Nickel affinity columns were washed twice with 600 μL wash buffer containing 50 mM sodium phosphate monobasic (NaH2PO4), 300 mM sodium chloride (NaCl) and 20 mM imidazole pH 8 by centrifugation at 890x*g* for 2 minutes and flow thorough were saved as wash 1 and wash 2 for SDS-PAGE analysis. At last, protein was eluted twice (elution 1 and 2) each time in 300 μL elution buffer with 50 mM NaH2PO4, 300 mM NaCl and 300 mM imidazole pH 8 by centrifugation at 3000 rpm for 2 minutes and eluate was collected. Protein quantification was done according to Bradford assay (Bradford 1976). Five different BSA concentrations (5, 10, 15, 20, 25 and 50 ug) were prepared and absorbance was determined at 595 nm. A graph was plotted using above BSA concentrations (mean of three) at x-axis and their absorbances at y-axis respectively, which gave a linear equation of straight line ($Y=0.112X + 0.0357$) and extinction coefficient ($R^2 = 0.999$). The NBS protein concentration (x) was calculated using above equation by replacing Y with

NBS absorbance recorded at 595 nm. Protein was resolved on 15% Mini-protean TGX stain-free precast gels. Gels were transferred on polyvinylidene difluoride (PVDF) membrane, equilibrated with TBS buffer (20 mM Tris, 150 mM NaCl pH 7.5) for 15 minutes on a rotatory shaker, followed by blocking of membrane for 1 h. with Tris-buffered saline (TBS) containing tween 20 (TBST) and 1% bovine serum albumin (BSA). PVDF membrane was incubated overnight at 4°C in TBST with 1% BSA containing 6x-His epitope tag antibody in 1:1000 dilution. After primary antibody incubation, membrane washed five times each for five minutes in TBST buffer and incubated for 1 h. in TBST containing 2.5% dry milk and goat anti-mouse horseradish peroxidase (GAM)-HRP conjugate in 1: 20,000 dilution. Immunoblot was washed six times for 5 minutes each in TBST buffer and developed using clarity Western ECL substrate from Bio-rad (USA).

Nuclear magnetic resonance spectroscopy and mass spectrometry. Proton and carbon NMR spectra were recorded on a Varian 200 MHz NMR apparatus. Chemicals shifts (δ) are recorded in parts per million (ppm). Coupling constant are expressed in Hz. Samples were dissolved in DMSO-d6 for data acquisition (δ 2.49 ppm for 1H NMR and 39.95 ppm for 13C NMR) using TMS as internal standard (δ 0.00 ppm). Multiplicities are described by the following abbreviations: s for singlet, d for doublet, dd for doublet of doublets, t for triplet and m for multiplet. Mass spectral were obtained from NanoQAM (Université du Québec à Montréal) using a Time-of-Flight LC/MS (LC/MS-TOF), Agilent Technologie, LC 1200 Series/6210 TOF-LCMS with electrospray ionization and positive mode (ESI+).

Norbelladine synthesis. Norbelladine was synthesized using a previously published protocol with modifications (Park 2014). A summary of the method is described below along with the result of the NMR spectral data (Figure D.1).

Step A: Synthesis of the imine norcraugsodine

An equimolar quantity of 3,4-dihydroxybenzaldehyde (251.6 mg, 1.82 mmol) and tyramine (249.8 mg, 1.82 mmol) were added as powders to a flask containing dichloromethane (7 mL). The solution was stirred gently for 6 hours at room temperature

to give the imine intermediate. The solvent was evaporated under reduced pressure with a rotatory evaporator and then, with a mechanical pump to remove solvent residue and water. The product obtained was sufficiently pure to be used as such in the next step. Crude yield, 99%; Yellow solid.

1H NMR (200 MHz, DMSO-d6) δ: 7.98 (1H, s, CH imine), 7.15 (1H, d, J = 2.0 Hz, CH-Ar), 6.99 (2 H, d, J = 8.6 Hz, 2 x CH-Ar), 6.90 (1 H, dd, J1= 2 Hz and J2 = 8.2 Hz, CH-Ar), 6.65 (1 H, d, J = 8.6 Hz, CH-Ar), 6.63 (2 H, d, J = 8.6 Hz, 2 x CH-Ar), 3.63 (2H, t, J = 7.2 Hz, CH=NCH2CH2), 2.73 (2H, t, J = 7.4 Hz, CH=NCH2CH2); 13C NMR (200 MHz, DMSO-d6) δ: 160.85, 155.90, 149.49, 146.08, 130.46, 130.12, 127.76, 121.88, 115.77, 115.42, 113.85, 62.41, 36.76; ESI+ HRMS: (M+H)+ calculated for C15H16NO3 = 258.1125; found = 258.1078.

Step B: Synthesis of norbelladine

The imine norcraugsodine (50.6 mg, 0.98 mmol) dissolved in methanol (5 mL) is hydrogenated to the amine norbelladine using 30 mol% palladium on carbon (Pd/C 10 %) under a H_2 atmosphere. The hydrogen is bubbled three times (t = 0, 30 and 60 minutes) during the hydrogenation process. The mixture is agitated for a total of 2 hours and then filtered on a silica gel to remove the Pd/C and impurities. Methanol was evaporated under reduced pressure with a rotatory evaporator and then, with a mechanical pump to give norbelladine. The product is a brownish solid. Yield: 98%.

1H NMR (200 MHz, DMSO-d6) δ: 6.94 (2 H, d, J = 8.6 Hz, CH-Ar), 6.63 (4 H, m, 4 x CH-Ar), 6.52 (1H, dd, J1 = 1.7 Hz and J2 = 7.7 Hz, CH-Ar), 3.49 (2H, s, Ar-CH2-NH), 2.58 (4 H, m, NH-CH2CH2-Ar); 13C NMR (200 MHz, DMSO-d6) δ: 155.8 , 145.4, 144.3, 132.1, 130.9, 129.8, 119.2, 116.0, 115.6, 115.5, 53.1, 51.1, 35.4; ESI+ HRMS: (M+H)+ calculated for C15H18NO3 = 260.1281; found = 260.1182.

Enzyme assays. Screening assays contained 10 μg of purified protein, 10 μM tyramine, 300 μM 3,4-DHBA, 100 mM Tris buffer in a total volume of 90 μL. The assay was incubated at 37°C for 2 hrs followed by termination using 3 μL of 20% trichloroacetic

acid (TCA). Negative control includes non-induced proteins from *E. coli* (0.0 mM IPTG) and purified *Np*NBS protein boiled at 95°C for 15 minutes.

The analysis of the enzymatic product of NBS was performed on a Waters 2690 high performance liquid chromatograph coupled to a Micromass Quattro LC mass spectrometer using a Kinetex C18 column (150 mm long x 4,6 mm inside diameter, 5 μM particle size). Samples were subjected to positive-mode electrospray ionization (ESI[+]) liquid chromatography (Zulak, Cornish et al.)-tandem mass spectrometry (MS/MS) for reaction product characterization, including collision-induced dissociation (CID) fragmentation analysis. 10 μL of each sample was injected onto the column and compounds were eluted at a flow rate of 0.25 mL/min using ammonium acetate 10 mM, pH 5.0 (solvent A) and acetonitrile 100% (solvent B). The LC program started with 40% solvent B, a gradient began at 0 min to 98% at 7 min, 98% at 9 min, 40% at 10 min, and 40% at 11 min. Total run time was 12 minutes per sample. Analytes were detected using a triple-quadrupole mass analyzer operating in positive ion mode (ESI+). For MS/MS analyses, norbelladine and norcraugsodine standards were characterized by the isolation of the parent mass in Q1, the specific fractionation of the parent molecules in a collision cell at a selected energy in q, and finally the scan of the characteristic ions fragments in Q2. The conditions of the MS/MS section were set to acquire in positive ion mode as follows: desolvation gas flow rate 708 L/hr, desolvation gaz temperature 400°C, source temperature 120°C, capillary voltage 1000 V, cone voltage 15 V, scan mass range from 100 to 265 +ESI and collision energy of 0-30 V. MassLynx software from Waters was used for data acquisition and processing.

Standard norbelladine (m/z 260) was analyzed using MS mode where the first two quadrupoles were set to radio frequency (RF) only and the third quadrupole scanned the mass range of 253-265 m/z. The obtained mass-to-charge (m/z) value and retention time were used to develop subsequent collision-induced dissociation (CID) experiments. Fragmentation spectrum was obtained using daughter mode at optimized collision energy and the third quadrupole scanned the mass range of 50-275 m/z. Norbelladine eluted at 5.5 minute with five significant [M+H]+ in daughter ion mode (260, 159, 138, 123, 121).

Abundant fragment 138 m/z (using cone voltage 20V and collision energy 10ev) and 123 m/z (cone voltage 15V and collision energy 20eV), 121 m/z (cone voltage 15V and collision energy 20eV) were optimized for qualifier and quantifier analysis in multi reaction monitoring mode (MRM) based on their signal intensity at an applied voltage (collision energy). MRM was used to measure the intensity of selected fragments to mark them as qualifier and quantifier ion. Norcraugsodine also elute at the similar retention time as norbelladine (5.5 min) with two significant [M+H]+ in daughter ion mode (258, 257 and 121). The fragment ion m/z 121 was detected in both standards and considered as the quantitative ion for analysis. CID fragmentation spectra for standards are available in figure D.2. Qualitative ions for norbelladine and norcraugsodine are m/z 138 and m/z 157 respectively for validation. The MRM method thus developed was subsequently used to analyze the enzymatic essay samples. Quantification of standards and enzyme assay were performed using integration function of m/z 121 (quantifier daughter ion) on MassLynx software to obtain area under the chromatogram peak. Quantification of norbelladine in the enzyme essay sample was achieved by using the standard curve of norbelladine which has a linear equation of y=31264x+2879 and a correlation coefficient (R2) of 0.999. Solutions of concentration between 0 and 10 ppm of norbelladine were prepared and each of them have been injected in triplicate in the LC-MS using the optimized MRM method for the quantitative daughter ion of norbelladine. The area under the chromatographic peaks for each solution was obtained by using the integration function of MassLynx software. Enzyme assay product was confirmed by comparison of LC-ESI-MS/MS data to standards.

3.5 Results

Candidate gene identification. RNA from *N. pseudonarcissus* 'King Alfred' bulb was isolated and corresponding cDNA was used for Illumina sequencing, which provided a deep transcriptome database to search for expressed genes encoding enzymes involved in AA biosynthesis (Singh and Desgagné-Penix 2017). To identify ortholog of *NCS* in *N. pseudonarcissus*, we performed a mining of previously obtained trinotate annotated transcriptome data which was aligned using sprot_Top_BLASTx search tools

for sequence similiarity search with *NCS* orthologs (*Thalictrum flavum, Papaver somniferum, Argenome Mexicana, Eschscholzia californica, Chelidonium majus and Coptis japonica)*. Our candidate gene selection was based on open reading frame (ORF) and translation product. Three resulting orthologous full-length/longest transcripts obtained were TR17354|c0_g1_i1, TR17354|c0_g1_i2 and TR17354|c0_g1_i3 consisiting of 593, 575 and 583 bp with corresponding open reading frames of 492, 480 and 483 bp, respectively. Their E value ranging between $1e^{-20}$ to $4e^{-20}$ and their corresponding translated amino acid sequences were ~97% identical. The candidate displaying a longest transcript length of 593 bp with predicted ORF of 492 bp was named *N. pseudonarcissus Norbelladine Synthase (NpNBS)* and was used for subsequent functional analysis. The ORF encoded protein consisted of 163 amino acids with an anticipated molecular weight of 19 kDa and a theoretical pI of 5.5. BLAST and motif scan analyses revealed the presence of conserved Bet v1 and Pathogenesis-Related (PR-10) protein domain. The PR-10/Bet v1 allergen proteins do not possess NCS activity although they share similarity to NCS protein. Thus, *Np*NBS, similarly to NCS, belongs to the PR10/Bet v1 protein family.

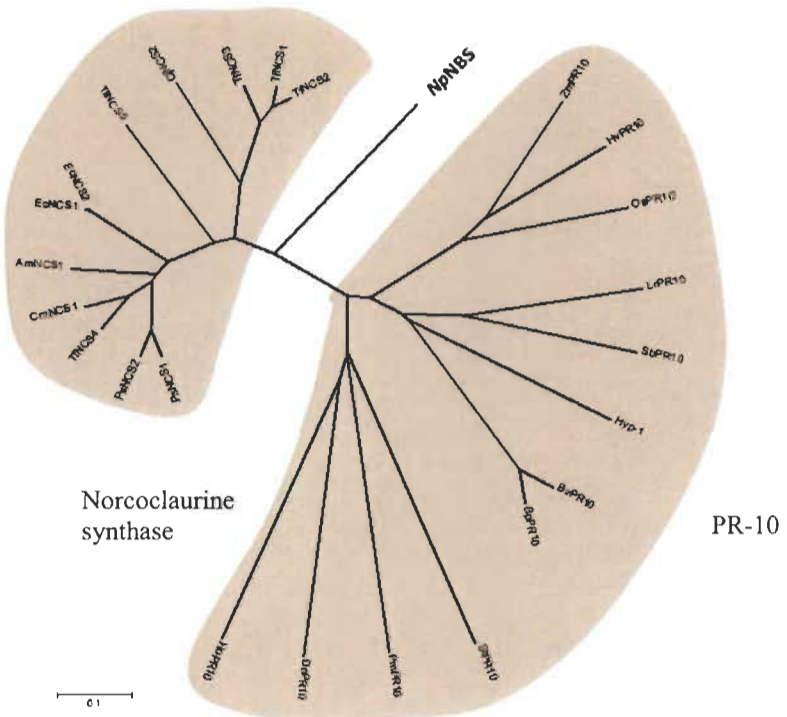


Figure 3.2 Phylogenetic relationship among several PR10/Bet v1 proteins from a variety of plants

Phylogenetic analysis. To investigate the evolutionary history of *Np*NBS and its relationship with PR-10/Bet v1 proteins, a phylogeny tree was constructed with the predicted amino acid sequence of *Np*NBS. Although all proteins are member of the PR-10 protein family, the phylogenetic tree was clearly divided into three main clusters: the PR-10 from monocots, PR-10 from dicots and the NCS (norcoclaurine synthases) (Figure 3.2). This type of clustering was reported previously (Lee and Facchini 2010). Notably, *Np*NBS shares higher homology with proteins from the NBS cluster including 38% amino acid sequence identity with *Thalictrum flavum Tf*NCS1/2 and 41% with *Papaver somniferum Ps*NCS1/2. *Np*NBS displayed comparatively lower identity to proteins from the PR-10 from dicots cluster including 19% identity to *Betula verrucosa Bv*PR10 and only 7% to *Hyacinthus orientalis Ho*PR10. Although *Np*NBS sequence appears to be at the edge of the PR-10 from monocots/NCS cluster, it is more closely related with the NCS group since it showed higher homology with *Tf*NCS (Figure 3.2). Altogether the results suggest that NBS, NCS and PR-10 evolved as individual gene over time.

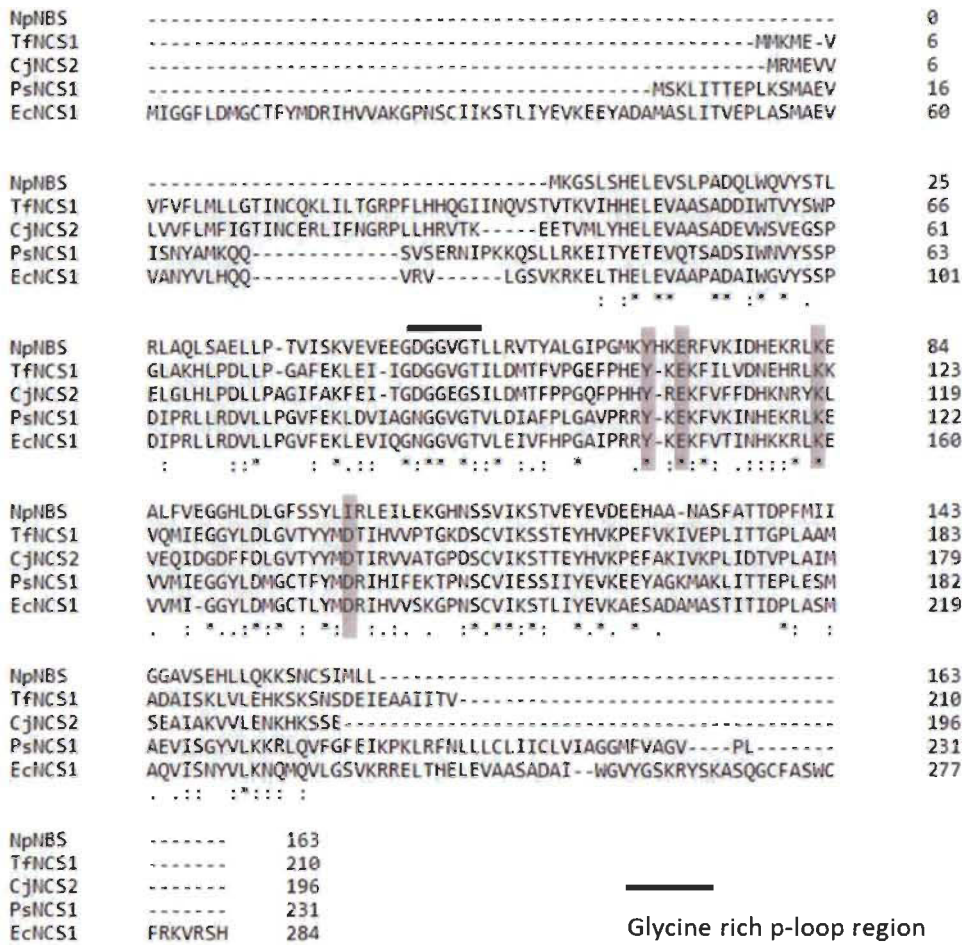


Figure 3.3 Clustal omega alignment of deduced amino acid sequence of *Narcissus pseudonarcissus* norbelladine synthase (*Np*NBS) with *norcoclaurine synthase* (NCS) amino acid sequences obtained from different species.

To identify regions of similarity that may be important for functional and structural relationship, multiple sequences alignment were performed using *Np*NBS and the NCS amino acid sequences from various plant species (Figure 3.3). Thus, *Ec*NCS1, *Ps*NCS1, *Cj*NCS2 (formely *Cj*PR10A) and *Tf*NCS1 genes shared over 38% identity in amino acid sequences with each other, while they showed the same (38% identity) with *Np*NBS. Alignment also showed the presence of NCS conserved catalytic residues Tyr^{108}, Glu^{110}, Lys^{122} and Asp^{141} (Figure 3.3). The residues were renamed as Tyr^{68}, Glu^{71}, Lys^{83}, according to their position in *Np*NBS protein sequence. However, the NCS Asp^{141} was found missing in *Np*NBS sequence. All aligned amino acid sequences contained a phosphate-binding loop (P-loop) glycine rich region (Figure 3.3). The glycine rich loop is

the conserved ligand-binding domain of Bet v1 protein family (Vilmolmadkang et al. 2016). Subcellular localization of the *Np*NBS amino acid sequence was predicted using several programs, including SignalP4.1 and WoLF PSORT, and the result indicated the absence of signal peptides. In addition, localization predictor program (PSLpred) identified *Np*NBS as a cytosolic protein. Altogether, the data suggest that *Np*NBS proteins are probably localized in cytoplasm.

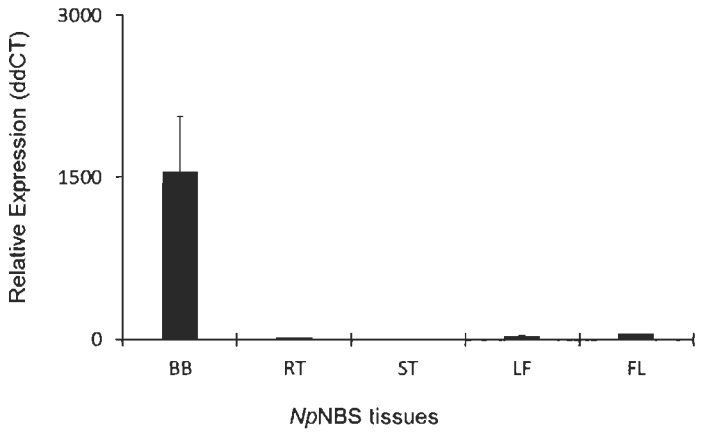


Figure 3.4 Quantitative real-time PCR results of *NpNBS* gene in different tissues of *N. pseudonarcissus* 'King Alfred'. Graph is plotted using normalized ddCT values scaled to lowest. Histone was used for internal reference. Expression fold change and error bars were calculated using the comparative $2^{-\Delta\Delta Ct}$ method from three independent experiments. Bars represent the mean standard deviation of three independent replicates. Abbreviations: tissues are BB, bulb; RT, root; ST, stem; LF, leaf; FL, flower.

RT-PCR analysis. The study of *Np*NBS expression profile in different tissues of *N. pseudonarcissus* 'King Alfred' by quantitative Real-time PCR revealed high expression of *Np*NBS in bulbs compared to roots, stems, leaves and flowers (Figure 3.4). The normalized ddCT expression of *Np*NBS in other tissues was detected below 60, whereas, its expression in bulbs was approximately 1500 folds higher, suggesting the importance of AA metabolism in *N. pseudonarcissus* bulbs.

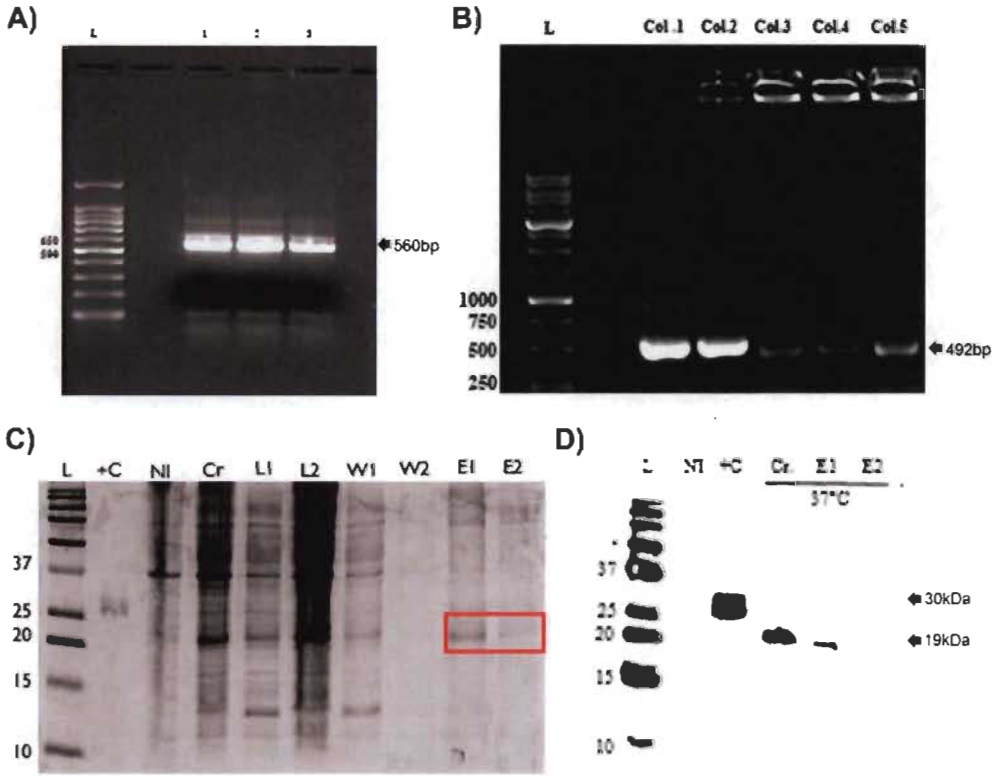


Figure 3.5 Heterologous expression of *Np*NBS in *E. coli*.

***Np*NBS gene cloning and heterologous expression.** To produce recombinant *Np*NBS, a full-length *Np*NBS cDNA was PCR amplified from *N. pseudonarcissus*. The amplified product of *Np*NBS shows a band of 566 bp on the 1% agarose electrophoresis gel (Figure 3.5A). *Np*NBS was gel purified, cloned into pET301 expression vector with C-terminal 6x-His tag, transformed into *E. coli* competent cells and positive colonies selected on LB with ampicillin and chloramphenicol. Colony PCR confirmed the presence of positive colonies (Figure 3.5B) which were induced with IPTG at 37°C and protein fractions were extracted. SDS-PAGE analysis was performed with different fractions obtained during *Np*NBS protein purification process which shows bands of different sizes, suggesting the presence of numerous bacterial cell proteins in fractions from the crude (Cr), lysate (L1 and L2) and wash buffer (W1 and W2). An apparent molecular mass of 19 kDa in elution buffer 1 (E1) was observed, which was absent in non-induced protein (NI) (Figure 3.5C). Western (immunoblot) analysis was performed using 6x His-tag monoclonal antibody. A His-tag *Np*NBS crude (Cr) and purified extracts (E1) isolated

from transformed bacterial cell culture grown at 37°C show molecular weight of 19 kDa. No protein expression was observed in cultures without IPTG (Figure 3.5D). The results showed successful production of recombinant *Np*NBS from IPTG-induced bacterial cell culture.

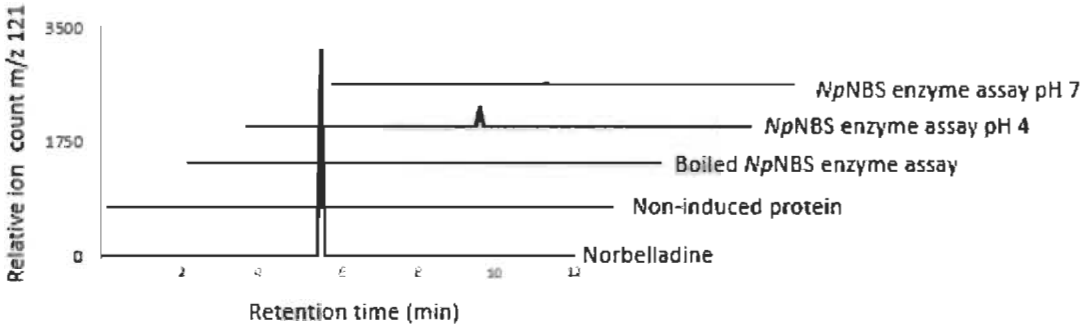


Figure 3.6 Extracted ion chromatogram showing the product of *Np*NBS assays

***Np*NBS enzyme assay.** For the biosynthesis of all AA, the precursor norbelladine is made from the condensation of tyramine and 3,4-DHBA which forms the imine norcraugsodine followed by its reduction to norbelladine. It is not clear if norbelladine formation is one stepwise reaction catalyze by a single enzyme or two separate reactions (condensation and reduction) catalyzed by different enzymes. To confirm *Np*NBS protein function, enzyme assays were carried out with purified *Np*NBS recombinant protein. The resulting assay product was subjected to LC-MS/MS analysis using a Positive Electrospray Ionization mode (+ESI). The QqQ dual +ESI source conditions were optimized using freshly synthesized alkaloid standards, norbelladine and norcraugsodine (Fig. D.1) to obtain a good signal and high sensitivity. The MS/MS parameters such as capillary voltage, spray voltage and skimmer voltage were enhanced to maximize the ionization in the source and sensitivity to identify and characterize all possible fragmentation products. With standards, we observed predicted major mass spectral fragments for the nobelladine m/z 260 [M + H]+ and norcraugsodine m/z 258 [M+H], both at 5.5 min (Fig. D.2). Fragmentation of norbelladine molecular ion m/z 260 [M+H]+ yielded ion fragments of m/z 159, 138, 123 and 121 (Fig. D.2). The ion fragment m/z 138 was obtaine by the elimination of the 4-ethylphenol moiety (122 Da) whereas m/z 121

was produced by loss of tyramine. Precisely for norbelladine, the transition [M+H]+ of 260→138 was selected for the qualifier ion fragment and the transition [M+H]+ of 260→121 was used as the quantifier ion (Fig D.2). Similar MS/MS ion fragments spectra were reported for norbelladine suggesting a good fragmentation of our standard (Kilgore, Holland et al. 2016).

Enzyme assays with *Np*NBS yielded a peak at 5.5 minute on LC-MS which was the same retention time than those of the norbelladine and norcraugsodine standards. The +ESI-MS/MS analysis of the *Np*NBS enzyme assay product showed the presence of the qualifier and quantifier ion fragments of norbelladine (Figure 3.6). The level of product was detected in the *Np*NBS assay pH 4 and at lower levels at pH 7. Norcraugsodine molecular ion and fragment ions were not detected in the assay. We obtained a low background reaction due to non-specific condensation reaction. The reaction product was also present at low level in control assay/non-induced protein assay (~3.05e3 peak intensity) compared to NBS assay at pH 4 (~2.09e5 peak intensity). *Np*NBS enzyme assay product shows that *Np*NBS catalyzes the condensation reaction between tyramine and 3,4 DHBA to produce norbelladine. Thus, our in vitro enzyme assay demonstrated the ability of *Np*NBS as norbelladine synthase with a preference at pH 4.

3.6 Discussion

The formation of norbelladine is a crucial step in the formation of Amaryllidaceae alkaloids. In this study, we identified the enzyme norbelladine synthase in wild daffodil (*N. pseudonarcissus*) responsible for the condensation of tyramine with 3,4-dihydroxybenzaldehyde (3,4-DHBA) to form norbelladine. *Np*NBS contains conserved Bet v1 domain and thus is a member of the Bet v1/PR10 protein family. The PR-10 family universally exits in plants of monocots and dicots group (Wen, Vanek-Krebitz et al. 1997, Wang, Huang et al. 1999, Yu, Ekramoddoullah et al. 2000). Apart from the PR-10 family, recently Bet v1 have shown an extended low level similarity with three more protein families; major latex and ripening-related proteins, norcoclaurine synthase proteins, and cytokinin-binding proteins from legumes (Radauer and Breiteneder 2007). Several studies

have also demonstrated that members of the Bet v1/PR-10, specifically of the NCS group, have a crucial role in alkaloid production (Samanani, Park et al. 2005, Zulak, Cornish et al. 2007). For example, *Ps*NCS from *Papaver sominferum* shares significant amino acid identity with PR-10 and Bet v1 protein and NCS has a major role in benzylisoquinoline alkaloids biosynthesis. Indeed, Virus-Induced Gene Silencing (VIGS) of PsNCS1 and PsNCS2 transcripts resulted in a 75% to 82% decrease of main alkaloids in *Papaver sominferum* latex compared to control (Lee and Facchini 2010).

NCS catalyzes the condensation of dopamine with 4-hydroxyphenylacetaldehyde (4-HPAA) to yield (S)-norcoclaurine, the central precursor of benzylisoquinoline alkaloids. *Np*NBS was identified from transcriptome data search using NCS sequence homology. *Np*NBS shared 41% amino acid sequence identity with *P. somniferum* NCS1 and NCS2 and lacked a signal peptide. Motif search suggested that *Np*NBS is a cytosolic protein and is not associated with a subcellular compartment such as endoplasmic reticulum, mitochondria or peroxisomes. The observed catalytic activity of *Np*NBS is possibly due to a lack of signal peptide, which is not necessarily required for enzyme's functional activity. Our results are consistent with studied conducted on *Tf*NCS which reported that a signal peptide had no role in the catalytic activity of an enzyme. Samanani et al. (2004) studied the activity of *Tf*NCS protein encoding truncation of first 10 and 19 N-terminal amino acids which resulted in formation of (S)-norcoclaurine. Both recombinant protein displayed enzymatic activity (Samanani, Liscombe et al. 2004). In addition, a study conducted on four N-terminally truncated variants of *Cj*NCS from first 10, 19, 29 and 42 amino acids revealed an increasing fold of enzymatic activity. For example, *Cj*NCS-Δ19 displayed 10-fold (0.30 U/mL) and *Cj*NCS-Δ29 had shown 40-fold (1.20 U/mL) higher activity compared to full length *Cj*NCS (0.03 U/mL) (Nishihachijo, Hirai et al. 2014). However, *Cj*NCS-Δ40 show lowest activity among all constructs (3-fold) which might be due to deletion of some important amino acid residues required for enzyme-substrate interaction. Recently, Li et al. (2016) reported on the isolation and characterization of NCS variants from several alkaloid-producing plant species. They found that cDNAs encoding NCS orthologs possess two, three or four repeated catalytic

domains containing catalytic residues Y,E,K and D. The presence of repeat domains was found proportional with increase in catalytic efficiency.

*Np*NBS enzyme catalyzes norbelladine synthesis by a condensation process between amine and aldehyde. The Glu^{110}/Glu^{71} was detected with +1 histidine shift in *Np*NBS might affect the interaction between Glu and Tyr residues and hinders the cyclization process. Moreover, *Np*NBS abilty to form norbelladine in absence of Asp141 catalytic residues suggests that Asp^{141} residue is not responsible for base catalysis and might be involved in imparting electrostatic stabilization (Lichman, Gershater et al. 2015).

*Np*NBS enzyme has optimum activity at pH 4 indicating that the reaction prefers an acidic environment. Similar pH of 6.2 has been reported for optimum acitivity of *Ps*NCS (Samanani and Facchini 2001). Further study could be performed by adding more pH conditions ranging from pH 4 to pH 8, to identify ideal pH requirement of *Np*NBS enzyme.

qPCR- This data is also supported by its high FPKM value (31543) obtained in the transcriptome. It was reported among the top expressed genes, which altogether comprises 16.12% of the *N. pseudonarcissus* 'King Alfred' transcriptome. *Np*NBS high expression was consistent with the expression of *NpN4OMT* in bulbs, which supported the abundant occurrence of AAs in bulb compared to other tissues, suggesting bulbs as an important site of enzyme localization and AAs synthesis (Singh and Desgagné-Penix 2017). Kilgore et al., 2014 also presented a similar correlation between high expression and read count of *N4OMT* with abundant galanthamine accumulation in bulb compared to other tissues (Kilgore, Augustin et al. 2014). Therefore, based on the expression data, we conclude that *Np*NBS has a crucial role in AAs metabolism. A similar High *Tf*NCS expression and enzymatic activity was reported in underground rhizome and root tissues compared to other parts, suggesting tissues a specific localization of NCS enzyme (Samanani, Liscombe et al. 2004).

The rapid progress in the discovery of AAs biosynthesis pathway genes along with newly characterized *Np*NBS enzyme is useful for reconstituting a short synthetic AAs biosynthesis pathway in yeasts or microbes that will allow feasible and high-scale production of targeted medicinally important Amaryllidaceae alkaloids which are gaining researchers interest for their new pharmaceutical applications, including an Alzheimer's treatment drug, galanthamine. The applied strategy could be implemented in characterization of genes from any class of specialized metabolites as our discovery shows the utility of the designed workflow.

3.7 Acknowledgement

We thank Professor Hugo Germain for helpful discussion and for the gateway cloning vectors systems and other lab materials. We thank Ariane Garand for help with recombinant protein expression, purification and assay. We also thank Professor André Lajeunesse for his helpful comments and suggestions to optimize the LC-MS/MS detection method. This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) award number RGPIN 05294-2014 (Discovery) to IDP. This work was also supported by the NSERC award number EQPEQ 472990-2015 (Research tools and instruments) for the acquisition of the HPLC-PDA and the qPCR.

3.8 Figure Legends

Figure 3.1: Norbelladine synthase (NBS) catalyzes the condensation of tyramine and 3,4-dihydroxybenzaldehyde (3,4-DHBA) to form norbelladine, the common precursor to all Amaryllidaceae alkaloids produced in plants including galanthamine, lycorine and haemanthamine.

Figure 3.2: Phylogenetic relationships among several PR10/Bet v 1 proteins from a variety of plants. Phylogeny tree was based on the amino acid sequences for *Betula verrucosa Bv*PR10, *Betula platyphylla Bp*PR10, *Papaver somniferum Ps*NCS, *Ps*NCS2, *Thalictrum flavum Tf*NCS, *Tf*NCS2,*Tf*NCS3, *Tf*NCS4, *Tf*NCS5, *Chelidonium majus Cm*NCS1, *Argenome mexicana Am*NCS, *Eschscholzia californica Ec*NCS1, *Ec*NCS2,

Coptis japonica *Cj*NCS2 (formerly *Cj*PR10A), *Daucus carota* *Dc*PR10, *Solanum tuberosum* *St*PR10, *Hyacinthus orientalis* *Ho*PR10, *Pinus monticola* *Pm*PR10, *Hordeum vulgare* *Hv*PR10, *Oryza sativa* *Os*PR10, *Lily regale* *Lr*PR10, *Sorghum bicolor* *Sb*PR10, *Zea mays* *Zm*PR10 and *Hypericum perforatum* Hyp-1. Accession numbers provided in table D.1. The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei 1987) with bootstrap value 500. The optimal tree with the sum of branch length = 5.78745187 is shown. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the p-distance method (Nei and Kumar 2000) and are in the units of the number of amino acid differences per site. The analysis involved 25 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 139 positions in the final dataset. Evolutionary analyses were conducted in MEGA 6 (Tamura, Stecher et al. 2013).

Figure 3.3: Clustal omega alignment of deduced amino acid sequence of *Narcissus pseudonarcissus* norbelladine synthase (*Np*NBS) with *norcoclaurine synthase* (NCS) amino acid sequences obtained from different species: *Eschscholzia californica* *Ec*NCS1, *Thalictrum flavum* *Tf*NCS1, *Papaver somniferum* *Ps*NCS1, and *Coptis japonica* *Cj*NCS2 (formely *Cj*PR10A). Grey boxes corresponding to Tyr, Lys, Asp and Glu residues forms the *NpNBS* catalytic residue. The black bar depicts a glycine rich P-loop domain conserved in *NCS* and *NpNBS* protein. The fully conserved residues are marked with an asterisks (*). The positions with conservation between amino acid residues of similar properties are marked with a colon (:) and the positions with conservations between amino acids residues of weakly similar properties are marked with a period (.). Numbers on the right indicate the residues position in the sequence.

Figure 3.4: Quantitative real-time PCR results of *NpNBS* gene in different tissues of *N. pseudonarcissus* 'King Alfred'. Graph is plotted using normalized ddCT values scaled to lowest. Histone was used for internal reference. Expression fold change and error bars were calculated using the comparative $2^{-\Delta\Delta Ct}$ method from three independent experiments. Bars represent the mean standard deviation of three independent replicates. Abbreviations: tissues are BB, bulb; RT, root; ST, stem; LF, leaf; FL, flower.

Figure 3.5: Heterologous expression of *Np*NBS in *Escherichia coli*. (A) Agarose gel electrophoresis of PCR amplified candidate *Np*NBS from *N. pseudonarcissus* cDNA library. Product size is 560 bp (ORF 492bp + Gateway adapters 68bp). Numbers on the left refer to the location of standard DNA molecular markers in bp. (B) Agarose gel electrophoresis of PCR amplified *Np*NBS performed with 5 positively transformed bacterial colonies (col.1 to col.5) shows a band of *Np*NBS ORF transcript size of 492 bp, L is for Ladder; (C) SDS-PAGE analysis of *Np*NBS protein. Purified protein was extracted from 0.9 mM IPTG induced *E. coli* Rosetta™ (DE3) pLysS host strain at 37°C for 8h in elution 1 with 300 mM of imidazole. L-Ladder; +C- positive control with a 30 kDa protein; NI-non-induced protein; L1- lysate 1 (10 mM imidazole); L2- lysate 2 (10 mM imidazole); W1-wash 1 (20 mM imidazole); W2-wash 2 (20 mM imidazole); E1-elution 1 (300 mM imidazole); E2-elution 2 (300 mM imidazole). (D) Western blot analysis performed using 6x HIS-tag epitope antibody shows expression of recombinant protein at 37°C in 0.9 mM IPTG induced crude and pure protein extracts. No detectable expression of recombinant protein at 25°C in 0.9 mM IPTG induced crude and pure protein extracts. Lane L-ladder; NI- protein non-induced ; +C-positive control protein with C-terminal His tags 30 kDa; Cr-crude protein 37°C/8h/0.9 mM; E1- elution 1, 37°C/8h/0.9 mM; E2- elution 2, 37°C/8h/0.9 mM; Cr-crude 25°C/8h/0.9 mM; E1- elution 1, 25°C/8h/0.9 mM; E2-elution 2, 25°C/8h/0.9 mM. Numbers on the left refer to the location of standard protein molecular weight markers in kDa.

Figure 3.6: Extracted ion chromatograms showing the product of *Np*NBS enzyme assays. The tested substrates used were 3,4-dihydroxybenzaldehyde (300 μM) and tyramine (10 μM), the extracted ion chromatogram corresponds to standard NB or assays conducted with the non-induced protein fraction, the heat-denatured recombinant *Np*NBS enzyme; and the complete assay performed with recombinant *Np*NBS at pH 4 and pH 7. Parent ion mass-to-charge (*m/z*) of 260 for norbelladine and 258 for norcraugsodine were subjected to collision-induced dissociation analysis for identification and quantitation (Fig D.2).

3.9 References

Barton, D. and T. Cohen (1957). "Festschrift A. Stoll." Birkhiiuser, Basel 117.

Barton, D., G. Kirby, J. Taylor and G. Thomas (1963). "866. Phenol oxidation and biosynthesis. Part VI. The biogenesis of amaryllidaceae alkaloids." Journal of the Chemical Society (Resumed): 4545-4558.

Battersby, A., H. Fales and W. Wildman (1961). "Biosynthesis in the Amaryllidaceae. Tyrosine and norbelladine as precursors of haemanthamine." Journal of the American Chemical Society 83(19): 4098-4099.

Bradford, M. M. (1976). "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding." Analytical biochemistry 72(1-2): 248-254.

Chan, W.-T., C. S. Verma, D. P. Lane and S. K.-E. Gan (2013). "A comparison and optimization of methods and factors affecting the transformation of Escherichia coli." Bioscience reports 33(6): e00086.

Eichhorn, J., T. Takada, Y. Kita and M. H. Zenk (1998). "Biosynthesis of the amaryllidaceae alkaloid Galanthaminefn2." Phytochemistry 49(4): 1037-1047.

El Tahchy, A., A. Ptak, M. Boisbrun, E. Barre, C. Guillou, F. o. Dupire, F. o. Chrétien, M. Henry, Y. Chapleur and D. Laurain-Mattar (2011). "Kinetic study of the rearrangement of deuterium-labeled 4′-O-methylnorbelladine in Leucojum aestivum shoot cultures by mass spectrometry. Influence of precursor feeding on Amaryllidaceae alkaloid accumulation." Journal of natural products 74(11): 2356-2361.

Grisebach, H. (1973). "Comparative biosynthetic pathways in higher plants." Pure and Applied Chemistry 34(3-4): 487-514.

He, M., C. Qu, O. Gao, X. Hu and X. Hong (2015). "Biological and pharmacological activities of amaryllidaceae alkaloids." RSC Advances 5(21): 16562-16574.

Hotchandani, T. and I. Desgagne-Penix (2017). "Heterocyclic Amaryllidaceae Alkaloids: Biosynthesis and Pharmacological Applications." Current topics in medicinal chemistry 17(4): 418-427.

Jin, Z. and X.-H. Xu (2013). Amaryllidaceae Alkaloids. Natural Products. K. G. Ramawat and J.-M. Mérillon, Springer Berlin Heidelberg: 479-522.

Kilgore, M. B., M. M. Augustin, G. D. May, J. A. Crow and T. M. Kutchan (2016). "CYP96T1 of Narcissus sp. aff. pseudonarcissus Catalyzes Formation of the Para-Para'CC Phenol Couple in the Amaryllidaceae Alkaloids." Frontiers in plant science 7: 225.

Kilgore, M. B., M. M. Augustin, C. M. Starks, M. O'Neil-Johnson, G. D. May, J. A. Crow and T. M. Kutchan (2014). "Cloning and characterization of a norbelladine 4′-O-methyltransferase involved in the biosynthesis of the Alzheimer's drug galanthamine in Narcissus sp. aff. pseudonarcissus." PloS one 9(7): e103223.

Kilgore, M. B., C. K. Holland, J. M. Jez and T. M. Kutchan (2016). "Identification of a noroxomaritidine reductase with Amaryllidaceae alkaloid biosynthesis related activities." Journal of Biological Chemistry 291(32): 16740-16752.

Lee, E.-J. and P. Facchini (2010). "Norcoclaurine synthase is a member of the pathogenesis-related 10/Bet v1 protein family." The Plant Cell 22(10): 3489-3503.

Li, L., H.-J. Dai, M. Ye, S.-L. Wang, X.-J. Xiao, J. Zheng, H.-Y. Chen, Y.-h. Luo and J. Liu (2012). "Lycorine induces cell-cycle arrest in the G0/G1 phase in K562 cells via HDAC inhibition." Cancer cell international 12(1): 49.

Lichman, B. R., M. C. Gershater, E. D. Lamming, T. Pesnot, A. Sula, N. H. Keep, H. C. Hailes and J. M. Ward (2015). "'Dopamine-first'mechanism enables the rational engineering of the norcoclaurine synthase aldehyde activity profile." The FEBS journal 282(6): 1137-1151.

Liscombe, D. K., B. P. MacLeod, N. Loukanina, O. I. Nandi and P. J. Facchini (2005). "Erratum to "Evidence for the monophyletic evolution of benzylisoquinoline alkaloid biosynthesis in angiosperms"[Phytochemistry 66 (2005) 1374–1393]." Phytochemistry 66(20): 2500-2520.

Nei, M. and S. Kumar (2000). Molecular evolution and phylogenetics, Oxford university press.

Nishihachijo, M., Y. Hirai, S. Kawano, A. Nishiyama, H. Minami, T. Katayama, Y. Yasohara, F. Sato and H. Kumagai (2014). "Asymmetric synthesis of tetrahydroisoquinolines by enzymatic Pictet–Spengler reaction." Bioscience, biotechnology, and biochemistry 78(4): 701-707.

Park, J. B. (2014). "Synthesis and characterization of norbelladine, a precursor of Amaryllidaceae alkaloid, as an anti-inflammatory/anti-COX compound." Bioorganic & medicinal chemistry letters 24(23): 5381-5384.

Prvulovic, D., H. Hampel and J. Pantel (2010). "Galantamine for Alzheimer's disease." Expert opinion on drug metabolism & toxicology 6(3): 345-354.

Radauer, C. and H. Breiteneder (2007). "Evolutionary biology of plant food allergens." Journal of Allergy and Clinical Immunology 120(3): 518-525.

Saitou, N. and M. Nei (1987). "The neighbor-joining method: a new method for reconstructing phylogenetic trees." Molecular biology and evolution 4(4): 406-425.

Saliba, S., A. Ptak and D. Laurain-Mattar (2015). "4′-O-Methylnorbelladine feeding enhances galanthamine and lycorine production by Leucojum aestivum L. shoot cultures." Engineering in Life Sciences 15(6): 640-645.

Samanani, N. and P. J. Facchini (2001). "Isolation and partial characterization of norcoclaurine synthase, the first committed step in benzylisoquinoline alkaloid biosynthesis, from opium poppy." Planta 213(6): 898-906.

Samanani, N., D. K. Liscombe and P. J. Facchini (2004). "Molecular cloning and characterization of norcoclaurine synthase, an enzyme catalyzing the first committed step in benzylisoquinoline alkaloid biosynthesis." The Plant Journal 40(2): 302-313.

Samanani, N., S.-U. Park and P. J. Facchini (2005). "Cell type–specific localization of transcripts encoding nine consecutive enzymes involved in protoberberine alkaloid biosynthesis." The Plant Cell 17(3): 915-926.

Singh, A. and I. Desgagné-Penix (2014). "Biosynthesis of the Amaryllidaceae alkaloids." Plant Science Today 1(3): 114-120.

Singh, A. and I. Desgagné-Penix (2017). "Transcriptome and metabolome profiling of Narcissus pseudonarcissus 'King Alfred'reveal components of Amaryllidaceae alkaloid metabolism." Scientific reports 7(1): 17356.

Tahchy, A. E., M. Boisbrun, A. Ptak, F. Dupire, F. Chrétien, M. Henry, Y. Chapleur and D. Laurain-Mattar (2010). "New method for the study of Amaryllidaceae alkaloid biosynthesis using biotransformation of deuterium-labeled precursor in tissue cultures." Acta Biochimica Polonica 57(1): 75.

Tamura, K., F. U. Battistuzzi, P. Billing-Ross, O. Murillo, A. Filipski and S. Kumar (2012). "Estimating divergence times in large molecular phylogenies." Proceedings of the National Academy of Sciences 109(47): 19333-19338.

Tamura, K., G. Stecher, D. Peterson, A. Filipski and S. Kumar (2013). "MEGA6: molecular evolutionary genetics analysis version 6.0." Molecular biology and evolution 30(12): 2725-2729.

Van Loon, L. C., M. Rep and C. M. Pieterse (2006). "Significance of inducible defense-related proteins in infected plants." Annu. Rev. Phytopathol. 44: 135-162.

Wang, C.-S., J.-C. Huang and J.-H. Hu (1999). "Characterization of two subclasses of PR-10 transcripts in lily anthers and induction of their genes through separate signal transduction pathways." Plant Molecular Biology 40(5): 807-814.

Wen, J., M. Vanek-Krebitz, K. Hoffmann-Sommergruber, O. Scheiner and H. Breiteneder (1997). "The Potential ofBetv1Homologues, a nuclear multigene family, as phylogenetic markers in flowering plants." Molecular phylogenetics and evolution 8(3): 317-333.

Yu, X., A. K. Ekramoddoullah and S. Misra (2000). "Characterization of Pin m III cDNA in western white pine." Tree physiology 20(10): 663-671.

Zulak, K. G., A. Cornish, T. E. Daskalchuk, M. K. Deyholos, D. B. Goodenowe, P. M. Gordon, D. Klassen, L. E. Pelcher, C. W. Sensen and P. J. Facchini (2007). "Gene transcript and metabolite profiling of elicitor-induced opium poppy cell cultures reveals the coordinate regulation of primary and secondary metabolism." Planta 225(5): 1085-1106.

Chapter IV contains a discussion of results described in Chapter II and III and future perspectives.

CHAPTER IV

CONCLUSIONS

4.1 Conclusion

Amaryllidaceae is the family of bulbous plants comprising of popular ornamental species such as daffodils (*Narcissus*), spider lily (*Crinum*), and snowdrops (*Galanthus*). A large number of these alkaloids are divided into three major categories based on their C-C phenol coupling reactions (Bastida Armengol, Berkov et al. 2011). Few examples of these alkaloids with their biological properties are the acetylcholine esterase inhibiting galanthamine, the anti-tumor activity of lycorine and the cytotoxic haemanthamine, anti-fungal property of hippeastrine and an anti-viral effect of pancratistatin (Jin and Xu 2013, He, Qu et al. 2015). Some Amaryllidaceae generas holding interest of researchers for the study of their bioactive AAs includes *Galanthus, Narcissus, Leucojum, Lycoris, Haemanthus, Hippeastrum Clivia, Crinum and Pancratium* (Fennell and Van Staden 2001, Louw, Regnier et al. 2002, Kornienko and Evidente 2008, Rønsted, Savolainen et al. 2008, Cedrón, Del Arco-Aguilar et al. 2010, de Andrade, Pigni et al. 2012). However, only few genes of AAs biosynthesis have been characterized yet. Therefore, the elucidation of their biosynthesis pathway genes will lay a foundation for synthetic biology and development of new plant cultivars.

This research represents a potent workflow for Amaryllidaceae alkaloids biosynthesis pathway genes elucidation that could be implemented for other unidentified genes involved in the biosynthesis of valuable metabolites in non-model plants. The hypothesis of this integrated omics study was designed to overcome the hurdles involved in dealing with the study of higher non-model plant that produce important specialized metabolites through their complex network of interacting components. Since the genetic information for commercially significant non-model *N. pseudonarcissus* 'King Alfred' was lacking, we performed a *de novo* sequencing and assembly of *N. pseudonarcissus*

'King Alfred' transcriptome through Illumina paired-end sequencing which provided the essential data required for the study of AAs metabolism. Paired-end sequencing of both fragments generates a good quality alignable data. It provides a superior alignment than the single-end sequencing by providing longer contigs for *denovo* assembly. Moreover, PE sequencing could also detect insertion, deletion and inversions. The transcriptome data obtained from *Narcissus pseudonarcissus* 'King Alfred' adds information to previously existing few transcriptomes of *Narcissus* species and provides opportunity for identification and characterization of other full-length transcripts of AAs biosynthesis pathway. We preferred transcriptome study over genomics because plants possess high percentage of introns or noncoding sequences in their genome which hinders a precise prediction of coding sequence, wherease transcriptome represents only expressed transcripts or genes (Xiao, Zhang et al. 2013). Also, our metabolite-guided RNA-seq approach helped us to identify the unknown genes of AAs pathway from *N. pseudonarcissus* 'King Alfred' transcriptome. The results obtained from this research proves that the integration of transcriptome and metabolome data is an effective strategy for cataloging the component of specialized metabolism such as uncharacterized genes.

Considering the chemical diversity of specialized metabolites, a single metabolite analysis technique is not sufficiently sensitive for determining all cellular metabolites (Weckwerth 2003). Therefore, we used several metabolomic techniques such as TLC, HPLC, and UPLC-QTOF-MS for *N.pseudonarcissus* 'King Alfred' metabolites to extend our current knowledge on AAs profile. Based on the concentration and variety of AAs represented by several species, *N. pseudonarcissus* 'King Alfred' was selected for further study. It displayed an array of AAs including lycorine and galanthamine present in all tissues, and absence of narciclasine. For TLC, we used 254 nm and 365 nm to visualize amaryllidaceae alkaloids as some alkaloids possess phenol ring structure which floresce at 365nm. The result shows a large number of AAs are present in bulbs compared to other tissues. These results provided initial confirmation about the presence of AAs in *N. pseudonarcissus* 'King Alfred'. Similar alkaloid distribution pattern was also observed in different developmental stages of *Narcissus confusus* plant where the bulb was shown to accumulate different kinds of AAs either in low or high concentrations (López, Bastida et

al. 2003). Further, we analyzed our AAs extracts on UPLC-QTOF-MS in positive and negative ion mode to confirm their identity based on their molecular masses. Most of the AAs were identified in positive ion mode because they possess basic functional groups such as amine whose nitrogen could be protonated and detected in a positive ion mode. Analysis resulted in the identication of several compounds including amides, sequiterpenes, esters phenolic acids and amino acids. We detceted 14 AAs with their corresponding masses and relative abundance. Crinamine was present in all tissues, most abundantly in leaves (6.07 min., m/z 302.1341) followed by roots (6.09 min., m/z 302.1383). Lycorenine was most abundant in leaves (5.36 min., m/z 318.1866). Epilycoramine was found in all tissues, while lycorine and homolycorine were only present in the bulb. The bulb also shows presence of all three different structural type (galanthamine type, lycorine type, and crinamine type) of AAs. Intrestingly, 4-hydroxybenzaldehyde was detected in bulb and root while narciclasine was detcetd in bulb in negative ion mode analysis. 4-hydroxybenzaldehyde is a neutral compound whose oxygen atom could be protonated and hydroxyl group could be deprotonated. The possibility to remove the proton would be higher than accepting the proton by oxygen which resulted in its detection in negative ion mode. Detection of narciclasine by UPLC-QTOF-MS indicates that it could detect low concentration of analytes which is not possible through HPLC and thus narciclasine was absent in HPLC analysis. UPLC-MS-QTOF data also revealed the dominance of galanthamine-type and crinamine-type alkaloids in all the tissues while only lycorine-type was dominant in leaves. These results suggest active biosynthetic routes within different tissues of the plant.

De novo transcriptome assembly was performed after the detection of AAs in the tissues of *N. pseudonarcissus* 'King Alfred'. Bulb RNA was selected for Illumina HiSeq. 2000 sequencing as it displayed a broad range of known and unknown AAs. Additionally, various studies on AAs profiling have shown that bulb compared to other parts (stems, roots, stem discs, leaves and flowers) contains large number and quantity of alkaloids, therefore, its advantageous to select the target organ when searching for maximum alkaloids. For example, alkaloid profiling of three different species of Amaryllidaceae using HPTLC at different developemental stages revealed that *N. papyraceus* bulb contain

large amount of lycorine-type alkaloid compared to other parts. Moreover, two unique alkaloids were identified in the bulb which were not found in any other organs. Also, similar profiling was performed using *Narcissus tazetta* tissues showing presence of several alkaloids in different organs but vittatine and 11-hydroxyvittatine of the crinine series were unique only to the *N. tazetta L.* bulb (Shawky, Abou-Donia et al. 2015) Bulbs and leaves of Amaryllidaceae family plants possess high alkaloid content to protect their carbohydrate reserves from pathogens and herbivores (Ruiz, Ward et al. 2002, Gómez, Azorín et al. 2003).

A rich transcriptome data was obtained with 73,081,603 reads where we identified all previously characterized (*e.g. N4OMT, CYP96T1, NR*) and proposed (*e.g. C4H, C3H, HCT, 4CL)* genes of AAs by BLASTx analysis. Out of 195,347 transcripts, BLAST search could annotate only 11,708 transcripts, which represents a big number of uncharacterized genes or transcripts specific to *N. pseudonarcissus* 'King Alfred'. The 125 most expressed transcripts of the *N. pseudonarcissus* 'King Alfred' transcriptome belonged to the plant defense, metabolism, transport, and cellular structure category. The most expressed genes were from *Narcissus mosaic virus*, a common virus infecting Amaryllidaceae plants. Ornamental plants of Narcissus family are vegetatively propagated and accumulate viruses of genera *Carlavirus*, *Macluravirus*, *Nepovirus*, *Potexvirus* and *Potyvirus* (Brunt 1996, Wylie, Li et al. 2014). Genes involved in plant defense including metallothionein, catalase and peroxidsases were among the top expressed genes which might be due to virus infected plant. Such type of gene expression was also seen in Potato plant. Niehl et. al. analysed a defense response in potato X virus infected leaves of *Solanum tuberosum* and found that PR-1, salicyclic acid biosynthesis and glutathione genes were upregulated in infected plant compared to the control. These genes are primarily involved in ethylene production and ROS dependent signaling (Niehl, Lacomme et al. 2006). Comparision of the most abundant transcripts profile with *Papaver somniferum* revealed similar expression pattern demonstrating that genes involved in plant's fundamental process are relatively consistent. Interestingly, we detected *NCS*-like sequence at rank 62 suggesting its involvement in norbelladine synthesis step in AAs pathway. At rank 115, a cysteine protease-like transcript was reported sharing similarity with a C3 chain-shortening

enzyme from *V. planifolia,* which demonstrated an ability to convert 4-coumaric acid to 4-hydroxybenzaldehyde. High accumulation of crinamine suggested by UPLC-QTOF-MS analysis related to the occurrence of *vittatine 11-hydroxylase* among the most expressed genes which is responsible for converting vittatine to 11-hydroxyvittatine, eventually leading to the end product, crinamine. Through a local BLASTx analysis, we were able to identify full-length genes and their isoforms. For example, two *PAL* transcripts which shared more than 98% similar to *PAL* sequences from *Lycoris radiata* and *Ornithogalum longebracteatum,* were identified as *PAL1* and *PAL2*. Similarly, we also identified two full-length transcripts for *TYDC.*

Fragment per kilobase million (FPKM) results revealed that the genes from precursor pathway to norbelladine such as *PAL, TYDC, C3H, C4H* displayed similar expression levels which suggest a coordinated gene regulation to form intermediate precursors. The genes falling into the category of the highest FPKM values were *N4OMT3* and cytochrome P450s (*CYP96T1* and *CYP96T2*), suggesting their role in the central biosynthetic process of AAs which is specific to the bulb tissues. High expression of genes from central pathway of AAs are specific only to AAs pathway (*NBS* and *N4OMT)* and their expression correlates with highest AAs diversity in bulbs. High expression of *NpNBS* in bulb tissues also relates to the high accumulation of 4'-*O*-methylnorbelladine in bulbs.

In order to conduct a relative gene expression study, we performed RT-qPCR analysis which shows a differential expression pattern of all AA biosynthetic genes in bulbs, roots, stems, leaves, and flowers. For example, *PAL* isoforms, *PAL1* and *PAL2* in stem tissues displayed a similar expression level but in bulbs, *PAL2* expression was higher than *PAL1*, which suggests a different regulatory mechanism for each isoform. In *Lycoris radiata*, *Lr*PAL2 was also found more expressed than *Lr*PAL1 under MeJA. UV, gibberlic acid and salicylic acid treatment suggesting different role of these isoforms (Jiang, Xia et al. 2013). Particularly in bulbs, *N4OMT, TYDC,* and *PAL2* were most highly expressed transcripts. Flowers shows the least expression of all genes. Overall expression study shows differential gene expression profile where the lowest expression was detected in leaves and flowers.

Bulbs show high *N4OMT* and *TYDC* gene expression with highest alkaloid content, supporting gene-metabolite correlation hypothesis. Transcripts encoding enzymes that are proposed to be involved in AA precursor (tyramine and 3,4-DHBA) biosynthesis were identified along with the 4-hydroxybenzaldehyde (4-HBA) in metabolite profile. Expression of genes belonging to phenylpropanoid pathway (*e.g. PAL, C4H, C3H, 4CL*, and *HCT*) does not correspond to high alkaloid accumulation in stem suggests metabolic channeling of intermediate product to the enzymes of other isoquinoline pathway such as anthocyanins, lignins, flavonoids etc. Also, a high expression of *CYP96T* and *NorRed* was found in tissues with low alkaloid content. These contrasting results suggest that transport of AAs and enzymes within plant tissues, which was also previously reported for benzylisoquinoline alkaloids (Bird, Franceschi et al. 2003, Ziegler and Facchini 2008, Onoyovwe, Hagel et al. 2013). Thus, complete study provided us with a detailed knowledge of AAs and the expression pattern of their biosynthesis genes in *Narcissus pseudonarcissus*.

Further, we focused our study on a proposed condensation reaction between two substrates, tyramine and 3,4 DHBA via Mannich reaction leading to the formation of norbelladine. We observed that this reaction step of AAs metabolism share analogy with Norcoclaurine synthase (NCS) catalyzed step in the benzylisoquinoline alkaloid metabolic pathway. NCS is known to catalyze the first step in BIA metabolism leading to an intermediate formation, (S)-norcoclaurine. Studies of NCS characterization have been reported from *Thalictrum flavum* and *Papaver somniferum* (Samanani and Facchini 2002, Samanani, Liscombe et al. 2004). We also detected *NCS*-like transcript among the most abundantly expressed genes (FPKM value of 31543) in the study reported in chapter II. Therefore, we performed an *NCS* ortholog BLAST search in *N. pseudonarcissus* 'King Alfred' transcriptome. Out of three highly similar candidates, which shared homology with NCS from *Thalictrum flavum*, one candidate (*Np*NBS) with longest ORF cloned and expressed for functional characterization. The candidate selection was also based on its expression correlation with alkaloid accumulation in bulb. Through our enzyme assay

experiment, we discovered that *Np*NBS was responsible for the formation of norbelladine and catalyzing the first committed step of AAs biosynthesis pathway.

To study *Np*NBS function in relation to other proteins, we performed a motif scan and found a conserved Bet v1 domain among *Np*NBS, PR-10, and NCS sequences. This demonstrated that *Np*NBS belongs to a Bet v1/PR-10 protein family. Unlike *Cj*NCS1, *Np*NBS does not share similiarity with 2-oxoglutarate-dependent dioxygenases. *Np*NBS lacks 2-oxoglutarate-binding domain and Fe^{2+} binding ligands(Minami, Dubouzet et al. 2007). Instead, *Np*NBS possess a glycine rich p-loop motif corresponding to consensus sequence GXGGXGX, which is detected in NCS and PR-10/Bet V1 proteins from several species including *Ps*NCS, *Pb*NCS, *Tf*NCS, *Cj*PR10A, *Am*PR10A, *Le*PR10A, and *Ah*PR10A (Lee and Facchini 2010). Despite of a common reaction mechanism between NCS and strictosidine synthase, like *Tf*NCS, *Np*NBS doesn't share common features with strictisidine synthase enzyme. For example, STR1 sequence was reported with a common repetitive motifs comprising of five residues: three hydrophobic residues, followed by a small residue (Ser, Thr, Val or Gly) and a hydrophillic residue (Asp or Glu) at last. Such repetitive motifs were absent in *Np*NBS and related NCS sequences (member of PR-10/Bet v1). These finding suggests that similar reaction mechanism of NCS and STR might be a result of convergent evolution(Stöckigt, Barleben et al. 2008).

*Np*NBS lacks a signal peptide and is a cytosolic protein. Among several NCS, *Np*NBS was found closest to *Ps*NCS1 and *Ps*NCS2 from *Papaver somniferum* with 41% similarity. High expression of *NpNBS* gene in bulb demonstrated by qRT-PCR analysis was consistent with the expression of *N4OMT*, *TYDC* genes and also with the high diversity of AAs accumulation in the bulb, suggesting bulb as an important part of a plant for enzymes and AAs localization (Chapter II). In addition, studies conducted on *Np*N4OMT and *Tf*NCS also revealed enzyme localization specifically in underground tissues such as roots and rhizomes (Samanani, Liscombe et al. 2004). These results are supported by previous studies conducted specifically on NCS groups of Bet v1/PR-10, For example, VIGS mediated gene silencing of *PsNCS1* and *PsNCS2* from *Papaver somniferum* resulted in a significant decrease of BIAs (Lee and Facchini 2010). *Np*NBS

activity in the absence of a signal peptide is in agreement with NCS activity study from NCS from *Thalictrum flavum*, which reports that the truncation of signal peptide contributes to the enzymatic activity an enzyme (Samanani, Liscombe et al. 2004, Nishihachijo, Hirai et al. 2014).

For protein expression, pET301/CT-DEST expression vector harboring *Np*NBS gene was transformed into Rosetta (DE3) pLys host strain carrying a chromosomal copy of a T7 RNA polymerase which is required for the transcription of *NpNBS* gene under the control of T7 promoter in the expression vector. Therefore, the strain is suitable for the production of protein from target gene cloned under pET vectors. It allows the expression of eukaryotic proteins by supplying codons that are rare in the *E.coli*.

A previous study conducted on *NR* gene has suggested norcraugsodine as one of the intermediate upstream norbelladine in AAs biosynthesis pathway(Kilgore, Holland et al. 2016). Therefore, it was necessary to identify the product of NBS through LC-MS/MS using norbelladine and norcraugsodine as standards. We determined *Np*NBS assay product using qualifier and quantifier ion selected for each standard. Enzyme assay performed with non-induced and heat-denatured *Np*NBS were used as negative controls. As expected, no product was detected in non-induce sample by SDS-PAGE analysis because no IPTG was added to induce expression. This control also indicates stable expression system in Rosetta (DE3) pLys host strain of *E.coli* and only upon the addition of IPTG *Np*NBS product was observed. We also added boiled *Np*NBS enzyme assay as a negative control to ensure that the product obtained was due to an active enzyme, which was not synthesized by a heat-denatured protein. However, a weak signal of background reaction was observed in control (with buffer and substrates), suggesting a non-enzymatic chemical reaction occurring between the the two substrates. At last, through LC-MS/MS analysis, we confirmed that *Np*NBS was capable of forming norbelladine by quantification of m/z 121. A qualifier ion (m/z 138) specific to norbelladine was detected and norcraugsodine qualifier ion (m/z 157) was absent in the assay product. The *Np*NBS enzyme showed higher activity at pH 4 compared to pH 7, suggesting that enzyme favors reaction at acidic pH. We also identified conserved catalytic residues Tyr^{68}, Glu^{71} and

Lys[83] in *Np*NBS responsible for its catalytic activity. However, absence of Asp[141] residue doesn't affect enzymes activity suggesting that Asp[141] residue is not responsible for base catalysis and might be involved in imparting electrostatic stabilization (Lichman, Gershater et al. 2015).

4.2 Future Perspectives

Advancement in omics technologies, their reduced cost of processing and integrated analysis, have accelerated the study of phytochemicals from non-model plants. For this project, we generated a transcriptome profile of *N. pseudonarcissus* 'King Alfred' as described in chapter two, that could be utilized for further investigation of new genes, gene isoforms, and transcription factors. A homology search for a target gene together with co-expressiion analysis of a corresponding biosynthesis pathway in the transcriptome will be a good strategy to identify unknown genes. Recently, *Noroxomaritidine reductase (NR)* gene was characterized using similar approach and was found responsible for catalyzing conversion of noroxomaritdine to oxomaritidine (Kilgore, Holland et al. 2016). Unknown enzymes of the AAs biosynthetic pathway leading to the end products such as haemanthamine, lycorenine and narwedine could also be targeted through a similar approach which will help to fill the gaps of a complex biosynthetic pathway. Moreover, metabolic engineering using identified genes will be possible by developing a short synthetic pathway for specific AA such as galanthamine or oxomaritinamine. For example, using 4'-*O* methylnorbelladine as a precursor for galanthamine production in a heterologous host, the discovery of only two enzymes is required: one enzyme catalyzing conversion of *N*-demethylnarwedine to norgalanthamine and second converting norgalanthamine to galanthamine. In the case of oxomaritinamine synthesis, all genes are characterized from intermediate 3,4-DHBA till oxomaritinamine. Therefore, metabolic engineering could be performed by constructing enzymatic pathway using *NBS* gene along with recently characterized enzymes (*N4OMT, CYP96T1* and *NR*) with 3,4-dihydroxybenzaldehyde and tyramine as precursors (Figure 1.13). Also, confirmation of two or three genes will allow the production of 11-hydroxyvittatine and crinamine from 4'-O-methylnorbelladine via metabolic engineering. Recently, 7-gene pathway *(SAS,*

CPR, SAR, SAT, CODM, T6ODM and COR) was reconstructed for the production of morphine and codeine in *S. cerevisiae* (Fossati, Narcross et al. 2015). Another study demonstrated that reconconsitution of 10-BIA biosynthesis genes in *S. cerevisiae* enabled the production of dihydrosanguinarine and sanguinarine (an oxide derivative of dihydrosanguinarine) from (R,S)-norlaudanosoline, with a product yield ranging from 37 % to 57%. This assembly of 10 genes was the first longest and complex reconstituted pathway ever reported in yeast and represents the potential of microbial engineering (Fossati, Ekins et al. 2014).

In addition, gene silencing will help to study the role of the discovered gene using VIGS, TALEN (transcription activator-like effector nucleases) and CRISPER/CAS9 systems. Moreover, overexpression of a known gene for its functional analysis using a transient expression system will be useful for developing transgenic plants or cell cultures for high alkaloid content analysis. VIGS has been used extensively in functional analysis of several BIA pathway genes including suppression of *4'OMT* and *7'OMT* genes, which resulted in reduction of total BIA content in stem of *Papaver somniferum*. On contrary, *4'OMT* and *7'OMT* genes overexpression was performed through agro-infiltration of LBA44 strain of *Agrobacterium tumefaciens* in stem, leaf and capsule tissues. The strain contains p-GEM-T vector harbouring 4'OMT and 7'OMT genes under 35S promoter. This resulted in high morphine concentration in stem but lesser amount in capsules (Gurkok, Ozhuner et al. 2016). Silencing of *Tf*NCS via VIGS resulted in approximately 75% reduced *Tf*NCS transcript expression and significant decrease in the accumulation of morphine, codeine, thebaine, noscapine and papaverine (Lee and Facchini 2010). VIGS has also been used for the suppression of six genes to study their putative role in papaverine biosynthesis. As a result, silencing of gene encoding *coclaurine N-methyltransferase* significantly increased papverine level while silencing of *(S)-3'-hydroxy-N-methylcoclaurine 4-O-methyltransferase* and *norreticuline 7-O-methyltransferase* reduced papverine concentration. Silencing of *6-O-methyltransferase* transcript resulted in reduction of total alkaloid content indicating the role of (S)-norcoclaurine as a key branch-point intermediate (Desgagné-Penix and Facchini 2012). VIGS mediated gene silencing of *CYP82Y1* gene decreased noscapine, narcotoline, and a

putative downstream secoberbine intermediate level notably and increased accumulation of upstream pathway intermediated such as scoulerine, tetrahydrocolum-bamine, canadine, and *N*-methylcanadine (Dang and Facchini 2014). Though CRISPR/Cas system has been successfully implanted on several eukaorytic organism but it has been reported only few times in plants such as *Arabidopsis thaliana*. Mutagenesis was obtained for seven targeted genes using CRISPR/Cas9 demonstrating the efficiency of the tool (Li, Norville et al. 2013). It has also shown an efficient targeted mutagenesis of *chlorophyll A oxygenase 1* (CAO 1) in *Oryza sativa* (Miao, Guo et al. 2013).

The results of UPLC-QTOF-MS analysis shows that some AAs were listed twice or thrice with slight different retention times, which demands a more populated database to differentiate between the structural isomers of these alkaloids. For example, AAs, crinamine is listed with different retention times (6.07 min and 6.12 min) in the UPLC-QTOF-MS data (chapter II, table II), which is possibly haemanthamine, an epimer of crinamine. Therefore, a further analysis of such AAs using chiral chromatography coupled to mass spectrometry could be performed to resolve the identity of structural isomers in order to obtain a consolidated knowledge on AAs. However, no such study has been reported so far but a chiral column with LC-MS analysis was used for confirming enzyme assay product obtained from a sterioselective native *Cj*NCS enzyme, which produce *(S)*-norcoclaurine. A chiral column was used to separate *(S)* and *(R)* forms of coclaurine standard and confirmed the production of *(S)*-coclaurine(Minami, Dubouzet et al. 2007).

Besides UPLC-MS based metabolomics, NMR combined with mass spectroscopy (MS) is a novel technique that could be used in future for identification of unknown metabolites. This technique first identifies molecular formula via MS and then produces a chemical structure in accordance to a molecular formula obtained through NMR (Bingol, Bruschweiler-Li et al. 2015).

To understand the molecular basis behind the condensation reaction mechanism between tyramine and 3,4-dihydroxybenzaldehyde (3,4-DHBA) catalyzed by norbelladine synthase and to study their binding, it will be interesting to determine the

crystal structure of *Np*NBS by X-ray crystallography. This will also provide information about substrates binding sites, enzyme's folding pattern through conformational study of α-helices, β-sheets and proteins subunits (monomer, dimer, trimer or tetramer). A crystallographic structure of *Tf*NCS shows that enzyme is a tetramer. Enzyme's secondary structure consists of a seven stranded anti-parallel β-sheet folded around C-terminal helix (α-3) and two small α-helices (α1 and α2). The monomer shows an extended tunnel between β-sheets and three α-helices. Each monomer conforms to the overall fold corresponding to the Betv1-like superfamily. However, *Tf*NCS has longer C-terminal helix than Bet-v1. NCS active sites consisted of four residues (Lys^{122}, Asp^{141}, Glu^{110} and Tyr^{108}) (Pasquo, Bonamore et al. 2008). In addition, X-ray crystallography has been performed for an AAs biosynthesis enzyme, Noroxomaritidine reductase (NR) which revealed that it reduces C-C double bond in noroxomaritidine to form oxomaritinamine and its binding to piperonal (3, 4-DHBA analogue) and tyramine. Through this analysis, NR was found to be a tetramer and each of its monomer comprises of 271 amino acids arranged in alternating α/β structure formed by seven parrallel β-sheets compressed between two layers of three α-helices (Kilgore, Holland et al. 2016).

As we have performed *Np*NBS transcript expression in different tissues of *N.pseudonarcissus* 'King Alfred' in this study, further investigation about *Np*NBS protein distribution in the plant will add to *Np*NBS protein information. For example, a comparative study of relative abundance of *NCS*2 transcript and protein in different plant tissues was performed by RNA gel blot hybridization using *NCS2* cDNA from *Papaver somniferum* and immunoblot analysis using polyclonal antibody against *Ps*NCS. *PsNCS* transcripts and protein was abundant in roots and stems in comparision to leaf and carpel. Its abundance in stem and roots was consistent with distribution of other BIA producing enzyme incating BIA metabolism occurs mainly in these parts (Lee and Facchini 2010).

The cellular localization of *Np*NBS could be determined through immunocytochemical staining. This will help to detect the cell-specific occurance of *Np*NBS in different tissues of *N. pseudonarcissus* 'King Alfred'. A similar study has been performed previously for *Tf*NCS, which revealed its localization in the sieve elements of

the phloem present adjacent to the laticifers in the vascular bundles of stems, root, leaf and carpel. The localization of *Tf*NCS was consistent with *6OMT* location in plant (Lee and Facchini 2010).

4.3 Final conclusions

Altogether, this research introduces OMICS tools guided workflow for identification and characterization of new genes and enzymes involved in the metabolism of non-model plants. Our research led to the identification of all proposed genes of the pathways including their full-length transcripts, which is an important information for various biotechnological applications. Additionally, our work shows the molecular characterization of a novel gene, *Np*NBS responsible for catalyzing the first committed step in AAs biosynthesis pathway. Discovery of *Np*NBS along with the previously discovered AA biosynthetic enzymes has opened the door for high-scale production of medicinal AAs via synthetic biology.

REFERENCES

Achakzai, A. K. K., P. Achakzai, A. Masood, S. A. Kayani and R. B. Tareen (2009). "Response of plant parts and age on the distribution of secondary metabolites on plants found in Quetta." Pakistan journal of Botany **41**(5): 2129-2135.

Achakzai, A. K. K., P. Achakzai, A. Masood, S. A. Kayani and R. B. Tareen (2009). "Response of plant parts and age on the distribution of secondary metabolites on plants found in Quetta." Pak. J. Bot **41**(5): 2129-2135.

Akhtar, S., J. Ahmad and A. Ahmad (2017). Molecular Network of Monoterpene Indole Alkaloids (MIAs) Signaling in Plants with Reference to Catharanthus roseus (L.) G. Don. Stress Signaling in Plants: Genomics and Proteomics Perspective, Volume 2, Springer**:** 37-67.

Akula, R. and G. A. Ravishankar (2011). "Influence of abiotic stress signals on secondary metabolites in plants." Plant signaling & behavior **6**(11): 1720-1731.

Alcantara, J., D. A. Bird, V. R. Franceschi and P. J. Facchini (2005). "Sanguinarine biosynthesis is associated with the endoplasmic reticulum in cultured opium poppy cells after elicitor treatment." Plant Physiology **138**(1): 173-183.

Aniszewski, T. (1994). "From iodine to enzyme: a critical review of chemical and biological methods of lupine alkaloid analysis." The science of legumes: an annual international scientific journal (Finland).

Aniszewski, T. (2015). Alkaloids: Chemistry, Biology, Ecology, and Applications, Elsevier.

Ashihara, H. and H. Kubota (1987). "Biosynthesis of purine alkaloids in Camellia plants." Plant and cell physiology **28**(3): 535-539.

Ashihara, H., H. Sano and A. Crozier (2008). "Caffeine and related purine alkaloids: biosynthesis, catabolism, function and genetic engineering." Phytochemistry **69**(4): 841-856.

Ashihara, H. and T. Suzuki (2004). "Distribution and biosynthesis of caffeine in plants." Front Biosci **9**(2): 1864-1876.

Barik, J., F. Dajas-Bailador and S. Wonnacott (2005). "Cellular responses to nicotinic receptor activation are decreased after prolonged exposure to galantamine in human neuroblastoma cells." Br J Pharmacol **145**(8): 1084-1092.

Barton, D. H. R., G. W. Kirby, G. B. Taylor and G. M. Thomas (1963). "Phenol oxidation and biosynthesis. Part VI. The biogenesis of Amaryllidaceae alkaloids." J. Chem. Soc.: 4545-4558.

Bastida Armengol, J., S. Berkov, L. Torras Claveria, N. B. Pigni, J. P. d. Andradre, V. Martínez, C. Codina Mahrer and F. Viladomat Meya (2011). "Chemical and biological aspects of Amaryllidaceae alkaloids." Recent Advances in Pharmaceutical Sciences, 2011, Chapter 3, p. 65-100. Editor: Diego Muñoz-Torrero.

Bastida, J., S. Berkov, L. Torras, N. B. Pigni, J. P. de Andrade, V. Vanessa Martínez, C. Codina and F. Viladomat (2011). Chapter 3: Chemical and biological aspects of Amaryllidaceae alkaloids. E. D. Muñoz-Torrero, T. R. Network, F. P. O. 37/661 (2), T.-. 023 and I. Kerala, Recent Advances in Pharmaceutical Sciences.

Bate-Smith, E. (1972). "Detection and determination of ellagitannins." Phytochemistry **11**(3): 1153-1156.

Battersby, A. R., H. M. Fales and W. C. Wildman (1961). "Biosynthesis in the Amaryllidaceae. Tyrosine and norbelladine as precursors of haemanthamine." J. Amer. Chem. Soc. **83**: 4098-4099.

Beaudoin, G. A. and P. J. Facchini (2014). "Benzylisoquinoline alkaloid biosynthesis in opium poppy." Planta **240**(1): 19-32.

Ben Amar, A., S. Daldoul, G. M. Reustle, G. Krczal and A. Mliki (2016). Reverse Genetics and High Throughput Sequencing Methodologies for Plant Functional Genomics.

Benowitz, N. L. (1990). "Clinical pharmacology of caffeine." Annual review of medicine **41**(1): 277-288.

Berkov, S., J. Bastida, F. Viladomat and C. Codina (2011). "Development and validation of a GC-MS method for rapid determination of galanthamine in Leucojum aestivum and Narcissus ssp.: a metabolomic approach." Talanta **83**(5): 1455-1465.

Berkov, S., V. Martínez-Francés, J. Bastida, C. Codina and S. Ríos (2014). "Evolution of alkaloid biosynthesis in the genus Narcissus." Phytochemistry **99**: 95-106.

Bernays, E. (1983). "Antifeedants in crop pest management." Natural Products for Innovative Pest Management. Pergamon Press, Oxford: 259-271.

Bhattacharya, A., P. Sood and V. Citovsky (2010). "The roles of plant phenolics in defence and communication during Agrobacterium and Rhizobium infection." Molecular plant pathology **11**(5): 705-719.

Bidlack, W. R., S. T. Omaye, M. S. Meskin and D. K. Topham (2000). Phytochemicals as bioactive agents, CRC press.

Bingol, K., L. Bruschweiler-Li, C. Yu, A. Somogyi, F. Zhang and R. Brüschweiler (2015). "Metabolomics beyond spectroscopic databases: a combined MS/NMR strategy for the rapid identification of new metabolites in complex mixtures." Analytical chemistry **87**(7): 3864-3870.

Bird, D. A., V. R. Franceschi and P. J. Facchini (2003). "A tale of three cell types: alkaloid biosynthesis is localized to sieve elements in opium poppy." The Plant Cell **15**(11): 2626-2635.

Bird, D. A., V. R. Franceschi and P. J. Facchini (2003). "A tale of three cell types: alkaloid biosynthesis is localized to sieve elements in opium poppy." Plant Cell **15**(11): 2626-2635.

Bolger, A. M., M. Lohse and B. Usadel (2014). "Trimmomatic: a flexible trimmer for Illumina sequence data." Bioinformatics **30**(15): 2114-2120.

Boschin, G. and D. Resta (2013). Alkaloids derived from lysine: quinolizidine (a focus on lupin alkaloids). Natural Products, Springer**:** 381-403.

Bourgaud, F., A. Gravot, S. Milesi and E. Gontier (2001). "Production of plant secondary metabolites: a historical perspective." Plant science **161**(5): 839-851.

Brown, C. T., A. Howe, Q. Zhang, A. B. Pyrkosz and T. H. Brom (2012). "A reference-free algorithm for computational normalization of shotgun sequencing data." arXiv preprint arXiv:1203.4802.

Brunt, A. (1996). "Plant Viruses Online: Descriptions and Lists from the VIDE Database. Ver. 20^< th." http://biology. anu. edu. au/Groups/MES/vide/.

Brunt, A. A., K. Crabtree, M. J. Dallwitz, A. J. Gibbs, L. Watson and E. J. Zurcher (1996). "Plant Viruses Online: Descriptions and Lists from the VIDE Database. Version: 16th January 1997." URL http://biology.anu.edu.au/Groups/MES/vide/.

Calabrò, S. (2015). Plant secondary metabolites. Rumen Microbiology: From Evolution to Revolution, Springer**:** 153-159.

Cázares-Flores, P., D. Levac and V. De Luca (2016). "Rauvolfia serpentina N-methyltransferases involved in ajmaline and Nβ-methylajmaline biosynthesis belong to a gene family derived from γ-tocopherol C-methyltransferase." The Plant Journal **87**(4): 335-342.

Cedrón, J. C., M. Del Arco-Aguilar, A. Estévez-Braun and Á. G. Ravelo (2010). Chemistry and biology of Pancratium alkaloids. The Alkaloids: Chemistry and Biology, Elsevier. **68:** 1-37.

Chang, L., J. M. Hagel and P. J. Facchini (2015). "Isolation and characterization of O-methyltransferases involved in the biosynthesis of glaucine in Glaucium flavum." Plant physiology **169**(2): 1127-1140.

Chase, M. W., J. L. Reveal and M. F. Fay (2009). "A subfamilial classification for the expanded asparagalean families Amaryllidaceae, Asparagaceae and Xanthorrhoeaceae." Botanical Journal of the Linnean Society **161**(2): 132-136.

Cheng, A. X., Y. G. Lou, Y. B. Mao, S. Lu, L. J. Wang and X. Y. Chen (2007). "Plant terpenoids: biosynthesis and ecological functions." Journal of Integrative Plant Biology **49**(2): 179-186.

Cimmino, A., M. Masi, M. Evidente, S. Superchi and A. Evidente (2017). "Amaryllidaceae alkaloids: Absolute configuration and biological activity." Chirality.
Cordell, G. A. (2008). The alkaloids: chemistry and biology, Academic Press.

Cragg, G. M. and D. J. Newman (2005). "Biodiversity: A continuing source of novel drug leads." Pure and applied chemistry **77**(1): 7-24.

Croteau, R., T. M. Kutchan and N. G. Lewis (2000). "Natural products (secondary metabolites)." Biochemistry and molecular biology of plants **24**: 1250-1319.

Dakora, F. and D. Phillips (1996). "Diverse functions of isoflavonoids in legumes transcend anti-microbial definitions of phytoalexins." Physiological and Molecular Plant Pathology **49**(1): 1-20.

Dang, T.-T. T. and P. J. Facchini (2014). "CYP82Y1 is N-methylcanadine 1-hydroxylase, a key noscapine biosynthetic enzyme in opium poppy." Journal of Biological Chemistry **289**(4): 2013-2026.

de Andrade, J. P., N. B. Pigni, L. Torras-Claveria, S. Berkov, C. Codina, F. Viladomat and J. Bastida (2012). "Bioactive alkaloid extracts from Narcissus broussonetii: Mass spectral studies." Journal of pharmaceutical and biomedical analysis **70**: 13-25.

de Andrade, J. P., N. B. Pigni, L. Torras-Claveria, Y. Guo, S. Berkov, R. Reyes-Chilpa, A. el Amrani, J. A. S. Zuanazzi, C. Codina and F. Viladomat (2012). "Alkaloids from the Hippeastrum genus: chemistry and biological activity." Revista latinoamericana de química **40**(2): 83-98.

De Felice, B., F. Manfellotto, R. D'Alessandro, O. De Castro, A. Di Maio and M. Trifuoggi (2013). "Comparative transcriptional analysis reveals differential gene expression between Sand Daffodil tissues." Genetica **141**(10-12): 443-452.

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

Delano-Frier, J. P., H. Aviles-Arnaut, K. Casarrubias-Castillo, G. Casiqu
e-Arroyo, P. A. Castrillon-Arbelaez, L. Herrera-Estrella, J. Massange-Sanchez, N. A. Martinez-Gallardo, F. I. Parra-Cota, E. Vargas-Ortiz and M. G. Estrada-Hernandez (2011). "Transcriptomic analysis of grain amaranth (Amaranthus hypochondriacus) using

454 pyrosequencing: comparison with A. tuberculatus, expression profiling in stems and in response to biotic and abiotic stress." BMC Genomics **12**: 363.

Der Marderosian, A. and J. Beutler (2002). Oregano, in" The Review of Natural Products".

DerMarderosian, A. and J. A. Beutler (2002). The review of natural products: the most complete source of natural product information, Facts and Comparisons.

Desgagné-Penix, I. and P. J. Facchini (2011). Benzylisoquinoline Alkaloid Biosynthesis. Plant Metabolism and Biotechnology, John Wiley & Sons, Ltd: 241-261.

Desgagne-Penix, I., S. C. Farrow, D. Cram, J. Nowak and P. J. Facchini (2012). "Integration of deep transcript and targeted metabolite profiles for eight cultivars of opium poppy." Plant Mol Biol **79**(3): 295-313.

Desgagné-Penix, I., S. C. Farrow, D. Cram, J. Nowak and P. J. Facchini (2012). "Integration of deep transcript and targeted metabolite profiles for eight cultivars of opium poppy." Plant molecular biology **79**(3): 295-313.

Desgagné-Penix, I., M. F. Khan, D. C. Schriemer, D. Cram, J. Nowak and P. J. Facchini (2010). "Integration of deep transcriptome and proteome analyses reveals the components of alkaloid metabolism in opium poppy cell cultures." BMC Plant Biology **10**(1): 252.

Desgagné-Penix, I. and P. J. Facchini (2011). "Benzylisoquinoline alkaloid biosynthesis." Plant Metabolism and Biotechnology: 241-261.

Desgagné-Penix, I. and P. J. Facchini (2012). "Systematic silencing of benzylisoquinoline alkaloid biosynthetic genes reveals the major route to papaverine in opium poppy." The Plant Journal **72**(2): 331-344.

Dewick, P. M. (2002). Medicinal natural products: a biosynthetic approach, John Wiley & Sons.

Dewick, P. M. (2009). Medicinal natural products: a biosynthetic approach., John Wiley and Sons Ltd.

Dewick, P. M. (2009). "The shikimate pathway: aromatic amino acids and phenylpropanoids." Medicinal Natural Products: A Biosynthetic Approach, 3rd Edition: 137-186.

Diamond, A. and I. Desgagne-Penix (2016). "Metabolic engineering for the production of plant isoquinoline alkaloids." Plant Biotechnol J **14**(6): 1319-1328.

Diamond, A. and I. Desgagné-Penix (2016). "Metabolic engineering for the production of plant isoquinoline alkaloids." Plant biotechnology journal **14**(6): 1319-1328.

Dias, D. A., S. Urban and U. Roessner (2012). "A historical overview of natural products in drug discovery." Metabolites **2**(2): 303-336.

Doskočil, I., A. Hošťálková, M. Šafratová, N. Benešová, J. Havlík, R. Havelek, J. Kuneš, K. Královec, J. Chlebek and L. Cahlíková (2015). "Cytotoxic activities of Amaryllidaceae alkaloids against gastrointestinal cancer cells." Phytochemistry Letters **13**: 394-398.

Dudareva, N., E. Pichersky and J. Gershenzon (2004). "Biochemistry of plant volatiles." Plant physiology **135**(4): 1893-1902.

Dumont, P., L. Ingrassia, S. Rouzeau, F. Ribaucour, S. Thomas, I. Roland, F. Darro, F. Lefranc and R. Kiss (2007). "The Amaryllidaceae isocarbostyril narciclasine induces apoptosis by activation of the death receptor and/or mitochondrial pathways in cancer cells but not in normal fibroblasts." Neoplasia **9**(9): 766-776.

Dumont, P., L. Ingrassia, S. Rouzeau, F. Ribaucour, S. Thomas, I. Roland, F. Darro, F. Lefranc and R. Kiss (2007). "The Amaryllidaceae isocarbostyril narciclasine induces apoptosis by activation of the death receptor and/or mitochondrial pathways in cancer cells but not in normal fibroblasts." Neoplasia **9**(9): 766-776.

Eichhorn, J., T. Takada, Y. Kita and M. H. Zenk (1998). "Biosynthesis of the Amaryllidaceae alkaloid galanthamine." Phytochemistry **49**(4): 1037-1047.

Evidente, A., A. Andolfi, A. H. Abou-Donia, S. M. Touema, H. M. Hammoda, E. Shawky and A. Motta (2004). "(−)-Amarbellisine, a lycorine-type alkaloid from Amaryllis belladonna L. growing in Egypt." Phytochemistry **65**(14): 2113-2118.

Evidente, A. and A. Kornienko (2009). "Anticancer evaluation of structurally diverse Amaryllidaceae alkaloids and their synthetic derivatives." Phytochemistry Reviews **8**(2): 449.

Facchini, P. J. (2001). "Alkaloid biosynthesis in plants: biochemistry, cell biology, molecular regulation, and metabolic engineering applications." Annual review of plant biology **52**(1): 29-66.

Facchini, P. J. and V. De Luca (1994). "Differential and tissue-specific expression of a gene family for tyrosine/dopa decarboxylase in opium poppy." Journal of Biological Chemistry **269**(43): 26684-26690.

Facchini, P. J., K. L. Huber-Allanach and L. W. Tari (2000). "Plant aromatic L-amino acid decarboxylases: evolution, biochemistry, regulation, and metabolic engineering applications." Phytochemistry **54**(2): 121-138.

Facchini, P. J., K. L. Huber-Allanach and L. W. Tari (2000). "Plant aromatic L-amino acid decarboxylases: evolution, biochemistry, regulation, and metabolic engineering applications." Phytochemistry **54**(2): 121-138.

Fang, X., C.-Q. Yang, Y.-K. Wei, Q.-X. Ma, L. Yang and X.-Y. Chen (2011). "Genomics grand for diversified plant secondary metabolites." Plant Div Res **33**(1): 53-64.

Farrow, S. C., J. M. Hagel and P. J. Facchini (2012). "Transcript and metabolite profiling in cell cultures of 18 plant species that produce benzylisoquinoline alkaloids." Phytochemistry **77**: 79-88.

Fattorusso, E. and O. Taglialatela-Scafati (2008). Modern alkaloids: structure, isolation, synthesis, and biology, John Wiley & Sons.

Fennell, C. and J. Van Staden (2001). "Crinum species in traditional and modern medicine." Journal of Ethnopharmacology **78**(1): 15-26.

Figueiredo, A. C., J. G. Barroso, L. G. Pedro and J. J. Scheffer (2008). "Factors affecting secondary metabolite production in plants: volatile components and essential oils." Flavour and Fragrance journal **23**(4): 213-226.

Fossati, E., A. Ekins, L. Narcross, Y. Zhu, J.-P. Falgueyret, G. A. Beaudoin, P. J. Facchini and V. J. Martin (2014). "Reconstitution of a 10-gene pathway for synthesis of the plant alkaloid dihydrosanguinarine in Saccharomyces cerevisiae." Nature communications **5**: 3283.

Fossati, E., L. Narcross, A. Ekins, J.-P. Falgueyret and V. J. Martin (2015). "Synthesis of morphinan alkaloids in Saccharomyces cerevisiae." PLoS One **10**(4): e0124459.

Frick, K. M., L. G. Kamphuis, K. H. Siddique, K. B. Singh and R. C. Foley (2017). "Quinolizidine alkaloid biosynthesis in lupins and prospects for grain quality improvement." Frontiers in plant science **8**.

Fürstenberg-Hägg, J., M. Zagrobelny and S. Bak (2013). "Plant Defense against Insect Herbivores." International Journal of Molecular Sciences **14**(5): 10242-10297.

Gabrielsen, B., T. P. Monath, J. W. Huggins, D. F. Kefauver, G. R. Pettit, G. Groszek, M. Hollingshead, J. J. Kirsi, W. M. Shannon and E. M. Schubert (1992). "Antiviral (RNA) activity of selected Amaryllidaceae isoquinoline constituents and synthesis of related substances." Journal of Natural Products **55**(11): 1569-1581.

Gahlan, P., H. R. Singh, R. Shankar, N. Sharma, A. Kumari, V. Chawla, P. S. Ahuja and S. Kumar (2012). "De novo sequencing and characterization of Picrorhiza kurrooa transcriptome at two temperatures showed major transcriptome adjustments." BMC genomics **13**(1): 126.

Gallage, N. J., E. H. Hansen, R. Kannangara, C. E. Olsen, M. S. Motawia, K. Jorgensen, I. Holme, K. Hebelstrup, M. Grisoni and B. L. Moller (2014). "Vanillin formation from ferulic acid in Vanilla planifolia is catalysed by a single enzyme." Nat Commun **5**: 4037.

Gangwar, M., M. K. Gautam, A. K. Sharma, Y. B. Tripathi, R. Goel and G. Nath (2014). "Antioxidant capacity and radical scavenging effect of polyphenol rich Mallotus philippenensis fruit extract on human erythrocytes: an in vitro study." The Scientific World Journal **2014**.

Ge, G., P. Xiao, Y. Zhang and L. Yang (2011). "The first insight into the tissue specific taxus transcriptome via Illumina second generation sequencing." PLoS one **6**(6): e21220.

Ghosal, S., A. Shanthy and S. K. Singh (1988). "Isocraugsodine, an n-arylidenephenethylamine from Crinum asiaticum and its ez isomerism." Phytochemistry **27**(6): 1849-1852.

Gomez, D., J. Azorin, J. Bastida, F. Viladomat and C. Codina (2003). "Seasonal and spatial variations of alkaloids in Merendera montana in relation to chemical defense and phenology." J Chem Ecol **29**(5): 1117-1126.

Gómez, D., J. Azorín, J. Bastida, F. Viladomat and C. Codina (2003). "Seasonal and spatial variations of alkaloids in Merendera montana in relation to chemical defense and phenology." Journal of chemical ecology **29**(5): 1117-1126.

Gouvea, D. R., L. Gobbo-Neto, H. T. Sakamoto, N. P. Lopes, J. L. C. Lopes, F. Meloni and J. G. Amaral (2012). "Seasonal variation of the major secondary metabolites present in the extract of Eremanthus mattogrossensis Less (Asteraceae: Vernonieae) leaves." Química Nova **35**(11): 2139-2145.

Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B. W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman and A. Regev (2011). "Full-length transcriptome assembly from RNA-Seq data without a reference genome." Nat Biotechnol **29**(7): 644-652.

Griffin, W. J. and G. D. Lin (2000). "Chemotaxonomy and geographical distribution of tropane alkaloids." Phytochemistry **53**(6): 623-637.

Grycová, L., J. Dostál and R. Marek (2007). "Quaternary protoberberine alkaloids." Phytochemistry **68**(2): 150-175.

Guillon, S., J. Trémouillaux-Guiller, P. K. Pati, M. Rideau and P. Gantet (2006). "Hairy root research: recent scenario and exciting prospects." Current opinion in plant biology **9**(3): 341-346.

Guo, G., G. Yao, G. Zhan, Y. Hu, M. Yue, L. Cheng, Y. Liu, Q. Ye, G. Qing and Y. Zhang (2014). "N-methylhemeanthidine chloride, a novel Amaryllidaceae alkaloid, inhibits pancreatic cancer cell proliferation via down-regulating AKT activation." Toxicology and applied pharmacology **280**(3): 475-483.

Gurkok, T., E. Ozhuner, I. Parmaksiz, S. Özcan, M. Turktas, A. Ipek, I. Demirtas, S. Okay and T. Unver (2016). "Functional Characterization of 4′ OMT and 7OMT Genes in BIA Biosynthesis." Frontiers in plant science **7**: 98.

Hagel, J. M. and P. J. Facchini (2005). "Elevated tyrosine decarboxylase and tyramine hydroxycinnamoyltransferase levels increase wound-induced tyramine-derived hydroxycinnamic acid amide accumulation in transgenic tobacco leaves." Planta **221**(6): 904-914.

Hagel, J. M., R. Mandal, B. Han, J. Han, D. R. Dinsmore, C. H. Borchers, D. S. Wishart and P. J. Facchini (2015). "Metabolome analysis of 20 taxonomically related benzylisoquinoline alkaloid-producing plants." BMC plant biology **15**(1): 220.

Hagel, J. M., R. Mandal, B. Han, J. Han, D. R. Dinsmore, C. H. Borchers, D. S. Wishart and P. J. Facchini (2015). "Metabolome analysis of 20 taxonomically related benzylisoquinoline alkaloid-producing plants." BMC Plant Biology **15**(1): 1-17.

Hagel, J. M., J. S. Morris, E. J. Lee, I. Desgagné-Penix, C. D. Bross, L. Chang, X. Chen, S. C. Farrow, Y. Zhang, J. Soh, S. C.W. and P. J. Facchini (2015). "Transcriptome analysis of 20 taxonomically related benzylisoquinoline alkaloid-producing plants." BMC Plant Biology **15**.

Han, R., A. Rai, M. Nakamura, H. Suzuki, H. Takahashi, M. Yamazaki and K. Saito (2016). "Chapter Two-De Novo Deep Transcriptome Analysis of Medicinal Plants for Gene Discovery in Biosynthesis of Plant Natural Products." Methods in enzymology **576**: 19-45.

Hao, D., P. Ma, J. Mu, S. Chen, P. Xiao, Y. Peng, L. Huo, L. Xu and C. Sun (2012). "De novo characterization of the root transcriptome of a traditional Chinese medicinal plant Polygonum cuspidatum." Science China Life Sciences **55**(5): 452-466.

Harken, R., C. Christensen and W. Wildman (1976). "Interconversions in the pluviine-lycorenine series." The Journal of Organic Chemistry **41**(14): 2450-2454.

Hartmann, T. (2004). "Plant-derived secondary metabolites as defensive chemicals in herbivorous insects: a case study in chemical ecology." Planta **219**(1): 1-4.

Hartmann, T., A. Ehmke, U. Eilert, K. von Borstel and C. Theuring (1989). "Sites of synthesis, translocation and accumulation of pyrrolizidine alkaloid N-oxides in Senecio vulgaris L." Planta **177**(1): 98-107.

Hartmann, T., H. Sander, R. Adolph and G. Toppel (1988). "Metabolic links between the biosynthesis of pyrrolizidine alkaloids and polyamines in root cultures of Senecio vulgaris." Planta **175**(1): 82-90.

Hartmann, T. and M. Zimmer (1986). "Organ-specific distribution and accumulation of pyrrolizidine alkaloids during the life history of two annual Senecio species." Journal of plant physiology **122**(1): 67-80.

Havelek, R., M. Seifrtova, K. Kralovec, L. Bruckova, L. Cahlikova, M. Dalecka, J. Vavrova, M. Rezacova, L. Opletal and Z. Bilkova (2014). "The effect of Amaryllidaceae alkaloids haemanthamine and haemanthidine on cell cycle progression and apoptosis in p53-negative human leukemic Jurkat cells." Phytomedicine **21**(4): 479-490.

Hawkins, K. M. and C. D. Smolke (2008). "Production of benzylisoquinoline alkaloids in Saccharomyces cerevisiae." Nature Chemical Biology **4**: 564.

He, M., C. Qu, O. Gao, X. Hu and X. Hong (2015). "Biological and pharmacological activities of amaryllidaceae alkaloids." RSC Advances **5**(21): 16562-16574.

Heinrich, M. and H. Lee Teoh (2004). "Galanthamine from snowdrop--the development of a modern drug against Alzheimer's disease from local Caucasian knowledge." J Ethnopharmacol **92**(2-3): 147-162.

Heinrich, M. and H. L. Teoh (2004). "Galanthamine from snowdrop—the development of a modern drug against Alzheimer's disease from local Caucasian knowledge." Journal of Ethnopharmacology **92**(2-3): 147-162.

Heinrich, M. and H. L. Teoh (2004). "Galanthamine from snowdrop—the development of a modern drug against Alzheimer's disease from local Caucasian knowledge." Journal of ethnopharmacology **92**(2): 147-162.

Hesse, M. (2002). Alkaloids: nature's curse or blessing?, John Wiley & Sons.

Hirai, M. Y., M. Yano, D. B. Goodenowe, S. Kanaya, T. Kimura, M. Awazuhara, M. Arita, T. Fujiwara and K. Saito (2004). "Integration of transcriptomics and metabolomics for understanding of global responses to nutritional stresses in Arabidopsis thaliana." Proceedings of the National academy of Sciences of the United States of America **101**(27): 10205-10210.

Hollingsworth, R. G., J. W. Armstrong and E. Campbell (2002). "Pest control: caffeine as a repellent for slugs and snails." nature **417**(6892): 915-916.

Hotchandani, T. and I. Desgagne-Penix (2017). "Heterocyclic Amaryllidaceae Alkaloids: Biosynthesis and Pharmacological Applications." Current Topics in Medicinal Chemistry **17**(4): 418-427.

Hu, M., S. Peng, Y. He, M. Qin, X. Cong, Y. Xing, M. Liu and Z. Yi (2015). "Lycorine is a novel inhibitor of the growth and metastasis of hormone-refractory prostate cancer." Oncotarget **6**(17): 15348.

Hyun, T. K., Y. Rim, H.-J. Jang, C. H. Kim, J. Park, R. Kumar, S. Lee, B. C. Kim, J. Bhak, B. Nguyen-Quoc, S.-W. Kim, S. Y. Lee and J.-Y. Kim (2012). "De novo transcriptome

sequencing of Momordica cochinchinensis to identify genes involved in the carotenoid biosynthesis." Plant Molecular Biology **79**(4): 413-427.

Jiang, Y., B. Xia, L. Liang, X. Li, S. Xu, F. Peng and R. Wang (2013). "Molecular and analysis of a phenylalanine ammonia-lyase gene (LrPAL2) from Lycoris radiata." Mol Biol Rep **40**(3): 2293-2300.

Jiang, Y., B. Xia, L. Liang, X. Li, S. Xu, F. Peng and R. Wang (2013). "Molecular and analysis of a phenylalanine ammonia-lyase gene (LrPAL2) from Lycoris radiata." Molecular biology reports **40**(3): 2293-2300.

Jiang, Y., N. Xia, X. Li, W. Shen, L. Liang, C. Wang, R. Wang, F. Peng and B. Xia (2011). "Molecular cloning and characterization of a phenylalanine ammonia-lyase gene (LrPAL) from Lycoris radiata." Mol Biol Rep **38**(3): 1935-1940.

Jiang, Y., N. Xia, X. Li, W. Shen, L. Liang, C. Wang, R. Wang, F. Peng and B. Xia (2011). "Molecular cloning and characterization of a phenylalanine ammonia-lyase gene (LrPAL) from Lycoris radiata." Molecular biology reports **38**(3): 1935-1940.

Jin, J., J. G. Shen, W. Cai, G. H. Xie, F. R. Liao, F. L. Gao, J. F. Ma, X. H. Chen and Z. J. Wu (2017). "Narcissus yellow stripe virus and Narcissus mosaic virus detection in Narcissus via multiplex TaqMan-based reverse transcription-PCR assay." J Appl Microbiol **122**(5): 1299-1309.

Jin, Z. (2013). "Amaryllidaceae and Sceletium alkaloids." Natural product reports **30**(6): 849-868.

Jin, Z. and X.-H. Xu (2013). Amaryllidaceae Alkaloids. Natural Products, Springer**:** 479-522.

Katiyar, C., A. Gupta, S. Kanjilal and S. Katiyar (2012). "Drug discovery from plant sources: An integrated approach." Ayu **33**(1): 10.

Katoch, D., D. Kumar, U. Sharma, N. Kumar, Y. S. Padwad, B. Lal and B. Singh (2013). "Zephgrabetaine: a new betaine-type amaryllidaceae alkaloid from Zephyranthes grandiflora." Natural product communications **8**(2): 161-164.

Kilgore, M., M. M. Augustin, G. D. May, J. A. Crow and T. M. Kutchan (2016). "CYP96T1 of Narcissus sp. aff. pseudonarcissus Catalyzes Formation of the Para-Para' C-C Phenol Couple in the Amaryllidaceae Alkaloids." Frontiers in Plant Science **7**.

Kilgore, M., C. Holland, J. M. Jez, J. A. Crow and T. M. Kutchan (2016). "Identification of a Noroxomaritidine Reductase with Amaryllidaceae Alkaloid Biosynthesis Related Activities." Journal of Biological Chemistry.

Kilgore, M. B., M. M. Augustin, G. D. May, J. A. Crow and T. M. Kutchan (2016). "CYP96T1 of Narcissus sp. aff. pseudonarcissus Catalyzes Formation of the Para-Para'CC Phenol Couple in the Amaryllidaceae Alkaloids." Frontiers in plant science **7**: 225.

Kilgore, M. B., M. M. Augustin, C. M. Starks, M. O'Neil-Johnson, G. D. May, J. A. Crow and T. M. Kutchan (2014). "Cloning and characterization of a norbelladine 4′-O-methyltransferase involved in the biosynthesis of the Alzheimer's drug galanthamine in Narcissus sp. aff. pseudonarcissus." PloS one **9**(7): e103223.

Kilgore, M. B., C. K. Holland, J. M. Jez and T. M. Kutchan (2016). "Identification of a noroxomaritidine reductase with Amaryllidaceae alkaloid biosynthesis related activities." Journal of Biological Chemistry **291**(32): 16740-16752.

Kilgore, M. B. and T. M. Kutchan (2016). "The Amaryllidaceae alkaloids: biosynthesis and methods for enzyme discovery." Phytochemistry reviews **15**(3): 317-337.

Kornienko, A. and A. Evidente (2008). "Chemistry, biology, and medicinal potential of narciclasine and its congeners." Chemical reviews **108**(6): 1982-2014.

Kornienko, A. and A. Evidente (2008). "Chemistry, biology, and medicinal potential of narciclasine and its congeners." Chem Rev **108**(6): 1982-2014.

Kukula-Koch, W. and J. Widelski (2017). Alkaloids. Pharmacognosy, Elsevier**:** 163-198.

Kurkin, V. A. (2003). "Phenylpropanoids from Medicinal Plants: Distribution, Classification, Structural Analysis, and Biological Activity." Chemistry of Natural Compounds **39**(2): 123-153.

Kutchan, T. M. (2009). Plant-derived natural products: synthesis, function, and application. u.k.

Lamoral-Theys, D., A. Andolfi, G. Van Goietsenoven, A. Cimmino, B. Le Calvé, N. Wauthoz, V. Mégalizzi, T. Gras, C. Bruyère and J. Dubois (2009). "Lycorine, the main phenanthridine Amaryllidaceae alkaloid, exhibits significant antitumor activity in cancer cells that display resistance to proapoptotic stimuli: an investigation of structure– activity relationship and mechanistic insight." Journal of medicinal chemistry **52**(20): 6244-6256.

Langmead, B., C. Trapnell, M. Pop and S. L. Salzberg (2009). "Ultrafast and memory-efficient alignment of short DNA sequences to the human genome." Genome biology **10**(3): R25.

Lattanzio, V., V. M. Lattanzio and A. Cardinali (2006). "Role of phenolics in the resistance mechanisms of plants against fungal pathogens and insects." Phytochemistry: Advances in research **661**: 23-67.

Lee, E.-J. and P. Facchini (2010). "Norcoclaurine synthase is a member of the pathogenesis-related 10/Bet v1 protein family." The Plant Cell **22**(10): 3489-3503.

Lee, E. J. and P. Facchini (2010). "Norcoclaurine synthase is a member of the pathogenesis-related 10/Bet v1 protein family." Plant Cell **22**(10): 3489-3503.

Lee, M. J., J. S. Pate, D. J. Harris and C. A. Atkins (2006). "Synthesis, transport and accumulation of quinolizidine alkaloids in Lupinus albus L. and L. angustifolius L." Journal of experimental botany **58**(5): 935-946.

Li, B. and C. N. Dewey (2011). "RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome." BMC bioinformatics **12**(1): 323.

Li, J.-F., J. E. Norville, J. Aach, M. McCormack, D. Zhang, J. Bush, G. M. Church and J. Sheen (2013). "Multiplex and homologous recombination–mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9." Nature biotechnology **31**(8): 688.

Li, X., P. Xu, C. Wang, N. Xu, A. Xu, Y. Xu, T. Sadahira, M. Araki, K. Wada and E. Matsuura (2017). "Synergistic effects of the immune checkpoint inhibitor CTLA-4 combined with the growth inhibitor lycorine in a mouse model of renal cell carcinoma." Oncotarget **8**(13): 21177.

Lichman, B. R., M. C. Gershater, E. D. Lamming, T. Pesnot, A. Sula, N. H. Keep, H. C. Hailes and J. M. Ward (2015). "'Dopamine-first'mechanism enables the rational engineering of the norcoclaurine synthase aldehyde activity profile." The FEBS journal **282**(6): 1137-1151.

Lincoln, T. and E. Zeiger (2006). "Secondary metabolites and plant defense." Plant Physiology. Fourth Edition. Sinauer Associates, Inc. Capítulo **13**: 125.

Liscombe, D. K., B. P. MacLeod, N. Loukanina, O. I. Nandi and P. J. Facchini (2005). "Erratum to "Evidence for the monophyletic evolution of benzylisoquinoline alkaloid biosynthesis in angiosperms"[Phytochemistry 66 (2005) 1374–1393]." Phytochemistry **66**(20): 2500-2520.

Liu, L., Y. Li, S. Li, N. Hu, Y. He, R. Pong, D. Lin, L. Lu and M. Law (2012). "Comparison of next-generation sequencing systems." BioMed Research International **2012**.

Livak, K. J. and T. D. Schmittgen (2001). "Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method." Methods **25**(4): 402-408.

López, S., J. Bastida, F. Viladomat and C. Codina (2002). "Acetylcholinesterase inhibitory activity of some Amaryllidaceae alkaloids and Narcissus extracts." Life Sciences **71**(21): 2521-2529.

López, S., J. Bastida, F. Viladomat and C. Codina (2003). "Galanthamine pattern in Narcissus confusus plants." Planta medica **69**(12): 1166-1168.

Louw, C., T. Regnier and L. Korsten (2002). "Medicinal bulbous plants of South Africa and their traditional relevance in the control of infectious diseases." Journal of Ethnopharmacology **82**(2-3): 147-154.

Lubbe, A., H. Gude, R. Verpoorte and Y. H. Choi (2013). "Seasonal accumulation of major alkaloids in organs of pharmaceutical crop Narcissus Carlton." Phytochemistry **88**: 43-53.

Lubbe, A., B. Pomahacova, Y. H. Choi and R. Verpoorte (2010). "Analysis of metabolic variation and galanthamine content in Narcissus bulbs by 1H NMR." Phytochem Anal **21**(1): 66-72.

Lubbe, A., R. Verpoorte and Y. H. Choi (2012). "Effects of fungicides on galanthamine and metabolite profiles in Narcissus bulbs." Plant Physiol Biochem **58**: 116-123.

Maelicke, A. (2017). Cholinergic enhancers with improved blood-brain barrier permeability for the treatment of diseases accompanied by cognitive impairment, Google Patents.

Martin, L. B., Z. Fei, J. J. Giovannoni and J. K. Rose (2013). "Catalyzing plant science research with RNA-seq." Frontiers in plant science **4**.

Meerow, A. and D. Snijman (1998). Amaryllidaceae. Flowering Plants· Monocotyledons, Springer**:** 83-110.

Miao, J., D. Guo, J. Zhang, Q. Huang, G. Qin, X. Zhang, J. Wan, H. Gu and L.-J. Qu (2013). "Targeted mutagenesis in rice using CRISPR-Cas system." Cell research **23**(10): 1233.

Michalska, K., H. Fernandes, M. Sikorski and M. Jaskolski (2010). "Crystal structure of Hyp-1, a St. John's wort protein implicated in the biosynthesis of hypericin." Journal of structural biology **169**(2): 161-171.

Minami, H., E. Dubouzet, K. Iwasa and F. Sato (2007). "Functional analysis of norcoclaurine synthase in Coptis japonica." Journal of Biological Chemistry **282**(9): 6274-6282.

Minami, H., E. Dubouzet, K. Iwasa and F. Sato (2007). "Functional analysis of norcoclaurine synthase in Coptis japonica." J Biol Chem **282**(9): 6274-6282.

Morris, J. S. and P. J. Facchini (2016). "Isolation and characterization of reticuline N-methyltransferase involved in biosynthesis of the aporphine alkaloid magnoflorine in opium poppy." Journal of Biological Chemistry **291**(45): 23416-23427.

Mukherjee, A. K., S. Basu, N. Sarkar and A. C. Ghosh (2001). "Advances in cancer therapy with plant based natural products." Current medicinal chemistry **8**(12): 1467-1486.

Muranaka, T. and K. Saito (2013). Phytochemical genomics on the way, Oxford University Press.

Ncube, B. and J. Van Staden (2015). "Tilting plant metabolism for improved metabolite biosynthesis and enhanced human benefit." Molecules **20**(7): 12698-12731.

Nessler, C. L. and R. J. Burnett (1992). "Organization of the major latex protein gene family in opium poppy." Plant molecular biology **20**(4): 749-752.

Niehl, A., C. Lacomme, A. Erban, J. Kopka, U. Krämer and J. Fisahn (2006). "Systemic Potato virus X infection induces defence gene expression and accumulation of β-phenylethylamine-alkaloids in potato." Functional plant biology **33**(6): 593-604.

Nishihachijo, M., Y. Hirai, S. Kawano, A. Nishiyama, H. Minami, T. Katayama, Y. Yasohara, F. Sato and H. Kumagai (2014). "Asymmetric synthesis of tetrahydroisoquinolines by enzymatic Pictet–Spengler reaction." Bioscience, biotechnology, and biochemistry **78**(4): 701-707.

O'Connor, S. E. and J. J. Maresh (2006). "Chemistry and biology of monoterpene indole alkaloid biosynthesis." Natural product reports **23**(4): 532-547.

Oksman-Caldentey, K.-M. and D. Inzé (2004). "Plant cell factories in the post-genomic era: new ways to produce designer secondary metabolites." Trends in plant science **9**(9): 433-440.

Oksman-Caldentey, K.-M. and K. Saito (2005). "Integrating genomics and metabolomics for engineering plant metabolic pathways." Current opinion in biotechnology **16**(2): 174-179.

Onoyovwe, A., J. M. Hagel, X. Chen, M. F. Khan, D. C. Schriemer and P. J. Facchini (2013). "Morphine biosynthesis in opium poppy involves two cell types: sieve elements and laticifers." Plant Cell **25**(10): 4110-4122.

Onoyovwe, A., J. M. Hagel, X. Chen, M. F. Khan, D. C. Schriemer and P. J. Facchini (2013). "Morphine biosynthesis in opium poppy involves two cell types: sieve elements and laticifers." The Plant Cell **25**(10): 4110-4122.

Osbourn, A. E. and V. Lanzotti (2009). Plant-derived natural products, Springer.
Pasquo, A., A. Bonamore, S. Franceschini, A. Macone, A. Boffi and A. Ilari (2008).

"Cloning, expression, crystallization and preliminary X-ray data analysis of norcoclaurine synthase from Thalictrum flavum." Acta Crystallographica Section F: Structural Biology and Crystallization Communications **64**(4): 281-283.

Pateraki, I. and A. K. Kanellis (2010). "Stress and developmental responses of terpenoid biosynthetic genes in Cistus creticus subsp. creticus." Plant cell reports **29**(6): 629-641.

Pelletier, S. W. (1983). "The nature and definition of an alkaloid." Alkaloids: Chemical and Biological Perspectives, John Wiley & Sons, New York, USA.

Pettit, G. R., N. Melody, D. L. Herald, J. M. Schmidt, R. K. Pettit and J. C. Chapuis (2002). "Synthesis of 10b (R)-hydroxypancratistatin, 10b (S)-hydroxy-1-epipancratistatin, 10b (S)-hydroxy-1, 2-diepipancratistatin and related isocarbostyrils." Heterocycles **56**(1-2): 139-155.

Pfaffenbach, M. and T. Gaich (2017). The Rhazinilam-Leuconoxine-Mersicarpine Triad of Monoterpenoid Indole Alkaloids. The Alkaloids: Chemistry and Biology, Elsevier. **77:** 1-84.

Plaitakis, A. and R. C. Duvoisin (1983). "Homer's moly identified as Galanthus nivalis L.: physiologic antidote to stramonium poisoning." Clinical neuropharmacology **6**(1): 1-5.

Pluskal, T., S. Castillo, A. Villar-Briones and M. Orešič (2010). "MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data." BMC Bioinformatics **11**(1): 395.

Podstolski, A., D. Havkin-Frenkel, J. Malinowski, J. W. Blount, G. Kourteva and R. A. Dixon (2002). "Unusual 4-hydroxybenzaldehyde synthase activity from tissue cultures of the vanilla orchid Vanilla planifolia." Phytochemistry **61**(6): 611-620.

Qi, H. and S. Li (2014). "Dose–response meta-analysis on coffee, tea and caffeine consumption with risk of Parkinson's disease." Geriatrics & gerontology international **14**(2): 430-439.

Rai, A. and K. Saito (2016). "Omics data input for metabolic modeling." Current opinion in biotechnology **37**: 127-134.

Rai, A., K. Saito and M. Yamazaki (2017). "Integrated omics analysis of specialized metabolism in medicinal plants." The Plant Journal **90**(4): 764-787.

Refaat, J., M. S. Kamel, M. A. Ramadan and A. A. Ali (2013). "Crinum; an endless source of bioactive principles: a review. part v. biological profile." International Journal of Pharmaceutical Sciences and Research **4**(4): 1239.

Riaño-Pachón, D. M., A. Nagel, J. Neigenfind, R. Wagner, R. Basekow, E. Weber, B. Müller-Röber, S. Diehl and B. Kersten (2008). "GabiPD: the GABI primary database—a plant integrative 'omics' database." Nucleic acids research **37**(suppl_1): D954-D959.

Rischer, H., M. Orešič, T. Seppänen-Laakso, M. Katajamaa, F. Lammertyn, W. Ardiles-Diaz, M. C. Van Montagu, D. Inzé, K.-M. Oksman-Caldentey and A. Goossens (2006).

"Gene-to-metabolite networks for terpenoid indole alkaloid biosynthesis in Catharanthus roseus cells." Proceedings of the National Academy of Sciences **103**(14): 5614-5619.

Ro, D.-K., E. M. Paradise, M. Ouellet, K. J. Fisher, K. L. Newman, J. M. Ndungu, K. A. Ho, R. A. Eachus, T. S. Ham, J. Kirby, M. C. Y. Chang, S. T. Withers, Y. Shiba, R. Sarpong and J. D. Keasling (2006). "Production of the antimalarial drug precursor artemisinic acid in engineered yeast." Nature **440**: 940.

Roberts, M. F. (2013). Alkaloids: biochemistry, ecology, and medicinal applications, Springer Science & Business Media.

Roberts, M. F., D. McCarthy, T. M. Kutchan and C. J. Coscia (1983). "Localization of enzymes and alkaloidal metabolites in Papaver latex." Archives of biochemistry and biophysics **222**(2): 599-609.

Rønsted, N., V. Savolainen, P. Mølgaard and A. K. Jäger (2008). "Phylogenetic selection of Narcissus species for drug discovery." Biochemical Systematics and Ecology **36**(5-6): 417-422.

Ruiz, N., D. Ward and D. Saltz (2002). "Responses of Pancratium sickenbergeri to simulated bulb herbivory: combining defence and tolerance strategies." Journal of Ecology **90**(3): 472-479.

Samanani, N., J. Alcantara, R. Bourgault, K. G. Zulak and P. J. Facchini (2006). "The role of phloem sieve elements and laticifers in the biosynthesis and accumulation of alkaloids in opium poppy." The Plant Journal **47**(4): 547-563.

Samanani, N. and P. J. Facchini (2001). "Isolation and partial characterization of norcoclaurine synthase, the first committed step in benzylisoquinoline alkaloid biosynthesis, from opium poppy." Planta **213**(6): 898-906.

Samanani, N. and P. J. Facchini (2002). "Purification and Characterization of Norcoclaurine Synthase THE FIRST COMMITTED ENZYME IN BENZYLISOQUINOLINE ALKALOID BIOSYNTHESIS IN PLANTS." Journal of Biological Chemistry **277**(37): 33878-33883.

Samanani, N., D. K. Liscombe and P. J. Facchini (2004). "Molecular cloning and characterization of norcoclaurine synthase, an enzyme catalyzing the first committed step in benzylisoquinoline alkaloid biosynthesis." Plant J **40**.

Samanani, N., D. K. Liscombe and P. J. Facchini (2004). "Molecular cloning and characterization of norcoclaurine synthase, an enzyme catalyzing the first committed step in benzylisoquinoline alkaloid biosynthesis." Plant J **40**(2): 302-313.

Samanani, N., D. K. Liscombe and P. J. Facchini (2004). "Molecular cloning and characterization of norcoclaurine synthase, an enzyme catalyzing the first committed step in benzylisoquinoline alkaloid biosynthesis." The Plant Journal **40**(2): 302-313.

Schnable, P. S., D. Ware, R. S. Fulton, J. C. Stein, F. Wei, S. Pasternak, C. Liang, J. Zhang, L. Fulton and T. A. Graves (2009). "The B73 maize genome: complexity, diversity, and dynamics." science **326**(5956): 1112-1115.

Schoental, R. (1968). "Toxicology and carcinogenic action of pyrrolizidine alkaloids." Cancer Research **28**(11): 2237-2246.

Schoental, R. and P. Magee (1959). "Further observations on the subacute and chronic liver changes in rats after a single dose of various pyrrolizidine (Senecio) alkaloids." The Journal of Pathology **78**(2): 471-482.

Schrader, K. K., A. Andolfi, C. L. Cantrell, A. Cimmino, S. O. Duke, W. Osbrink, D. E. Wedge and A. Evidente (2010). "A survey of phytotoxic microbial and plant metabolites as potential natural products for pest management." Chemistry & biodiversity **7**(9): 2261-2280.

Sener, B., I. Orhan and J. Satayavivad (2003). "Antimalarial activity screening of some alkaloids and the plant extracts from Amaryllidaceae." Phytother Res **17**(10): 1220-1223.

Shawky, E., A. H. Abou-Donia, F. A. Darwish, S. M. Toaima, S. S. Takla, N. B. Pigni and J. Bastida (2015). "HPTLC and GC/MS study of Amaryllidaceae alkaloids of two Narcissus species." Chemistry & biodiversity **12**(8): 1184-1199.

Silva, A., J. De Andrade, K. Machado, A. Rocha, M. Apel, M. Sobral, A. Henriques and J. Zuanazzi (2008). "Screening for cytotoxic activity of extracts and isolated alkaloids from bulbs of Hippeastrum vittatum." Phytomedicine **15**(10): 882-885.

Singh, A. and I. Desgagne-Penix (2014). "Biosynthesis of the Amaryllidaceae alkaloids." Plant Science Today **1**(2): 114-120.

Singh, A. and I. Desgagné-Penix (2014). "Biosynthesis of the Amaryllidaceae alkaloids." Plant Science Today **1**(3): 114-120.

Singh, A. and I. Desgagné-Penix (2015). Chapter 3 : Biosynthesis of Amaryllidaceae alkaloids: A biochemical outlook. In: Alkaloids: Biosynthesis, Biological Roles and Health benefits., Nova Science Publishers.

Singh, A. and I. Desgagné-Penix (2017). "Transcriptome and metabolome profiling of Narcissus pseudonarcissus 'King Alfred'reveal components of Amaryllidaceae alkaloid metabolism." Scientific reports **7**(1): 17356.

Sobarzo-Sanchez, E. (2015). "Biosynthesis, biological roles and health benefits."
Springob, K. and T. M. Kutchan (2009). Introduction to the different classes of natural products. Plant-Derived Natural Products, Springer**:** 3-50.

Steiner-Lange, S., A. Fischer, A. Boettcher, I. Rouhara, H. Liedgens, E. Schmelzer and W. Knogge (2003). "Differential defense reactions in leaf tissues of barley in response to

infection by Rhynchosporium secalis and to treatment with a fungal avirulence gene product." Molecular plant-microbe interactions **16**(10): 893-902.

Stöckigt, J., L. Barleben, S. Panjikar and E. A. Loris (2008). "3D-Structure and function of strictosidine synthase–the key enzyme of monoterpenoid indole alkaloid biosynthesis." Plant Physiology and Biochemistry **46**(3): 340-355.

Strack, D. (1997). "10 Phenolic Metabolism." Plant biochemistry: 387.

Sundarasekar, J., G. Sahgal and S. Subramaniam (2012). "Anti-candida activity by Hymenocallis littoralis extracts for opportunistic oral and genital infection Candida albicans." Bangladesh Journal of Pharmacology **7**(3): 211-216.

Szlavik, L., A. Gyuris, J. Minarovits, P. Forgo, J. Molnar and J. Hohmann (2004). "Alkaloids from Leucojum vernum and antiretroviral activity of Amaryllidaceae alkaloids." Planta Med **70**(9): 871-873.

Szlávik, L., Á. Gyuris, J. Minárovits, P. Forgo, J. Molnár and J. Hohmann (2004). "Alkaloids from Leucojum vernum and antiretroviral activity of Amaryllidaceae alkaloids." Planta medica **70**(09): 871-873.

Takos, A. M. and F. Rook (2013). "Towards a molecular understanding of the biosynthesis of amaryllidaceae alkaloids in support of their expanding medical use." Int J Mol Sci **14**(6): 11713-11741.

Takos, A. M. and F. Rook (2013). "Towards a molecular understanding of the biosynthesis of Amaryllidaceae alkaloids in support of their expanding medical use." International journal of molecular sciences **14**(6): 11713-11741.

Thomsen, T. and H. Kewitz (1990). "Selective inhibition of human acetylcholinesterase by galanthamine in vitro and in vivo." Life Sci **46**(21): 1553-1558.

Tijssen, P., M. Agbandje-McKenna, J. M. Almendral, M. Bergoin, T. W. Flegel, K. Hedman and P. Tattersall (2011). "Virus taxonomy 9th Report of the International Committee on Taxonomy of Viruses. ." 405–425.

Tohge, T., Y. Nishiyama, M. Y. Hirai, M. Yano, J. i. Nakajima, M. Awazuhara, E. Inoue, H. Takahashi, D. B. Goodenowe and M. Kitayama (2005). "Functional genomics by integrated analysis of metabolome and transcriptome of Arabidopsis plants over-expressing an MYB transcription factor." The Plant Journal **42**(2): 218-235.

Tojo, E. (1991). "(+)-Narcidine, a new alkaloid from Narcissus pseudonarcissus." Journal of Natural Products **54**(5): 1387-1388.

Torras-Claveria, L., S. Berkov, C. Codina, F. Viladomat and J. Bastida (2014). "Metabolomic analysis of bioactive Amaryllidaceae alkaloids of ornamental varieties of

Narcissus by GC–MS combined with k-means cluster analysis." Industrial Crops and Products **56**: 211-222.

Urbanczyk-Wochniak, E., A. Luedemann, J. Kopka, J. Selbig, U. Roessner-Tunali, L. Willmitzer and A. R. Fernie (2003). "Parallel analysis of transcript and metabolic profiles: a new approach in systems biology." EMBO reports **4**(10): 989-993.

Van Dijk, E. L., H. Auger, Y. Jaszczyszyn and C. Thermes (2014). "Ten years of next-generation sequencing technology." Trends in genetics **30**(9): 418-426.

Van Goietsenoven, G., J. Hutton, J.-P. Becker, B. Lallemand, F. Robert, F. Lefranc, C. Pirker, G. Vandenbussche, P. Van Antwerpen and A. Evidente (2010). "Targeting of eEF1A with Amaryllidaceae isocarbostyrils as a strategy to combat melanomas." The FASEB journal **24**(11): 4575-4584.

Verma, N. and S. Shukla (2015). "Impact of various factors responsible for fluctuation in plant secondary metabolites." Journal of Applied Research on Medicinal and Aromatic Plants **2**(4): 105-113.

Wang, L., T. Beuerle, J. Timbilla and D. Ober (2012). "Independent recruitment of a flavin-dependent monooxygenase for safe accumulation of sequestered pyrrolizidine alkaloids in grasshoppers and moths." PLoS One **7**(2): e31796.

Wang, P., H.-H. Yuan, X. Zhang, Y.-P. Li, L.-Q. Shang and Z. Yin (2014). "Novel lycorine derivatives as anticancer agents: synthesis and in vitro biological evaluation." Molecules **19**(2): 2469-2480.

Wang, R., S. Xu, Y. Jiang, J. Jiang, X. Li, L. Liang, J. He, F. Peng and B. Xia (2013). "De novo sequence assembly and characterization of Lycoris aurea transcriptome using GS FLX titanium platform of 454 pyrosequencing." PLoS One **8**(4): e60449.

Weckwerth, W. (2003). "Metabolomics in systems biology." Annual review of plant biology **54**(1): 669-689.

Wink, M. (2002). "Production of quinolizidine alkaloids in in vitro cultures of legumes." Biotechnology in agriculture and forestry: 118-136.

Wylie, S. J., H. Li, K. Sivasithamparam and M. G. Jones (2014). "Complete genome analysis of three isolates of narcissus late season yellows virus and two of narcissus yellow stripe virus: three species or one?" Arch Virol.

Wylie, S. J., H. Li, K. Sivasithamparam and M. G. Jones (2014). "Complete genome analysis of three isolates of narcissus late season yellows virus and two of narcissus yellow stripe virus: three species or one?" Archives of virology **159**(6): 1521-1525.

Xiao, M., Y. Zhang, X. Chen, E.-J. Lee, C. J. Barber, R. Chakrabarty, I. Desgagné-Penix, T. M. Haslam, Y.-B. Kim and E. Liu (2013). "Transcriptome analysis based on next-

generation sequencing of non-model plants producing specialized metabolites of biotechnological interest." Journal of biotechnology **166**(3): 122-134.

Yamazaki, M., K. Mochida, T. Asano, R. Nakabayashi, M. Chiba, N. Udomson, Y. Yamazaki, D. B. Goodenowe, U. Sankawa and T. Yoshida (2013). "Coupling deep transcriptome analysis with untargeted metabolic profiling in Ophiorrhiza pumila to further the understanding of the biosynthesis of the anti-cancer alkaloid camptothecin and anthraquinones." Plant and cell physiology **54**(5): 686-696.

Yang, H., J. Barros-Rios, G. Kourteva, X. Rao, F. Chen, H. Shen, C. Liu, A. Podstolski, F. Belanger, D. Havkin-Frenkel and R. A. Dixon (2017). "A re-evaluation of the final step of vanillin biosynthesis in the orchid Vanilla planifolia." Phytochemistry **139**: 33-46.

Ziegler, J. and P. J. Facchini (2008). "Alkaloid biosynthesis: metabolism and trafficking." Annu. Rev. Plant Biol. **59**: 735-769.

Ziegler, J. and P. J. Facchini (2008). "Alkaloid biosynthesis: metabolism and trafficking." Annu Rev Plant Biol **59**: 735-769.

Zulak, K. G., D. K. Liscombe, H. Ashihara and P. J. Facchini (2006). "Alkaloids." Plant Secondary Metabolites: Occurrence, Structure and Role in the Human Diet: 102-136.

ANNEX A

BIOSYNTHESIS OF AMARYLLIDACEAE ALKALODS

APARNA SINGH AND ISABEL DESGAGNE PENIX

Published on 10th of July 2014 in *Plant Science Today*

A.1 Abstract

Amaryllidaceae alkaloids (Grabherr, Haas et al.) are a structurally diverse group of plant specialized metabolites with powerful biological activities. The medicinal properties of many AAs have been identified including the anti-tumor agent narciclasine and galanthamine, used for Alzheimer's disease. Tracer studies have led to proposed pathways but AA biosynthesis remains molecularly uncharacterized. The use of systems biology-based approaches could lead to the unraveling of AA metabolic pathways. The elucidation of AA biosynthesis will provide necessary tools required to enhance AA production in plants as well as the development of microbial production platforms as an alternative to plants as a commercial source of valuable AAs.

Keywords: amaryllidaceae alkaloid; plant secondary metabolism; galanthamine biosynthesis; natural products; systems biology.

A.2 Introduction

For centuries, a large number of plants have been used as essential resource for therapeutic agents against various human illnesses. Among these traditional sources of medicines, the bulbous monocot Amaryllidaceae family, comprising over 1100 species, are among the top 20 of the most widely applied plant families (Jin, 2013). In addition to their pharmacological interests, the well-known horticultural ornamental value of

Amaryllidaceae plants (*Amaryllis, Narcissus* and *Galanthus*), has long been recognized and exploited.

The medicinal properties of the Amaryllidaceae are owed to the presence of specialized metabolites, the Amaryllidaceae alkaloids (Grabherr, Haas et al.), which are specific to this plant family. To date, more than 500 structurally diverse AAs have been isolated (Jin, 2013) and modern phytochemical studies have shown the various pharmacological activities of numerous AAs. These include analgesic, anti-cancer, and anti-microbial activities (Bastida et al., 2011; Heinrich & Lee Teoh, 2004; Kornienko & Evidente, 2008; Nair & van Staden, 2013). For example, narciclasine displays potent anti-tumor activity (Kornienko & Evidente, 2008) whereas galanthamine, an acetylcholinesterase inhibitor, is used to treat symptoms of Alzheimer's disease (Heinrich & Lee Teoh, 2004). As it is the case for many valuable plant natural products, the commercial use of most AAs is restricted by their limited availabilities due to their low concentrations in plants. For example, the yield of galanthamine from leaves of Leucojum aestivum is less than 0.2% DW and varies between 0.059-0.166% DW from plants of different geographical origin (Berkov, Bastida, Viladomat, & Codina, 2011). Narciclasine levels are higher in Narcissus tazetta (65mg/kg) compared with other Amaryllidaceaea species such as L. aestivum (30mg/kg) and Galanthus nivalis (10mg/kg) however this yield is still too low for lucrative applications (Kornienko & Evidente, 2008). Currently, only galanthamine is used clinically and almost entirely produced from Leucojum and Narcissus cultivated plant species (Takos & Rook, 2013). Most AAs available commercially are obtained from their natural sources and only a few are produced synthetically. Although several chemical syntheses for galanthamine, narciclasine and others AAs have been described (Kornienko & Evidente, 2008; Takos & Rook, 2013), chemists are still struggling to devise a reasonably short synthetic pathway that can be used to produce these compounds on an economically viable commercial scale. Since most AAs display promising biological activities [but are produced in small quantities relative to the potential clinical and commercial demand], there are obvious interests in engineering AA production in plants or in heterologous expression systems. However, the lack of information on AA biosynthetic pathways, regulation and transport makes this task

very challenging. To date, several AA pathways remain hypothetical and enzymes and metabolic intermediates await discovery. Recently, systems biology-based approaches have facilitated the discovery of biosynthetic genes involved in natural products metabolic pathways in other plant families through the integration of multiple 'omics' (genomic, transcriptomic, proteomic and metabolomic) datasets. This strategy appears to be a promising avenue for the elucidation of AA biosynthesis in order to identify and characterize biosynthetic enzymes. In the scope of this review, we are focusing on the molecular basis of AA biosynthesis and we will propose putative gene families involved in AA biosynthetic pathways.

A.3 Biosynthesis

In contrast to the extensive literature on the pharmaceutical effects of AAs, information on their molecular genetics and biochemical pathways is incomplete. Previous studies using radiolabeled precursors led to the biochemical elucidation of the first steps in AA biosynthesis (Barton & Cohen, 1957; Barton, Kirby, Taylor, & Thomas, 1963; Battersby, Fales, & Wildman, 1961; Eichhorn, Takada, Kita, & Zenk, 1998). Despite the vast structural diversity, AAs share a common central intermediate, norbelladine, which is formed through the condensation of aromatic amino acids derivatives; 3,4-dihydroxybenzaldehyde (3,4-DHBA also named protocatechuic aldehyde) and tyramine (Fig. 1). Interestingly, similarities to several plant specialized biosynthetic pathways can be observed. For example, the formation of 3,4-DHBA features the initial steps of the phenylpropanoid biosynthesis whereas the formation of the tyramine precursor is analogous to isoquinoline alkaloid biosynthesis. The resulting tetrahydroisoquinoline core of norbelladine is central to the biosynthesis of many structural types of AAs.

A.3.1 Initial biosynthesis reaction

The substrates and enzymes necessary for the first committed steps in plant specialized metabolism often appear to have been recruited from primary metabolic

pathways (Chu, Wegel, & Osbourn, 2011). Substrates for the initial reactions of alkaloids biosynthesis are derived from primary metabolism, especially the aromatic amino acids L-phenylalanine and L-tyrosine. Additionally, since gene duplication is one of the most important components for the evolution of specialized metabolism, there are several "later" enzymes in these pathways that have ancestors in primary metabolism. For example, the enzyme homospermidine synthase, which catalyzes the first committed step to pyrrolizidine alkaloids, has most likely been developed by gene duplication from deoxyhypusin synthase, an enzyme from the polyamine biosynthetic pathway involved in cell division and growth (Moll et al., 2002).

In the initial stages of AA biosynthesis, the enzyme phenylalanine ammonia lyase (PAL) catalyzes the elimination of ammonia to generate trans-cinnamic acid (Fig. 1). Two hydroxylation reactions catalyzed by cytochrome P450s, cinnamate-4-hydroxylase (Ca4H) and coumarate-3-hydroxylase (Ca3H) followed by the loss of two carbon atoms leads to the formation of the C6 C1 precursor, 3,4-DHBA (Bastida et al., 2011; Eichhorn et al., 1998; Grisebach, 1973). These reactions are also part of phenylpropanoid metabolism, so that AA biosynthesis nicely shows the connections between two plant specialized pathways, alkaloid and phenylpropanoid. On the other hand, tyrosine is decarboxylated to tyramine by tyrosine decarboxylase (TYDC) (Bastida et al., 2011; Eichhorn et al., 1998; Grisebach, 1973; Takos & Rook, 2013). TYDC is a unique and key regulatory enzyme in many alkaloid-producing plants. In addition to controlling the transition from primary to specialized metabolism, TYDC ensures an adequate supply of tyramine for the synthesis of various isoquinoline alkaloids including the well-known narcotic analgesics morphine and codeine (Desgagne-Penix & Facchini, 2011; Facchini, 2001; Facchini, Huber-Allanach, & Tari, 2000; Hagel & Facchini, 2013; Ziegler & Facchini, 2008). Although the enzymatic reactions leading to the formation of AA precursors still need confirmation, feeding experiments in N. pseudonarcissus (Eichhorn et al., 1998) and recent transcriptome analyses of Lycoris aurea (Wang et al., 2013) have proven the presence of metabolite intermediates and transcripts for PAL, C4H, C3H and TYDC. These findings support the current proposed pathways. However, no gene sequences or enzymes have been isolated and characterized to date.

A.3.2 Formation of norbelladine skeleton

All AAs are viewed as derivatives of the common skeleton structure of norbelladine (Fig. 1). The combination of 3,4-DHBA and tyramine results in the formation of a Schiff base intermediate which following reduction yields norbelladine (Dewick, 2009; Ghosal, Shanthy, & Singh, 1988). Several similar condensation reactions have been described in other plant alkaloid biosynthetic pathways. For example, the first committed step in the formation of benzylisoquinoline alkaloids (BIA) is the combination of two L-tyrosine derivatives (4-hydroxyphehylacetaldehyde and dopamine) by a Pictet-Spengler condensation catalyzed by norcoclaurine synthase (NCS) to produce the trihydroxytetrahydroisoquinoline alkaloid norcoclaurine (Luk, Bunn, Liscombe, Facchini, & Tanner, 2007). NCS catalyzes a two-step reaction mechanism; a condensation reaction followed by an intramolecular cyclization (Luk et al., 2007). The electron-donating oxygen of the hydroxyl group of dopamine (missing in tyramine) is essential for the reaction to proceed and may provide a mechanistic explanation for the absence of cyclization in the condensation step of AA biosynthesis. Two different protein families (pathogenesis-related PR10-Betv1 and 2-oxoglutarate-dependent dioxygenase) have been reported to catalyze this reaction in vitro. However, the PR10/Betv1-NCS was shown to be the major one involved in BIA biosynthesis in *Papaver somniferum* and *Thalictrum flavum* plants (Lee & Facchini, 2010; Minami, Dubouzet, Iwasa, & Sato, 2007; Samanani, Liscombe, & Facchini, 2004). Another example of a Pictet-Spengler condensation is the combination of tryptamine and secologanin resulting in strictosidine, the general precursor of monoterpenoid indole alkaloids. In *Rauvolfia serpentina*, this reaction is catalyzed by strictosidine synthase, a member of the six-bladed four-stranded β-propeller fold protein family (Stockigt, Barleben, Panjikar, & Loris, 2008).

The examples above show that the initial condensation reactions in alkaloid biosynthesis can be catalyzed by members of very different protein families. This suggests that the type of enzyme recruited for the non-cyclizing condensation reaction in AA

biosynthesis may belong to one of these families. However, additional classes of enzymes should not be

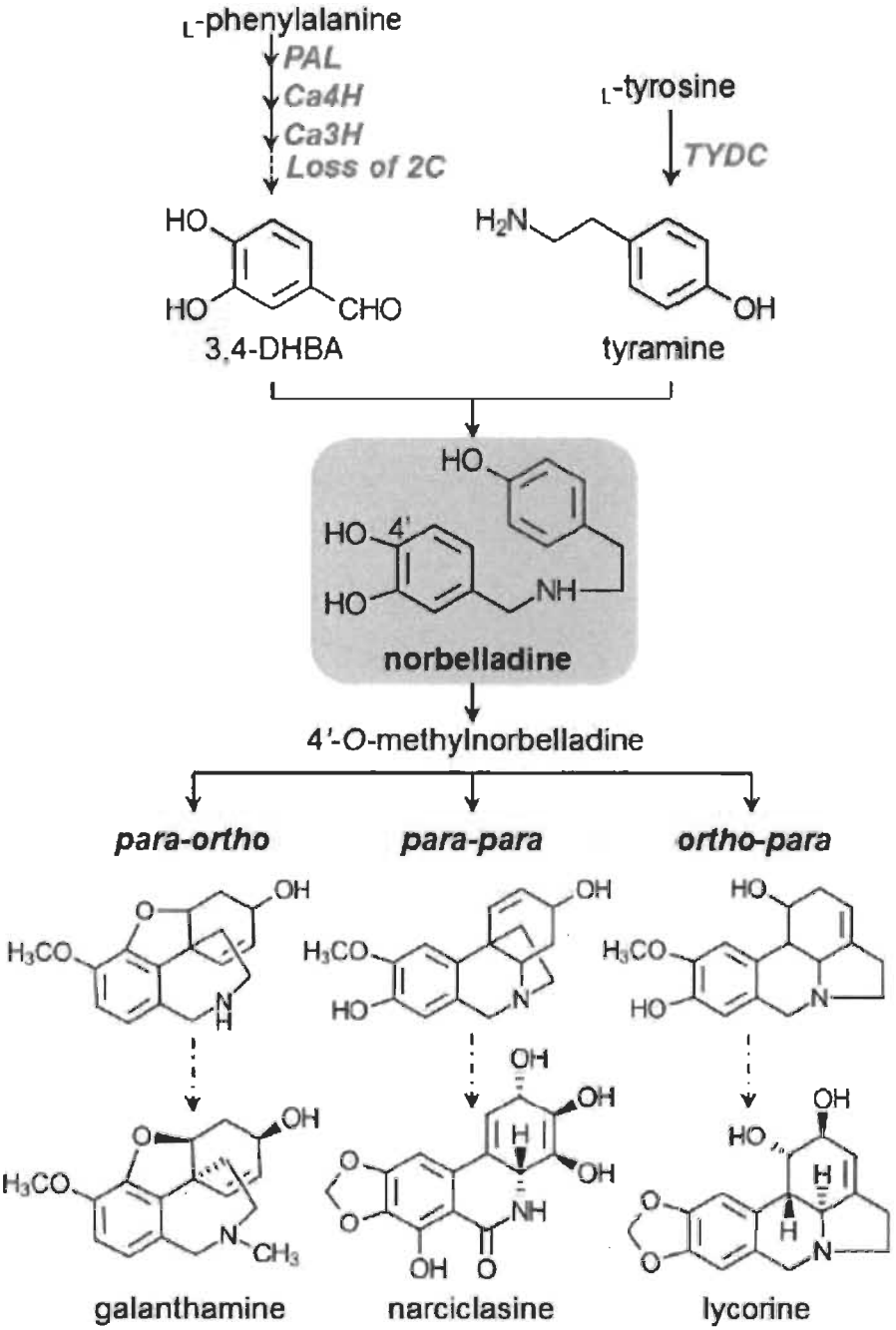


Figure. 1 **Proposed AA biosynthetic pathway.** Top: Early biosynthetic reactions leading to the norbelladine precursor (shaded gray). Middle: representation of the three backbone structures derived from the alternative phenol coupling reactions. Bottom: examples of chemical structures of AAs in each subgroups. Abbreviations defined in text

ruled out. It has been proposed that O-methylation of norbelladine precedes the oxidative phenol coupling, which yields the structurally diverse AAs. Plant O-methylation reactions are common transformation in the biosynthesis of alkaloids and are most often catalyzed by S-adenosyl-L-methionine (SAM)-dependent methyltransferases (MTs) (Liscombe, Louie, & Noel, 2012). Thus, it is assumed that norbelladine must be 4'-O-methylated to form 4'-O-methylnorbelladine. This compound then serves as the central intermediate from which multiple biosynthetic pathways lead to variousstructural types of AAs (Fig. 1).

A.3.3 Phenol coupling reaction and formation of diverse AA subgroup backbones

A crucial step in AA biosynthesis is the cyclization of 4'-O-methylnorbelladine by oxidative C-C phenol coupling, which can occur in *ortho-para'*, *para-ortho'* and *para-para'* positions. These alternative phenol coupling reactions generate three backbone structures. Accordingly, the different AA subgroups are referred to as *ortho-para*, *para-ortho'* and *para-para'* (Fig. 1). For example, anti-tumor AAs of the narciclasine-type are derived by a *para-para'* phenol coupling step (Bastida et al., 2011; Kornienko & Evidente, 2008; Takos & Rook, 2013). Recently, several enzymes have been discovered in various alkaloids pathways, which catalyze oxidative C-C phenol coupling reactions. All these enzymes were found to be members of the plant cytochromes P450 (CYP450s) (Mizutani & Sato, 2011). CYP450s catalyze a wide variety of monooxygenation/hydroxylation reaction in specialized metabolism and some of them are involved in unusual reactions such as methylenedioxy-bridge formation, oxidative C-C bond cleavage, oxidative rearrangement of carbon skeletons and phenol coupling reactions. For example, the intramolecular C-C phenol coupling catalyzed by CYP80G2 participates in the formation of aporphine-type alkaloid in *Coptis japonica* (Ikezawa,Iwasa, & Sato, 2008). Substrate specificity studies of CYP80G2 revealed the promiscuity of this enzyme as it reacted with several derivatives of reticuline such as (R,S)-orientaline, (R,S)-codamine, (S)-N-methylcoclaurine, and (S)-coclaurine informing that it could accept different functional group on the C-ring. However, CYP80G2 did not react with derivatives whose functional group were different from reticuline on the A-ring suggesting that the A-ring plays an

important role in substrate recognition (Ikezawa et al., 2008). Another example is salutaridine synthase, a member of the CYP719B1, which catalyzes the *para-ortho'* coupling of reticuline in the biosynthesis of morphinan alkaloids in *P. somniferum*. CYP719B1 is highly stereospecific as it catalyzes the C-C phenol coupling of only (R)-reticuline resulting in formation of salutaridine (Gesell et al., 2009). These examples demonstrate that specific CYP450s are able to catalyze intramolecular C-C phenol coupling reactions in alkaloid biosynthesis and that members of two CYP450 families, CYP80 and CYP719, have been identified. A recent comparative analysis of plant CYP450 sequences has shown that only the CYP80 family occurs in monocot plant although no Amaryllidaceae species were represented (Nelson & Werck-Reichhart, 2011). The CYP80 family has been associated with phenolic coupling reactions in several species whereas CYP719s are restricted to Ranunculales plants including *P. somniferum* and *C. japonica*. In addition, the latter steps of AA biosynthesis involve multiple C-C and C-O bond creation including methyledioxy-bridge formation. In other alkaloid pathways such as the BIA, these reactions are catalyzed by CYP719s and CYP80s (Diaz Chavez, Rolf, Gesell, & Kutchan, 2011; Gesell et al., 2009; Ikezawa et al., 2008). Altogether, these studies suggest that the phenol coupling and the methyledioxy-bridge formation steps in AA biosynthesis are likely to involve CYP450 of the CYP80 family. However, additional families of enzymes should not be ruled out. Interestingly, AAs from all three subgroups of phenol coupling can co-occur in a single plant, whereas some species and cultivars contains only one subgroup of AA (Berkov, Martinez-Frances, Bastida, Codina, & Rios, 2014; Kornienko & Evidente, 2008), This further suggests the existence of several CYP450 genes with different substrate, and product specificity with respect to the positions of intramolecular oxidative C-C phenol coupling.

A.3.4 Core structure modifications and decorations

The three backbone structures obtained from the phenol coupling steps form the scaffolds of further alkaloid diversity leading to more than 500 AAs known to date (Jin,2013). For example, in BIA biosynthesis, the reticuline skeleton is a basic building block for several types of isoquinoline alkaloids divided in several categories including

aporphines, benzophenanthridines, protoberberines and morphinans (Desgagne-Penix & Facchini, 2011). A complex network of enzymatic reactions will produce a spectrum of compounds that accumulate species-, cultivar-, tissue-, and development-specific. These chemical modifications are achieved by a multitude of enzymes catalyzing various types of reactions, such as O and N-methylations (OMTs, NMTs), C-C and C-O bond formations, oxidations and reductions, demethylations, and hydroxylations resulting in a variety of different structures. Fig. 1 shows an example for each phenol coupling subgroup of AA. The numerous AA pathways for which the complete set of reactions and biosynthetic genes have not yet been compiled and remain a significant challenge. However, phylogenetic analyses with candidates of analogous alkaloid pathways, such as the isoquinolinealkaloids, appear to be a logical and promising strategy.

A.4 Systems Biology and future directions

Until recently, non-model organisms such as medicinal plants were recalcitrant to modern molecular biology approaches for gene and pathway discovery due to the lack of genomic information and experimental protocols. However, the recent advances in 'omics' platforms and systems biology have revolutionized our understanding of natural products metabolism in non-model species (Schilmiller, Pichersky, & Last, 2012; Xiao et al., 2013) and correlations between plant transcriptome, proteome and metabolome have been successfully used for the identification of novel genes involved in alkaloid biosynthesis (Desgagne-Penix & Facchini, 2012; Desgagne-Penix, Farrow, Cram, Nowak, & Facchini, 2012; Desgagne-Penix et al., 2010; Liscombe, Ziegler, Schmidt, Ammer, & Facchini, 2009; Ziegler, Diaz-Chavez, Kramell, Ammer, & Kutchan, 2005; Ziegler et al., 2006). Each Amaryllidaceae species studied to date displays a specific AA profile, often with a few dominant compounds and a larger number of compounds at lower concentrations (Bastida et al., 2011; Berkov et al., 2014; Kornienko & Evidente, 2008). Although the molecular origin of this chemical diversity has not yet been clarified, these profiles likely result from differences in the expression level and substrate specificity of the various biosynthetic enzymes. The development and subsequent integration of 'omics' databases for Amaryllidaceae plants displaying different AA profiles would allow for the

identification of candidate genes involved in AA biosynthesis. For example, searches by sequence similarity to Arabidopsis, orthologous genes involved in synthesis of the precursors *3,4-DHBA* and tyramine, e.g. *PAL, Ca4H, Ca3H* and *TYDC*, can be identified. Those could be targeted for silencing or over-expression in plants to assess their role(s) in AA biosynthesis. Recently, the transcriptome of the Amaryllidaceae *Lycoris aurea* was sequenced and assembled and putative genes involved in AAs biosynthesis (*PAL, TYDC, OMT, NMT, P450*) were identified based on sequence analysis (Wang et al., 2013) however none of the biosynthetic enzymes involved in these pathways have been isolated or functionally characterized.

A.5 Conclusions

Over the past several years, extensive phytochemical and pharmacological analyses have reported the numerous biological activities of Amaryllidaceae alkaloids and numerous AAs display interesting and valuable pharmacological capabilities. Earlier biochemical tracer studies have led to the current AA biosynthetic proposed pathways. However, the molecular identity of the biosynthetic enzymes remains unknown. The technological advances brought as part of the post-genomics era have revolutionized the study of alkaloid metabolism and discoveries made over the past years have relied largely on 'omics' tools including transcriptome libraries, proteomic analyses and metabolomics methods. Integration of these 'omics' resources such as comparative analyses of metabolite and transcript data from Amaryllidaceae species will lead to the identification and isolation of numerous enzymes involved in AA biosynthesis. A deeper understanding of the molecular mechanisms involved in AA biosynthesis will support breeding efforts to produced cultivars of Amaryllidaceae species with enhanced AA production. In addition, it will pave the way for the successful metabolic engineering of microbial systems for the production of valuable AAs.

A.6 Acknowledgement

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada [Discovery] and the Universite du Quebec a Trois-Rivieres [Fonds Institutionnel de Recherche].

A.7 References

Barton, D. H. R., Kirby, G. W., Taylor, G. B., & Thomas, G. M. (1963). Phenol oxidation and biosynthesis. Part VI. The biogenesis of Amaryllidaceae alkaloids. Journal of the Chemical Society, 4545-4558. http://dx.doi.org/10.1039/jr9630004545

Barton, D.H.R., & Cohen, T. (1957). Some biogenetic aspects of phenol oxidation. Festschrift Prof. Dr. Arthur Stoll. Basel: Birkha¨user Verlag.

Bastida, J., Berkov, S., Torras, L., Pigni, N. B., de Andrade, J. P., Vanessa Martinez, V., ... Viladomat, F. (2011). Chemical andbiological aspects of Amaryllidaceae alkaloids. In RecentAdvances in Pharmaceutical Sciences. D. Muñoz-Torrero (Ed). Trivandrum, India: Transworld Research Network.

Battersby, A. R., Fales, H. M., & Wildman, W. C. (1961). Biosynthesis in the Amaryllidaceae. Tyrosine and norbelladine as precursors of haemanthamine. Journal of the American Chemical Society, 83, 4098 4099.http://dx.doi.org/10.1021/ja01480a037

Berkov, S., Bastida, J., Viladomat, F., & Codina, C. (2011). Development and validation of a GC-MS method for rapid determination of galanthamine in Leucojum aestivum and Narcissus ssp.: a metabolomic approach. Talanta, 83(5), 1455-1465. http://dx.doi.org/10.1016/j.talanta.2010.11.029

Berkov, S., Martinez-Frances, V., Bastida, J., Codina, C., & Rios, S. (2014). Evolution of alkaloid biosynthesis in the genus Narcissus. Phytochemistry, 99, 95-106 http://dx.doi.org/10.1016/j.phytochem.2013.11.002

Chu, H. Y., Wegel, E., & Osbourn, A. (2011). From hormones to secondary metabolism: the emergence of metabolic gene clusters in plants. Plant Journal, 66(1), 66-79. http://dx.doi.org/10.1111/j.1365-313X.2011.04503.x

Desgagne-Penix, I., & Facchini, P. J. (2011). Benzylisoquinoline Alkaloid Biosynthesis. In H. Ashihara, A. Crozier,& A. Komamine (Ed.), Plant Metabolism and Biotechnolog (pp. 241–261). Chichester, UK: John Wiley & Sons.

Desgagne-Penix, I., & Facchini, P. J. (2012). Systematic silencing of benzylisoquinoline alkaloid biosynthetic genes reveals the major route to papaverine in opium poppy. Plant Journal, 72(2), 331-344. http://dx.doi.org/10.1111/j.1365-313X.2012.05084.x

Desgagne-Penix, I., Farrow, S. C., Cram, D., Nowak, J., & Facchini, P. J. (2012). Integration of deep transcript and targeted metabolite profiles for eight cultivars of opium poppy. Plant Molecular Biology, 79(3), 295-313. http://dx.doi.org/10.1007/s11103-012-9913-2

Desgagne-Penix, I., Khan, M. F., Schriemer, D. C., Cram, D., Nowak, J., & Facchini, P. J. (2010). Integration of deep transcriptome and proteome analyses reveals the components of alkaloid metabolism in opium poppy cell cultures. BMC Plant Biology, 10, 252. http://dx.doi.org/10.1186/1471-2229-10-252

Dewick, P. M. (2009). Medicinal natural products: a biosynthetic apporach (3rd ed.). John Wiley and Sons Ltd. http://dx.doi.org/10.1002/9780470742761

Diaz Chavez, M. L., Rolf, M., Gesell, A., & Kutchan, T. M. (2011). Characterization of two methylenedioxy bridge-forming cytochrome P450-dependent enzymes of alkaloid formation in the Mexican prickly poppy Argemone mexicana. Archives of Biochemestry and Biophysics, 507(1), 186-193. http://dx.doi.org/10.1016/j.abb.2010.11.016

Eichhorn, J., Takada, T., Kita, Y., & Zenk, M. H. (1998). Biosynthesis of the amaryllidaceae alkaloid Galanthamine. Phytochemistry, 49(4), 1037-1047. http://dx.doi.org/10.1016/S0031-9422(97)01024-8

Facchini, P. J. (2001). ALKALOID BIOSYNTHESIS IN PLANTS: Biochemistry, Cell Biology, Molecular Regulation, and Metabolic Engineering Applications. Annual Review of Plant Physiology and Plant Molecular Biology, 52, 29-66. http://dx.doi.org/10.1146/annurev.arplant.52.1.29

Facchini, P. J., Huber-Allanach, K. L., & Tari, L. W. (2000). Plant aromatic L-amino acid decarboxylases: evolution, biochemistry, regulation, and metabolic engineering applications. Phytochemistry, 54(2), 121-138. http://dx.doi.org/10.1016/S0031-9422(00)00050-9

Gesell, A., Rolf, M., Ziegler, J., Diaz Chavez, M. L., Huang, F. C., & Kutchan, T. M. (2009). CYP719B1 is salutaridine synthase, the C-C phenol-coupling enzyme of morphine biosynthesis in opium poppy. Journal of Biological Chemistry, 284(36), 24432-24442. http://dx.doi.org/10.1074/jbc.M109.033373

Ghosal, S., Shanthy, A., & Singh, S. K. (1988). Isocraugsodine, an narylidenephenethylamine from Crinum asiaticum and its e-z isomerism. Phytochemistry, 27(6), 1849-1852.http://dx.doi.org/10.1016/0031-9422(88)80457-6

Grisebach, H. (1973). Comparative biosynthetic pathways in higher plants. Pure and Applied Chemistry, 34(3-4), 487-513.http://dx.doi.org/10.1351/pac197334030487

Hagel, J. M., & Facchini, P. J. (2013). Benzylisoquinoline Alkaloid Metabolism: A Century of Discovery and a Brave New World.Plant and Cell Physiology, 54(5), 647-672.http://dx.doi.org/10.1093/pcp/pct020

Heinrich, M., & Lee Teoh, H. (2004). Galanthamine from snowdrop--the development of a modern drug againstAlzheimer's disease from local Caucasian knowledge. Journal
of Ethnopharmacology, 92(2-3), 147-162.http://dx.doi.org/10.1016/j.jep.2004.02.012

Ikezawa, N., Iwasa, K., & Sato, F. (2008). Molecular cloning and characterization of CYP80G2, a cytochrome P450 that catalyzes an intramolecular C-C phenol coupling of(S)-reticuline in magnoflorine biosynthesis, from cultured Coptis japonica cells. Journal of Biological Chemistry, 283(14),8810-8821. http://dx.doi.org/10.1074/jbc.M705082200

Jin, Z. (2013). Amaryllidaceae and Sceletium alkaloids. Natural Product Reports, 30(6), 849-868.http://dx.doi.org/10.1039/c3np70005d

Kornienko, A., & Evidente, A. (2008). Chemistry, biology, and medicinal potential of narciclasine and its congeners. Chemical Reviews, 108(6), 1982-2014. http://dx.doi.org/10.1021/cr078198u

Lee, E. J., & Facchini, P. (2010). Norcoclaurine synthase is a member of the pathogenesis-related 10/Bet v1 protein family. Plant Cell, 22(10), 3489-3503.
http://dx.doi.org/10.1105/tpc.110.077958

Liscombe, D. K., Louie, G. V., & Noel, J. P. (2012). Architectures, mechanisms and molecular evolution of natural product methyltransferases. Natural Product Reports, 29(10), 1238-1250. http://dx.doi.org/10.1039/c2np20029e

Liscombe, D. K., Ziegler, J., Schmidt, J., Ammer, C., & Facchini, P. J. (2009). Targeted metabolite and transcript profiling for elucidating enzyme function: isolation of novel N-methyltransferases from three benzylisoquinoline alkaloid-producing species. Plant Journal, 60(4), 729-743.http://dx.doi.org/10.1111/j.1365-313X.2009.03980.x

Luk, L. Y., Bunn, S., Liscombe, D. K., Facchini, P. J., & Tanner, M. E. (2007). Mechanistic studies on norcoclaurine synthase of benzylisoquinoline alkaloid biosynthesis: an enzymatic Pictet-Spengler reaction. Biochemistry, 46(35), 10153-10161.
http://dx.doi.org/10.1021/bi700752n

Minami, H., Dubouzet, E., Iwasa, K., & Sato, F. (2007). Functional analysis of norcoclaurine synthase in Coptis japonica. Journal of Biological Chemistry, 282(9), 6274-6282. http://dx.doi.org/10.1074/jbc.M608933200

Mizutani, M., & Sato, F. (2011). Unusual P450 reactions in plant secondary metabolism. Archives of Biochemestry and Biophysics, 507(1), 194-203.
http://dx.doi.org/10.1016/j.abb.2010.09.026

Moll, S., Anke, S., Kahmann, U., Hansch, R., Hartmann, T., & Ober, D. (2002). Cell-specific expression of homospermidine synthase, the entry enzyme of the pyrrolizidine alkaloid pathway in Senecio vernalis, in comparison with its ancestor, deoxyhypusine synthase. Plant Physiology, 130(1), 47-57. http://dx.doi.org/10.1104/pp.004259

Nair, J. J., & van Staden, J. (2013). Pharmacological and toxicological insights to the South African Amaryllidaceae. Food Chemistry and Toxicology, 62C,262-275. http://dx.doi.org/10.1016/j.fct.2013.08.042

Nelson, D., & Werck-Reichhart, D. (2011). A P450-centric view of plant evolution. Plant Journal, 66(1), 194-211. http://dx.doi.org/10.1111/j.1365-313X.2011.04529.x

Samanani, N., Liscombe, D. K., & Facchini, P. J. (2004). Molecular cloning and characterization of norcoclaurine synthase, an enzyme catalyzing the first committed step in benzylisoquinoline alkaloid biosynthesis. Plant Journal, 40(2), 302-313. http://dx.doi.org/10.1111/j.1365-313X.2004.02210.x

Schilmiller, A. L., Pichersky, E., & Last, R. L. (2012). Taming the hydra of specialized metabolism: how systems biology and comparative approaches are revolutionizing plant biochemistry. Current Opinions in Plant Biology, 15(3), 338-344. http://dx.doi.org/10.1016/j.pbi.2011.12.005

Stockigt, J., Barleben, L., Panjikar, S., & Loris, E. A. (2008). 3D-Structure and function of strictosidine synthase--the key enzyme of monoterpenoid indole alkaloid biosynthesis. Plant Physiology and Biochemistry, 46(3), 340-355. http://dx.doi.org/10.1016/j.plaphy.2007.12.011

Takos, A. M., & Rook, F. (2013). Towards a molecular understanding of the biosynthesis of amaryllidaceae alkaloids in support of their expanding medical use. International Journal of Molecular Science, 14(6), 11713-11741. http://dx.doi.org/10.3390/ijms140611713

Wang, R., Xu, S., Jiang, Y., Jiang, J., Li, X., Liang, L., ... Xia, B. (2013). De novo sequence assembly and characterization of Lycoris aurea transcriptome using GS FLX titanium platform of 454 pyrosequencing. PLoS One, 8(4), e60449. http://dx.doi.org/10.1371/journal.pone.0060449

Xiao, M., Zhang, Y., Chen, X., Lee, E.-J., Barber, C. J. S., Chakrabarty, R., ... Sensen, C. W. (2013). Transcriptome analysis based on next-generation sequencing of non-model plants producing specialized metabolites of biotechnological interest. Journal of Biotechnology, 166(3), 122-134. http://dx.doi.org/10.1016/j.jbiotec.2013.04.004

Ziegler, J., & Facchini, P. J. (2008). Alkaloid biosynthesis: metabolism and trafficking. Annual Review of Plant Biology, 59, 735-769. http://dx.doi.org/10.1146/annurev.arplant.59.032607.092730

Ziegler, J., Diaz-Chavez, M. L., Kramell, R., Ammer, C., & Kutchan, T. M. (2005). Comparative macroarray analysis of morphine containing Papaver somniferum and eight morphine free Papaver species identifies an O-methyltransferase involved in benzylisoquinoline biosynthesis. Planta, 222(3), 458-471. http://dx.doi.org/10.1007/s00425-005-1550-4

Ziegler, J., Voigtlander, S., Schmidt, J., Kramell, R., Miersch, O., Ammer, C., ... Kutchan, T. M. (2006). Comparative transcript and alkaloid profiling in Papaver species identifies a short chain dehydrogenase/reductase involved in morphine biosynthesis. Plant Journal, 48(2), 177-192. http://dx.doi.org/10.1111/j.1365-313X.2006.02860.x

The following study, published in a book '*Alkaloids: Biosynthesis, Biological Roles and Health benefits*', highlights the medicinal properties, biosynthesis, challenges, and advancement in new technologies for the study of Amaryllidaceae alkaloids.

ANNEX B

BIOSYNTHESIS OF AMARYLLIDACEAE ALKALOID: A BIOCHEMICAL OUTLOOK

APARNA SINGH AND ISABEL DESGAGNE PENIX

Published on 16th of April 2015 in *Alkaloids: Biosynthesis, Biological Roles and Health Benefits*

B.1 Abstract

Amaryllidaceae alkaloids (Grabherr, Haas et al.) are a diverse group of biologically active specialized metabolites produced mainly in Amaryllidaceae plant family. Several AAs possess potent pharmaceutical properties making them interesting target for drug development. For example, the AA galanthanmine, an acetylcholinesterase inhibitor, is used to treat neurodegenerative disorder including Alzheimer's disease. Also, AAs such as lycorine possess anti-microbial activity, whereas others such as crinine and narciclasine are potentially anti-cancer agents. Ironically, more is understood about the effects of alkaloids on humans than on their biosynthesis or their roles in plants. Understanding the biochemical genetics underpinning AA biosynthesis could enable improved production of these important pharmaceutical both in plants and in other systems.

Keywords: Amaryllidaceae alkaloid; plant secondary metabolism; alkaloid biosynthesis; natural products; systems biology; galanthamine

Groupe de Recherche en Biologie Végétale
Department of chemistry, biochemistry and physics
3351 Boulevard des Forges, 2403-LP
Trois-Rivières, Québec, Canada G9A 5H7
Email: Isabel.Desgagne-Penix@uqtr.ca

B.2 Introduction

Alkaloids are low-molecular weight, nitrogenous specialized metabolites occurring in approximately 20% of plant species. Many of the about 12,000 structurally elucidated alkaloids show potent biological activity [1, 2]. Particularly, the pharmacological properties of Amaryllidaceae alkaloids (Grabherr, Haas et al.) have been exploited for centuries, and modern medicine continues to rely on plant-derived compounds such as acetylcholinesterase inhibitor galanthamine used to treat neurodegenerative disorder including Alzheimer's disease [3-7]. AAs such as lycorine possess anti-microbial activity, whereas others such as crinine and narciclasine are potentially anti-cancer agents. Ironically, more is understood about the effects of alkaloids on humans than on the roles of these compounds in the plants that produce them. Although not considered essential for normal growth and development, AAs likely play key roles in plant defense as several AAs possess potent pharmacological activities (Table 1), which in turn may be a good indication of their biological function. For example, the potency of crinine as a hallucinogenic or haemanthamine as an analgesic suggest their role as herbivores deterrents. Similarly, the anti-microbial properties of lycorine and cytotoxicity of narciclasine suggest that they may confer protection against pathogens.

The bulbous monocot Amaryllidaceae (*Amaryllis*, *Narcissus, Galanthus,* etc.), comprising over 1100 species, are among the top 20 of the most widely used plant families as traditional source of medicines [6, 7]. For example, in India, *Crinum bulbispermum* has been used for treating rheumatic disorder whereas in Nigeria, bulbs of *Crinum jagus* and *C. glaucum* has been traditionally recommended for various mental illness. The presence of AAs such as crinine and vittatine in *Crinum* bulbs (Table 1), with potency to inhibit acetylcholinesterase has been documented, extending their therapeutic usage towards Alzheimer's disease [6]. In Egypt, mixture from roots of *Amaryllis belladonna* and *Clivia miniata* were used to treat and prevent infection of wounds such as insect and snake bites. The anti-microbial activities are achieved by virtue of certain types of AAs namely lycorine, homolycorine, haemanthamine and tazettine present in the root mixture [10, 15, 19].

AA	Group	Plant source	Biological activity	Ref.
Galanthamine	*para-ortho'*	*Narcissus pseudonarcissus* *Leucojum aestivum* *Galanthus niveali* *Galanthus elewesii* *Pancratium maritimum*	acetylcholinesterase inhibitor, treatment for Alzheimer's disease	[4, 8-10]
Narciclasine	*para-para'*	*Narcissus spp.* *Leucojum aestivum* *Lycoris radiata,* *Pancratium maritimum* *Hymenocallis expansa*	cytotoxic	[3, 4]
Montanine	*papa-para'*	*Hippeastrum vittatum*	anxiolytic, antidepressant	[4, 11]
Pretazettine	*para-para'*	*Narcissus spp.*	Anti-viral, cytotoxic, analgesic	[12]
Crinine	*para-para'*	*Crinum spp.* *Galanthus reginae–olgae* *Brunsvigia gregaria* *Pancratinum maritimum* *Boophone disticha*	cytotoxic, hallucinogenic	[3, 4, 9, 13]
Haemanthamine	*para-para'*	*Narcissus pseudonarcissus* *Narcissus primigenius* *Pancratinum maritimum* *Clivia miniata*	cytotoxic, analgesic, anti-malarial, anti-retroviral, fungicide hypotensive	[4, 13-15]
Galanthine	*ortho-para'*	*Galanthus elewesii* *Lycoris sunguinea*	cytotoxic	[16, 17]
Narcissidine	*ortho-para'*	*Leucojium autumnale*	cytotoxic	[4]
Lycorine	*ortho-para'*	*Narcissus spp.*	Anti-viral	[3, 4,]

Table 1. Examples of important AA, plant source and biological activity.

Following on the commercial success of galanthamine as a prescription drug in the treatment of Alzheimer's disease, much work has focus on the discovery and production of derivatives as drug candidates. For example, the improved ability of galanthamine derivatives to cross blood-brain barrier coupled with stronger analgesic and analeptic properties make them interesting pharmaceutical compounds for the treatment of numerous neurological conditions such as paralysis and schizophrenia [20-23]. Given the interesting biological properties of most AAs, Amaryllidaceae plants provide a diverse and accessible platform for phytochemical-based drug discovery. For example, lycorine- and crinine-types of AAs have shown much promise as remarkably potent and selective anti-cancer agents [24]. As it is the cases for many valuable plant natural products, the commercial development of most AAs is restricted by their limited availabilities due to their low concentrations in plants. Currently, among the ~500 known AAs, only galanthamine is used clinically and is majoritarily obtained from *Leucojum* and *Narcissus* plant sources [5, 7]. Although, several complex chemical syntheses and semi-syntheses have been developed successfully, they are not commercially viable for the pharmaceutical industry [3]. Compared with the large knowledge for the potential pharmaceutical activities of AAs, the biochemistry of AA biosynthesis remains elusive. However, an impressive array of biosynthetic enzymes must function together to yield the vast diversity of AAs. Understanding the biochemical genetics underpinning AA biosynthesis could enable improved production of these important pharmaceutical and related molecules both in plants and in other systems.

B.3 Biosynthesis

In contrast to the extensive literature on the biological effects of AAs, information on their biochemical pathways and molecular genetics is incomplete. The majority of the research done on AA biosynthesis is supported by early feeding experiments using radiolabeled precursors. These studies led to the biochemical elucidation of the initial steps in AA biosynthesis [25-31]. Despite their vast structural diversity, AAs share a common biosynthetic origin, norbelladine, which is formed through the condensation of amino acids derivatives; 3,4-dihydroxybenzaldehyde (3,4-DHBA also named

protocatechuic aldehyde) and tyramine. The resulting norbelladine is central to the biosynthesis of many structural types of AAs (Figure 1).

Currently, there is no consensus about the names or number of classes of AAs however it may be influenced by the alkaloids present in the particular species. As the number of AA discovered increases, new core structures are discovered. Evidente and Kornienko, recognized twelve distinct ring types [3] of AAs whereas Jin [6] distinguished eighteen subgroups and this expansion is mainly due to the continued isolation of AAs with rare skeleton types. Bastida *et al.* recognized nine subgroups but mostly focused on *Narcissus spp.* [4]. A classification based on the biosynthetic origins of the different AAs as three groups (*para-ortho'*, *para-para'*, *ortho-para'*) and nine AA-types (Table 2) is depicted in this chapter in order to discuss the biosynthesis of AAs (Figure 2).

	Group	**AA-type**
1		Norbelladine
2	*para-ortho'*	Galanthamine
3	*para-para'*	Crinine
4	*para-para'*	Narciclasine
5	*para-para'*	Haemanthamine
6	*para-para'*	Montanine
7	*ortho-para'*	Homolycorine
8	*ortho-para'*	Lycorine
9	*ortho-para'*	Narcissidine

Table 2. Proposed classification based on biosynthetic origin.

B.3.1 Initial biosynthetic reactions

Despite the enormous variety of plant specialized metabolites, the number of corresponding basic biosynthetic pathways is restricted and distinct. The precursor substrates and enzymes necessary for the first committed steps often appear to have been recruited from primary metabolic pathways, such as glycolysis, the Krebs cycle, the pentose phosphate pathway and the shikimate pathway [32]. For example, the aromatic amino acids L-phenylalanine and L-tyrosine, produced by the shikimate pathway, are

precursors for a wide spectrum of natural products including phenylpropanoids, flavonoids, lignins, coumarins, cyanogenic glycosides, glucosinolates, and alkaloids [33].

AA biosynthesis starts with the formation of two precursors units: 1) the C6-C1 unit from L-phenylalanine named 3,4-DHBA and 2) the C6-C2-N unit from L-tyrosine called tyramine (Figure 3). The sequence of reactions leading to the conversion of L-phe into 3,4-DHBA are not known but studies suggested that 3,4-DHBA originates from phenylpropanoid C6-C3 compounds as in the formation of the fragrant phenolic compound, vanillin, in *Vanilla planifolia* plants [34-38]. Interestingly, vanillin is a metabolic intermediate involved in the biosynthesis of capsaicinoids which are benzylamine alkaloids responsible of the pungent flavor in chili pepper *Capsicum spp.* [39]. Altogether, these results suggest that initial reactions and biosynthetic enzymes of the phenylpropanoid pathway are common in the synthesis of AA precursor 3,4-DHBA. The first reaction of phenylpropanoid biosynthesis is catalyzed by the L-phenylalanine ammonia-lyase (PAL), a key regulatory enzyme promoting the deamination of L-phe to generate *trans*-cinnamic acid (Figure 3A). PAL reaction not only controls carbon flux into this pathway but also links specialized metabolism to primary metabolism leading to the production of a variety of phenolic compounds including cell wall structural polymer lignin [38]. This enzyme is among the most studied in plant specialized metabolism and many *PAL* genes have been cloned and characterized from various species including Amaryllidaceae species [40, 41]. Recently, *Lycoris radiata PAL1* gene was isolated, characterized and a correlation between *LrPAL1* expression and content of galanthamine was reported supporting its role in AA biosynthesis [40]. The hydroxylation of the benzyl ring of *trans*-cinnamic acid is catalyzed by the cytochrome P450-dependent monooxygenase (CYP) cinnamate-4-hydroxylase (C4H) which belongs to the CYP73 subfamily, and results in the formation of *p*-coumaric acid (Figure 3A) [34, 37, 42]. CYPs participate in a variety of biochemical pathways and the number of *CYP* genes in plant genomes is estimated to be up to 1%, implying that plants are huge sources for various CYP-dependent reactions [43]. CYP catalyze a variety of monooxygenation/hydroxylation reactions along with some unusal reactions such as methyldioxy bridge formation, phenol coupling reactions and oxidative rearrangement of

carbon skeletons. cDNA sequences encoding *C4H* have been identified in numerous plants and were shown to be highly conserved across species [42, 44]. Additionally, *C4H* sequences were present in the transcriptome database of the Amaryllidaceae *L. radiata* [41].

From this stage, as it is the case for vanillin biosynthesis, two separate routes are proposed as to how *p*-coumaric acid is converted to 3,4-DHBA: the oxidative 'ferulate' and the non-oxidative 'benzoate' pathways (Figure 3A). The ferulate pathway, proposed by Zenk in 1965 [34], suggested that the aromatic ring of *p*-coumaric acid undergoes hydroxylation by the coumarate-3-hydroxylase (C3H), giving rise to caffeic acid. The later undergo chain shortening (loss of two carbon atoms) to form the precursor 3,4-DHBA, via a mechanism similar to the β-oxidation sequence for the catabolic degradation of fatty acids, [4, 28, 31]. The other route, the benzoate pathway, proposes a non-oxidative chain shortening of *p*-coumaric acid first, followed by the hydroxylation of the aromatic ring to yield 3,4-DHBA. Recently, an unusual cysteine protease named 4-hydroxybenzaldehyde synthase (HBS) was shown to catalyze the conversion of *p*-coumaric acid to *p*-hydroxybenzaldehyde. However, the role of HBS in 3-4-DHBA and vanillin biosynthesis remains under investigation [36, 45].

The other precursor of AA biosynthesis, the C6-C2-N unit, originates from the decarboxylation of L-tyr to tyramine by the L-tyrosine decarboxylase (TYDC) [4, 5, 28, 31]. TYDC belongs to the DOPA decarboxylase family of the pyridoxal phosphate (PLP)-dependent aspartate aminotransferase superfamily. In addition to controlling the transition from primary to specialized metabolism, TYDC ensures an adequate supply of tyramine for the synthesis of various alkaloids including the well-known benzylisoquinoline alkaloids (Delano-Frier, Aviles-Arnaut et al.) such as the narcotic analgesics morphine and codeine [1, 2, 46-48]. Several molecular clones for TYDC have been isolated from multiple organisms (*e.g.* microorganisms, fungi, plants, animals) [46]. Plant TYDCs share great sequence similarity to each other within and across species [46, 49]. Transcriptome analysis of of *L. aurea* [41] have confirmed the presence of *TYDC* sequences in Amaryllidaceae.

Although, the enzymatic reactions leading to the formation of precursors are still hypothetical, biochemical evidences from biotransformation studies with labeled precursors [25-31] coupled with molecular evidences from recent transcriptome analyses [40, 41] have confirmed the presence of metabolite intermediates and transcripts for *PAL, C4H, C3H* and *TYDC* in Amaryllidaceae species supporting, in parts, the current proposed pathways.

B.3.2 Norbelladine-type alkaloids

The first committed step in AA biosynthesis in plants starts with the coupling of the two precursors, 3,4-DHBA and tyramine, defining the entry point of primary metabolites into AA biosynthetic pathway. The condensation of the aldehyde (3,4-DHBA) and the amine (tyramine), called Pictet-Spengler condensation, results in a Schiff's base intermediate which following reduction yields norbelladine (Figure 4) [50, 51]. Pictet-Spengler reactions are widely used in plant alkaloid biosynthesis to yield either a β–carboline or a tetrahydroquinoline product from the condensation of an aldehyde and an aromatic amine [52]. For example, the first committed step in the formation of BIAs is the combination of two L-tyr derivatives, namely 4-hydroxyphehylacetaldehyde and dopamine, by a Pictet-Spenglerase, the norcoclaurine synthase (NCS), to produce the trihydroxyisoquinoline alkaloid norcoclaurine [53-59]. NCS catalyzes a two-step reaction mechanism; a condensation followed by an intramolecular cyclization [53, 57]. The electron-donating oxygen of the hydroxyl group of dopamine (missing in tyramine) is essential for the reaction to proceed and may provide a mechanistic explanation for the absence of cyclization in the condensation step of AA biosynthesis. To date, two different protein families (pathogenesis-related PR10-Betv1 and 2-oxoglutarate-dependant dioxygenase) have been reported to catalyze this reaction *in vitro*. However, the PR10/Betv1-NCS was shown to be the major one involved *in vivo* in BIA biosynthesis of *Papaver somniferum* and *Thalictrum flavum* plants [54]. Another example of a Pictet-Spengler condensation is the combination of tryptamine and secologanin to form strictosidine (β-carboline product), the general precursor to terpene indole alkaloids [60,

61]. In *Rauvolfia serpentina*, this reaction is catalyzed by strictosidine synthase, a member of the six-bladed four-stranded β-propeller fold protein family [61]. These examples show that the initial Pictet-Spengler reactions in alkaloid biosynthesis can be catalyzed by members of very different protein families and suggests that the type of enzyme recruited for the non-cyclizing condensation reaction in AA biosynthesis may belong to one of these families or to an additional class of enzymes.

In Amaryllidaceae plants, an *O*-methylation of norbelladine on the '3,4-DHBA' aromatic ring takes place prior to intramolecular oxidative coupling which results in the formation of different skeletons of AAs. Thus, the key biosynthetic intermediate of AAs is 4'-*O*- norbelladine (Figure 4). Plant *O*-methylation reactions are common transformation in the biosynthesis of alkaloids and are most often catalyzed by *S*-adenosyl-L-methionine (SAM)-dependent methyltransferases (MTs) [62-72]. Thus, norbelladine must be 4'-*O*-methylated to form 4'-*O*-methylbelladine, a central intermediate from which multiple biosynthetic pathways lead to various structural types of AAs (Figures 1-2). Mann *et al.* isolated a catechol-*O*-methyltransferase from the Amaryllidaceae *Nerine bowdenii* bulbs that catalyzed the highly regiospecific methylation of norbelladine at the *para* position at a yield of more than 90% but it also methylated a wide range of catechol substrates [73]. Additionally, protein extracts from leaves of *L. vernum* showed specific activity with regiospecificities favoring the 4' position of norbelladine over the 3' position in a 18:1 ratio [28] and transcript sequences for *OMTs* were detected in the transcriptome of *L. radiata* [41].

B.3.3 Phenol coupling reaction and formation of diverse AA-type backbones

Many specialized metabolites are produced by the coupling of two or more phenolic rings. The reactions involve hydrogen abstraction from a phenol giving the radical followed by the delocalization of the unpaired electron via resonance forms in which the free electron is dispersed to positions *ortho* and *para* to the original oxygen

function. These phenol-derived radicals are quenched by coupling with other radicals and the coupling of two of these resonance structures, in various combinations, gives a range of dimeric systems. Thus, C-C bonds involving positions *ortho* or *para* to the original phenols may be formed. The reactive dienone systems formed as intermediates may in some cases be rearranged and further modified extending the range or structures ultimately derived from the basic reaction.

A crucial step in AA biosynthesis is the cyclization of 4'-*O*-methylbelladine by three different ways of intramolecular C-C oxidative phenol coupling named *para-ortho'*(*p-o'*), *para-para'* (*p-p'*) and *ortho-para'* (*o-p'*) and these alternative phenol coupling reactions generate three backbone structures (Figure 2). The different AA subgroups are referred to as *p-o'*, *p-p'* and *o-p'* are derived from these different backbones (Figure 1; Tables 1-2). For instance, anti-tumor AAs of the narciclasine-type are derived by a *p-p'* phenol coupling step whereas galanthamine originates from the *p-o'* coupling (Figure 2) [3-5].

Recently, several enzymes discovered to be responsible for the formation of intramolecular C-C phenol-couples in various alkaloids pathways, were found to belong to the plant CYP superfamily [43]. For example, the intramolecular C-C phenol-coupling catalyzed by CYP80G2 participates in the formation of aporphine-type alkaloid in *Coptis japonica* [74]. In the formation of morphinan alkaloids from *P.somniferum*, the enzyme salutaridine synthase, a member of the CYP719B1, catalyzed a pivotal bridge-forming reaction through the C-C phenol coupling of (*R*)-reticuline into salutaridine, the *p-o'*-coupled of morphine [75].

Altogether, these examples indicate that specific CYP are involved in the C-C phenol coupling reaction in alkaloid biosynthesis. Specifically, members of two CYP families, CYP80 and CYP719, have been identified. A recent comparative analysis of plant CYP sequences has shown that only the CYP80 family occurs in monocot plant species however no members of the Amaryllidaceae family were represented in this study [44]. This suggests that the phenol coupling steps in AA biosynthesis is likely to involve

CYP and may be of the CYP80 family. Interestingly, AAs from all three groups of phenol coupling can co-occur in a single plant, whereas some species and cultivars contains only one group of AA [3, 76] suggesting that multiples genes and/or different isoforms for the phenol coupling enzyme exist. Biotransformation of deuterated-4'-*O*-methylnorbelladine in tissue cultures of *L. aestivum* showed incorporation in all three groups of AAs further supporting the three modes of intramolecular oxidative phenol coupling for cyclization in this Amaryllidaceae.

The three core skeletons obtained from the phenol coupling steps form the basis of further alkaloid diversity. A complex network of enzymatic reactions exists to produce a spectrum of compounds that differs between species, varieties and cultivars and even between the different tissues and vegetative phases of the same plant. These biochemical modifications are achieved by a multitude of enzymes catalyzing various types of reactions, such as C-C and C-O bond formations, *O*- and *N*-methylations, demethylations, hydroxylations, oxidations and reductions. The various products obtained from these reactions yields the several hundred of structurally related AAs known to date (Table 1, Figures 1-2) [3-6, 13].

B.3.4 Alkaloids derived from the *para-ortho'*-coupling

The *p-o'* phenol coupling reaction of 4'-*O*-methylnorbelladine leads to the formation of galanthamine-type of AAs (Figure 2). Feeding experiments using 13C-labelled and deuterium-labelled *O*-methylnorbelladine to *L. aestivum* plants showed that the biosynthesis of galanthanmine involves the phenol oxidative coupling of *O*-methylnorbelladine to a dienone intermediate, which undergoes spontaneous closure of the ether bridge to yield *N*-demethylnarwedine (Figure 5) [4, 28-30]. The stereo-specific reduction of *N*-demethylnarwedine yields norgalanthanmine which is *N*-methylated to form galanthamine (Figure 5) [28-30]. Feeding studies also showed that narwedine is not the direct precursor of galanthanmine, and possibly exist in equilibrium with galanthamine [28], a reaction catalyzed by a hypothetically reversible oxido-reductase. Norgalanthamnine can also be further reduced to yield norlycoramine which is *N*-

methylated into lycoramine (Figure 5). Chlidanthine might arise from galanthamine by *O*-demethylation, a reaction confirmed with 3H-labeled incorporation into chlidanthine [77, 78].

Recently, opium poppy (*P. somniferum*) has emerged as a model system to investigate alkaloid metabolism. Great progress towards the elucidation of BIA metabolism including the well-known morphinan alkaloids has been made at the biochemical and molecular levels and a restricted number of enzyme families have been implicated. Enzymes of the lyase, oxidoreductase and transferase categories have been characterized from several BIA-producing plants and reviewed [2, 47, 48]. The analogy between morphine and galanthamine biosyntheses is quite stricking and is highlighted in Figure 6. The similarities between reactions suggested that AA biosynthetic enzymes may belong to similar gene families. As described above, the Pictet-Spengler condensation of precursor 3,4-DHBA and tyramine, into norbelladine resembles the condensation of 4-HPAA and dopamine to form norcoclaurine (Figure 6). Similarly, the 4'-*O*-methylation catalyzed by the 3'-hydroxy-*N*-methylcoclaurine-4'-*O*-methyltransferase (4'OMT) and the phenol coupling reaction catalyzed by SalSyn yielding salutaridine mirror the *O*-methylation and phenol coupling reaction in galanthamine biosynthesis (Figure 6). Additionally, the reduction reaction of *N*-demethylnarwedine is similar to the reduction of salutaridine (Figure 6). Reduction reactions are common in alkaloid metabolism and two major families have been reported to be involved in morphinan alkaloid biosynthesis, the short chain dehydrogenases/reductases (SDRs) and the aldo-keto reductases (AKRs) [2, 47, 48]. For example, the stereo-specific reduction of salutaridine to form salutaridinol is catalyzed by the cytosolic NADPH-dependent salutaridine reductase (SalR) which belongs to the SDR family whereas the reduction of codeinone or morphinone by the NADPH-dependent codeinone reductase (CoR), member of the AKRs, yields codeine and morphine respectively [79-81]. Recently, the last step in noscapine biosynthesis, a phtalideisoquinoline alkaloid from *P. somniferum* with anti-cancer properties, as shown to be catalyzed by a reductase member of the SDRs [82]. From this step, norgalanthanmine is *N*-methylated to galanthamine whereas the *N*-methylation reaction in BIA biosynthesis occurs at an earlier step, *i.e.* prior to 4'-*O*-methylation, and is catalyzed by the coclaurine

N-methyltransferase (Figure 6) [1, 2, 47, 48]. In some Amaryllidaceae plants such as in *Harmanthus multiflorus*, galanthamine is *O*-demethylated to form *O*-demethylgalanthamine. This reaction is analogous to the antepenultimate and final steps in morphine biosynthesis catalyzed by the 2-oxoglutarate/Fe(II)-dependent dioxygenases, thebaine 6-*O*-demethylase (T6ODM) and codeine *O*-demethylase (CODM) (Figure 6) [83-85]. Since the comparison of morphine and galanthamine biosynthetic pathways showed great similarities, it may be a strategy to identify interesting biosynthetic gene candidates for the elucidation of AA biosynthesis.

B.3.5 Alkaloids derived from the *para-para'*-subgroup

Great structural diversity arises from the *p-p'* phenol coupling of 4'-*O*-methylnorbelladine including crinine-, haemanthamine-, narciclasine- and montanine-type of AAs (Figure 2). Experiments using labelled precursors suggested that the oxidation of *p-p'* core skeleton promoting the formation of methylenedioxy bridge leads to the formation of two possible intermediates, crinine and vittatine (Figure 7) [4]. Different biosynthetic routes from the vittatine intermediate are possible. For example, vittatine (and crinine) may be converted into narciclasine via several oxidation reactions whereas one *O*-methylation of vittatine yields montanine (Figure 7). Hydroxylation of vitattine yields 11-hydroxyvittatine which is subsequently reduced into pancracine and *O*-methylated to form montanine (Figure 7). Alternatively, *O*-methylation of 11-hydroxyvittatine gives haemanthamine which can be hydroxylated and *N*-methylated to yield pretazettine (Figure 7). Enzymes from the methyltransferase families (OMTs and NMTs) are likely to catalyze the methylation reactions.

The formation of C-C or C-O bonds to establish the various AA backbone structure is catalyzed by oxidoreductases such as CYPs and FAD-dependent enzymes. For example, the methyledioxy bridge formation of the *p-p'* intermediate leading to the formation of crinine and vittatine may be catalyzed by enzymes of the CYP superfamily (Figure 7). For example, members of the CYP719A subfamily were reported to catalyze methylenedioxy bridge formation in the biosynthesis of alkaloids such as berberine and sanguinarine

including *C. japonica* CYP719A1, *E. californica* CYP719A2 and CYP719A3, and *Argemone mexicana* CYP719A13 and CYP719A14 [86-88]. This suggests that members of the CYP719A subfamily are likely to catalyze oxidation reactions in the biosynthesis of AAs. However, other enzyme family are also possible including other CYP families or FAD-dependent oxidoreductases such as berberine bridge enzyme (BBE), (*S*)-tetrahydroprotoberberine oxidase (STOX) and dihydrobenzophenanthridine oxidase (DBOX) which are known to be involved in alkaloid metabolism [2, 89, 90].

B.3.6 Alkaloids derived from the *ortho-para'*-subgroup

The *o-p'* phenolic coupling of 4'-*O*-methylnorbelladine results in the formation of the key branchpoint intermediate norpluviine, the precursor of lycorine-, narcissidine- and homolycorine-type of AAs. Feeding of isotopically labeled precursors to different cultivars of *Narcissus* including 'King Alfred', Sempre Avanti', 'Deanna Durbin' and 'Texas' revealed that norpluviine may be processed by three possible route to form the different AA-types [4]. In the first route, norорpluviine is *O*-methylated to pluviine which is hydroxylated to 9-*O*-methylpseudolycorine and then *O*-methylated to galanthine (Figure 8). Galantine may be further converted via oxidation-reduction reactions to form narcissidine. In another route, norpluviine is converted to lycorine via caranine with steps involing oxidation (methylenedioxy bridge formation) and hydroxylation of the benzyl ring (Figure 8). Alternatively, norpluviine may undergoes hydroxylation (benzylic oxidation at position 6), ring opening, rotation and closure, followed by subsequent *O*- and *N*-methylations to form lycorenine, which after oxidation yields homolycorine (Figure 8). The *O*- and *N*-methylation reactions involved in these pathways are likely to be catalyzed by similar enzyme families (OMTs, NMTs) than those discussed earlier (Figure 6). Similarly, the reduction reactions may be catatalyzed by enzyme of the SDR and the AKR families.

C-C or C-O bond formation is catalyzed by oxidoreductases such as CYPs and FAD-dependent enzymes. For example, the hydroxylation of norpluviine leads to the formation of a C-O bond which disturbed the electronic state of the hydroxynorpluviine

rings subsequently promoting ring opening, rotation and rearrangement of the core structure (Figure 8). The hydroxylation reactions in alkaloid biosynthesis are catalyzed by CYP enzymes. Characterized members of the CYP80 family including berbamunine synthase (*Berberis stolonifera* CYP80A1) and (*S*)-*N*-methylcolcaurine 3'-hydroxylase (*Eschscholzia californica* CYP80B1, *P. somniferum* CYP80B3 and *C. japonica* CYP80G2) [2, 74]. In addition to hydroxylation, members of the CYP80 family catalyze intermolecular C-O and intramolecular C-C phenol coupling reaction. Recently, members of another CYP family were reported to be involved in hydroxylation reactions of alkaloid biosynthesis. Specifically, the characterization of the *C. japonica* CYP82N2v2 and *P. somniferum* CYP82 showed hydroxylase activity in the last steps of sanguinarine biosynthesis [91, 92]. The methylenedioxy bridge formation from norpluviine to caranine may also be catalyzed by CYP enzymes specifically the CYP719A subfamily as discussed earlier (Figures 7-8). Altogether, these results showed that members of the CYP80 and CYP82 are involved in hydroxylation reactions whereas CYP719A members promoted methylenebridge formation in alkaloid biosynthetic pathways. This further supports the involvement of members of these CYP families in the AA biosynthesis. Finally, despite the vast structural diversity of AAs, a restricted number of enzyme families seems to be involved in AA biosynthesis. Additionally, some enzymes may exhibit a relatively broad substrate range whereas others are may be highly substrate specific further participating in the formation of a plethora of AAs.

B.4 Metabolic engineering and synthetic biology

The commercial demand to increase certain AA in plants is very difficult to achieve by conventional breeding techniques and nowadays technologies may help to reach these goals. Until recently, non-model organisms such as Amaryllidaceae plants were recalcitrant to modern molecular biology approaches for gene and pathway discovery due in part to the lack of transformation protocols. However, rapid progress in the development of integrative approaches such as genomics, transcriptomics, proteomics, metabolomics and fluxomics, has provided essential information for the understanding of many complex biological processes at different levels [93]. Comparative and integrative analyses of such 'omics' databases has facilitated the determination of gene function and

pinpoint rate limiting steps and factors in the biosynthetic pathway of natural products in non-model species [93, 94]. Correlations between plant transcriptome, proteome and metabolome have been successfully used for the identification of novel genes involved in alkaloid biosynthesis [69, 70, 79, 95-97].

Each Amaryllidaceae species studied to date displays specific alkaloid composition (AA types and levels), often with a few dominant AAs and a larger number at lower concentrations [3, 4, 76]. Although the molecular origin of this biochemical diversity has not yet been defined, these AA compositions likely result from differences in the expression level and substrate specificity of the various biosynthetic enzymes. The development and subsequent integration of 'omics' databases for Amaryllidaceae plants displaying different AA composition would allow for the identification of candidate genes involved in AA biosynthesis. For example, searches by sequence similarity to *P. somniferum*, of orthologous genes involved in synthesis of the precursors 3,4-DHBA and tyramine, *e.g. PAL, C4H, C3H, HBS* and *TYDC,* can be identified. Recently, the transcriptome of the Amaryllidaceae *L. aurea* was sequenced, assembled and putative genes involved in AAs biosynthesis (*PAL, TYDC, OMT, NMT, CYP*) were identified based on sequence analysis [41] however none of the biosynthetic enzymes involved in these pathways have been isolated or functionally characterized. Those could be targeted for silencing or over-expression study in plants to confirm their role(s) in AA biosynthesis and to alter AA composition. Additionally, information obtained would help in designing rational plant breeding methods to raise new cultivars with optimized alkaloid production beneficial for pharmaceutical industries and other research applications. For example, acetylcholinesterase inhibitor galanthamine and anticancerous agent narciclasine and crinine, could be produced commercially by such newly designed plant breeding programs.

Currently, large-scale production of AA-based pharmaceuticals is limited to extraction of alkaloids that accumulate in plants. However the synthesis of AAs in microbial hosts could bypass such limitations and transform industrial production of AAs (Figure 9). Thus, a precise knowledge of AA biosynthetic networks is mandatory to the

emerging field of synthetic biology, which aims to reconstruct plant pathways in microorganisms with the ultimate goal of creating alternative systems for the production of valuable bioproducts. To date, several strategies have been developed to engineered microbial hosts for the production of alkaloids specifically related to the BIAs [98-102]. Recently, the introduction of ten BIA biosynthetic genes in yeast led to the production of dihydrosanguinarine and its oxidized derivative sanguinarine from (*R,S*)-norlaudanosoline further demonstrating the feasibility of production of high-value alkaloids in microbial systems [102].

B.5 Conclusion

Over the past several years, extensive phytochemical and pharmacological analyses have reported the numerous biological activities of Amaryllidaceae alkaloids and numerous AAs display interesting and valuable pharmacological capabilities. Earlier biochemical tracer studies have led to the current AA biosynthetic proposed pathways however the molecular identity of the biosynthetic enzymes remains unknown. The technological advances brought as part of the post-genomics era have revolutionized the study of alkaloid metabolism and discoveries made over the past years have relied largely on 'omics' tools including transcriptome libraries, proteomic analyses and targeted metabolomics methods. Integration of these 'omics' resources such as comparative analyses of metabolite and transcript data from Amaryllidaceae species will lead to the identification and isolation of numerous enzymes involved in AA biosynthesis. A deeper understanding of the molecular mechanisms involved in AA biosynthesis will support breeding efforts to produced cultivars of Amaryllidaceae species with enhanced AA production. In addition, it will pave the way for the successful metabolic engineering of microbial systems for the production of valuable AAs.

B.6 List of figures

6OMT, norcoclaurine 6-*O*-methyltransferase; CNMT, coclaurine *N*-methyltransferase; NMCH, *N*-methylcoclaurine 3'-hydroxylase; 4'OMT, 3'-hydroxyl-*N*-methylcoclaurine 4'-*O*methyltransferase; SalSyn, salutaridine synthase; SalR, salutaridine reductase; SalAT, salutaridinol 7-*O*-acetyltransferase; T6ODM, thebaine 6-O-demethylase; CoR, codeinone reductase; CODM, codeine *O*-demethylase; OMT, *O*-methyltransferase; CYP, cytochrome P450; R, reductase; NMT, *N*-methyltransferase; ODM, *O*-demethylase.

Figure 7. Biosynthesis of *p-p'* alkaloids such as crinine, haemanthamine, narciclasine and montanine. Broken arrow represents more than one biochemical reaction. Abbreviations: R, reduction; O, oxidation; NMT, *N*-methylation; OMT, *O*-methylation.

Figure 8. Biosynthesis of *o-p'* alkaloids such as lycorine, homolycorine and narcissidine. Broken arrow represents more than one biochemical reaction. Abbreviations: R, reduction; O, oxidation; NMT, *N*-methylation; OMT, *O*-methylation.

Figure 9. Representation of the strategy for the metabolic engineering of microorganism for AA drug production

B.7 References

Facchini, P.J., *ALKALOID BIOSYNTHESIS IN PLANTS: Biochemistry, Cell Biology, Molecular Regulation, and Metabolic Engineering Applications.* Annu Rev Plant Physiol Plant Mol Biol, 2001. **52**: p. 29-66.

Ziegler, J. and P.J. Facchini, *Alkaloid biosynthesis: metabolism and trafficking.* Annu Rev Plant Biol, 2008. **59**: p. 735-69.

Kornienko, A. and A. Evidente, *Chemistry, biology, and medicinal potential of narciclasine and its congeners.* Chem Rev, 2008. **108**(6): p. 1982-2014.

Bastida, J., et al., *Chap3 Chemical and biological aspects of Amaryllidaceae alkaloids*, E.D. Muñoz-Torrero, et al., Editors. 2011, Recent Advances in Pharmaceutical Sciences,Recent Advances in Pharmaceutical Sciences, 2011: 65-100 ISBN: 978-81-7895-528-5.

Takos, A.M. and F. Rook, *Towards a molecular understanding of the biosynthesis of amaryllidaceae alkaloids in support of their expanding medical use.* Int J Mol Sci, 2013. **14**(6): p. 11713-41.

Jin, Z., *Amaryllidaceae and Sceletium alkaloids.* Nat Prod Rep, 2013. **30**(6): p. 849-68.

Singh, A. and I. Desgagne-Penix, *Biosynthesis of the Amaryllidaceae alkaloids.* Plant Science Today, 2014. **1**(2): p. 114-120.

Heinrich, M. and H. Lee Teoh, *Galanthamine from snowdrop--the development of a modern drug against Alzheimer's disease from local Caucasian knowledge.* J Ethnopharmacol, 2004. **92**(2-3): p. 147-62.

Conforti, F., et al., *Quantitative determination of Amaryllidaceae alkaloids from Galanthus reginae-olgae subsp. vernalis and in vitro activities relevant for neurodegenerative diseases.* Pharm Biol, 2010. **48**(1): p. 2-9.

Lubbe, A., et al., *Seasonal accumulation of major alkaloids in organs of pharmaceutical crop Narcissus Carlton.* Phytochemistry, 2013. **88**: p. 43-53.

da Silva, A.F., et al., *Anxiolytic-, antidepressant- and anticonvulsant-like effects of the alkaloid montanine isolated from Hippeastrum vittatum.* Pharmacol Biochem Behav, 2006. **85**(1): p. 148-54.

Furusawa, E. and S. Furusawa, *Effect of pretazettine and viva-natural, a dietary seaweed extract, on spontaneous AKR leukemia in comparison with standard drugs.* Oncology, 1988. **45**(3): p. 180-6.

Nair, J.J. and J. van Staden, *Pharmacological and toxicological insights to the South African Amaryllidaceae.* Food Chem Toxicol, 2013. **62C**: p. 262-275.

Lubbe, A., R. Verpoorte, and Y.H. Choi, *Effects of fungicides on galanthamine and metabolite profiles in Narcissus bulbs.* Plant Physiol Biochem, 2012. **58**: p. 116-23.

Sener, B., I. Orhan, and J. Satayavivad, *Antimalarial activity screening of some alkaloids and the plant extracts from Amaryllidaceae.* Phytother Res, 2003. **17**(10): p. 1220-3.

Pettit, G.R., et al., *Antineoplastic agents, 99. Amaryllis belladonna.* J Nat Prod, 1984. **47**(5): p. 796-801.

Weniger, B., et al., *Cytotoxic activity of Amaryllidaceae alkaloids.* Planta Med, 1995. **61**(1): p. 77-9.

Hua, D.H., S. Saha, and D.J. Takemoto, *Anticancer activities of 2,5,8,9-substituted 6-oxo-1,2,3,4,5,6-hexahydrophenanthridines on multi-drug-resistant phenotype cells.* Anticancer Res, 1997. **17**(4A): p. 2435-41.

Lamoral-Theys, D., et al., *Lycorine and its derivatives for anticancer drug design.* Mini Rev Med Chem, 2010. **10**(1): p. 41-50.

Koola, M.M., et al., *Potential role of the combination of galantamine and memantine to improve cognition in schizophrenia.* Schizophr Res, 2014.

Konrath, E.L., et al., *Alkaloids as a source of potential anticholinesterase inhibitors for the treatment of Alzheimer's disease.* J Pharm Pharmacol, 2013. **65**(12): p. 1701-25.

Birks, J. and D. Craig, *Galantamine for vascular cognitive impairment.* Cochrane Database Syst Rev, 2013. **4**: p. CD004746.

Furst, R. and I. Zundorf, *Plant-Derived Anti-Inflammatory Compounds: Hopes and Disappointments regarding the Translation of Preclinical Knowledge into Clinical Progress.* Mediators Inflamm, 2014. **2014**: p. 146832.

Nair, J.J., et al., *Cytotoxic agents of the crinane series of amaryllidaceae alkaloids.* Nat Prod Commun, 2012. **7**(12): p. 1677-88.

Barton, D.H.R. and T. Cohen, In Festschrift Arthur Stoll p.117 Birkhauser, Basel., 1957.

Barton, D.H.R., et al., *Phenol oxidation and biosynthesis. Part VI. The biogenesis of Amaryllidaceae alkaloids.* J. Chem. Soc., 1963: p. 4545-4558.

Battersby, A.R., H.M. Fales, and W.C. Wildman, *Biosynthesis in the Amaryllidaceae. Tyrosine and norbelladine as precursors of haemanthamine.* J. Amer. Chem. Soc., 1961. **83**: p. 4098-4099.

Eichhorn, J., et al., *Biosynthesis of the Amaryllidaceae alkaloid galanthamine.* Phytochemistry, 1998. **49**(4): p. 1037-1047.

El Tahchy, A., et al., *New method for the study of Amaryllidaceae alkaloid biosynthesis using biotransformation of deuterium-labeled precursor in tissue cultures.* Acta Biochim Pol, 2010. **57**(1): p. 75-82.

El Tahchy, A., et al., *Kinetic study of the rearrangement of deuterium-labeled 4'-O-methylnorbelladine in Leucojum aestivum shoot cultures by mass spectrometry. Influence of precursor feeding on amaryllidaceae alkaloid accumulation.* J Nat Prod, 2011. **74**(11): p. 2356-61.

Grisebach, H., *Comparative biosynthetic pathways in higher plants.* Pure and Applied chemistry, 1973. **34**(3-4): p. 487-513.

Chu, H.Y., E. Wegel, and A. Osbourn, *From hormones to secondary metabolism: the emergence of metabolic gene clusters in plants.* Plant J, 2011. **66**(1): p. 66-79.

Wink, M., *Biochemistry of plant secondary metabolism.* Annual plant reviews, 2010. **40**.

Zenk, M., *Biosynthese von vanillin in Vanilla planifolia Andr. Z.* Pflanzen Physiol., 1965. **53**: p. 404-414.

Walton, N.J., M.J. Mayer, and A. Narbad, *Vanillin.* Phytochemistry, 2003. **63**(5): p. 505-515.

Podstolski, A., et al., *Unusual 4-hydroxybenzaldehyde synthase activity from tissue cultures of the vanilla orchid Vanilla planifolia.* Phytochemistry, 2002. **61**(6): p. 611-620.

Negishi, O., K. Sugiura, and Y. Negishi, *Biosynthesis of vanillin via ferulic acid in Vanilla planifolia.* J Agric Food Chem, 2009. **57**(21): p. 9956-61.

Havkin-Frenkel, D. and F. Belanger, *Application of Metabolic Engineering to Vanillin Biosynthetic Pathways in Vanilla Planifolia*, in *Applications of Plant Metabolic Engineering*, R. Verpoorte, A.W. Alfermann, and T.S. Johnson, Editors. 2007, Springer Netherlands. p. 175-196.

Aza-Gonzalez, C., H.G. Nunez-Palenius, and N. Ochoa-Alejo, *Molecular biology of capsaicinoid biosynthesis in chili pepper (Capsicum spp.).* Plant Cell Rep, 2011. **30**(5): p. 695-706.

Jiang, Y., et al., *Molecular and analysis of a phenylalanine ammonia-lyase gene (LrPAL2) from Lycoris radiata.* Mol Biol Rep, 2013. **40**(3): p. 2293-300.

Wang, R., et al., *De novo sequence assembly and characterization of Lycoris aurea transcriptome using GS FLX titanium platform of 454 pyrosequencing.* PLoS One, 2013. **8**(4): p. e60449.

Ehlting, J., et al., *Cytochromes P450 in phenolic metabolism.* Phytochemistry Reviews, 2006. **5**(2-3): p. 239-270.

Mizutani, M. and F. Sato, *Unusual P450 reactions in plant secondary metabolism.* Arch Biochem Biophys, 2011. **507**(1): p. 194-203.

Nelson, D. and D. Werck-Reichhart, *A P450-centric view of plant evolution.* Plant J, 2011. **66**(1): p. 194-211.

Sircar, D. and A. Mitra, *Accumulation of p-hydroxybenzoic acid in hairy roots of Daucus carota 2: confirming biosynthetic steps through feeding of inhibitors and precursors.* J Plant Physiol, 2009. **166**(13): p. 1370-80.

Facchini, P.J., K.L. Huber-Allanach, and L.W. Tari, *Plant aromatic L-amino acid decarboxylases: evolution, biochemistry, regulation, and metabolic engineering applications.* Phytochemistry, 2000. **54**(2): p. 121-38.

Hagel, J.M. and P.J. Facchini, *Benzylisoquinoline Alkaloid Metabolism: A Century of Discovery and a Brave New World.* Plant and Cell Physiology, 2013. **54**(5): p. 647-672.

Desgagné-Penix, I. and P.J. Facchini, *Benzylisoquinoline Alkaloid Biosynthesis*, in *Plant Metabolism and Biotechnology*. 2011, John Wiley & Sons, Ltd. p. 241-261.

Torrens-Spence, M.P., et al., *Biochemical evaluation of a parsley tyrosine decarboxylase results in a novel 4-hydroxyphenylacetaldehyde synthase enzyme.* Biochemical and Biophysical Research Communications, 2012. **418**(2): p. 211-216.

Dewick, P.M., *Medicinal natural products: a biosynthetic apporach*. 3rd ed. 2009: John Wiley and Sons Ltd. 539.

Ghosal, S., A. Shanthy, and S.K. Singh, *Isocraugsodine, an n-arylidenephenethylamine from Crinum asiaticum and its e-z isomerism.* phytochemistry, 1988. **27**(6): p. 1849-1852.

Maresh, J.J., et al., *Strictosidine synthase: mechanism of a Pictet-Spengler catalyzing enzyme.* J Am Chem Soc, 2008. **130**(2): p. 710-23.

Luk, L.Y., et al., *Mechanistic studies on norcoclaurine synthase of benzylisoquinoline alkaloid biosynthesis: an enzymatic Pictet-Spengler reaction.* Biochemistry, 2007. **46**(35): p. 10153-61.

Lee, E.J. and P. Facchini, *Norcoclaurine synthase is a member of the pathogenesis-related 10/Bet v1 protein family.* Plant Cell, 2010. **22**(10): p. 3489-503.

Minami, H., et al., *Functional analysis of norcoclaurine synthase in Coptis japonica.* J Biol Chem, 2007. **282**(9): p. 6274-82.

Ilari, A., et al., *Structural basis of enzymatic (S)-norcoclaurine biosynthesis.* J Biol Chem, 2009. **284**(2): p. 897-904.

Bonamore, A., et al., *Norcoclaurine synthase: mechanism of an enantioselective pictet-spengler catalyzing enzyme.* Molecules, 2010. **15**(4): p. 2070-8.

Pasquo, A., et al., *Cloning, expression, crystallization and preliminary X-ray data analysis of norcoclaurine synthase from Thalictrum flavum.* Acta Crystallogr Sect F Struct Biol Cryst Commun, 2008. **64**(Pt 4): p. 281-3

Samanani, N., D.K. Liscombe, and P.J. Facchini, *Molecular cloning and characterization of norcoclaurine synthase, an enzyme catalyzing the first committed step in benzylisoquinoline alkaloid biosynthesis.* Plant J, 2004. **40**(2): p. 302-13.

Kutchan, T.M., *Strictosidine: From alkaloid to enzyme to gene.* Phytochemistry, 1993. **32**(3): p. 493-506.

Stockigt, J., et al., *3D-Structure and function of strictosidine synthase--the key enzyme of monoterpenoid indole alkaloid biosynthesis.* Plant Physiol Biochem, 2008. **46**(3): p. 340-55.

Liscombe, D.K., G.V. Louie, and J.P. Noel, *Architectures, mechanisms and molecular evolution of natural product methyltransferases.* Nat Prod Rep, 2012. **29**(10): p. 1238-50.

Lam, K.C., et al., *Structure, function, and evolution of plant O-methyltransferases.* Genome, 2007. **50**(11): p. 1001-13.

Rueffer, M., N. Nagakura, and M.H. Zenk, *Partial Purification and Properties of S-Adenosylmethionine: (R), (S)-Norlaudanosoline-6-O-Methyltransferase from Argemone platyceras Cell Cultures.* Planta Med, 1983. **49**(11): p. 131-7.

Frick, S. and T.M. Kutchan, *Molecular cloning and functional expression of O-methyltransferases common to isoquinoline alkaloid and phenylpropanoid biosynthesis.* Plant J, 1999. **17**(4): p. 329-39.

Morishige, T., et al., *Molecular characterization of the S-adenosyl-L-methionine:3'-hydroxy-N-methylcoclaurine 4'-O-methyltransferase involved in isoquinoline alkaloid biosynthesis in Coptis japonica.* J Biol Chem, 2000. **275**(30): p. 23398-405.

Morishige, T., et al., *Molecular cloning of columbamine O-methyltransferase from cultured Coptis japonica cells.* Eur J Biochem, 2002. **269**(22): p. 5659-67.

Ounaroon, A., et al., *(R,S)-Reticuline 7-O-methyltransferase and (R,S)-norcoclaurine 6-O-methyltransferase of Papaver somniferum - cDNA cloning and characterization of methyl transfer enzymes of alkaloid biosynthesis in opium poppy.* Plant J, 2003. **36**(6): p. 808-19.

Ziegler, J., et al., *Comparative macroarray analysis of morphine containing Papaver somniferum and eight morphine free Papaver species identifies an O-methyltransferase involved in benzylisoquinoline biosynthesis.* Planta, 2005. **222**(3): p. 458-71.

Desgagne-Penix, I., et al., *Integration of deep transcript and targeted metabolite profiles for eight cultivars of opium poppy.* Plant Mol Biol, 2012. **79**(3): p. 295-313.

Pienkny, S., et al., *Functional characterization of a novel benzylisoquinoline O-methyltransferase suggests its involvement in papaverine biosynthesis in opium poppy (Papaver somniferum L).* Plant J, 2009. **60**(1): p. 56-67.

Nomura, T. and T.M. Kutchan, *Three new O-methyltransferases are sufficient for all O-methylation reactions of ipecac alkaloid biosynthesis in root culture of Psychotria ipecacuanha.* J Biol Chem, 2010. **285**(10): p. 7722-38.

Mann, J.D., H.M. Fales, and S.H. Mudd, *ALKALOIDS AND PLANT METABOLISM. VI. O-METHYLATION IN VITRO OF NORBELLADINE, A PRECURSOR OF AMARYLLIDACEAE ALKALOIDS.* J Biol Chem, 1963. **238**: p. 3820-3.

Ikezawa, N., K. Iwasa, and F. Sato, *Molecular cloning and characterization of CYP80G2, a cytochrome P450 that catalyzes an intramolecular C-C phenol coupling of (S)-reticuline in magnoflorine biosynthesis, from cultured Coptis japonica cells.* J Biol Chem, 2008. **283**(14): p. 8810-21.

Gesell, A., et al., *CYP719B1 is salutaridine synthase, the C-C phenol-coupling enzyme of morphine biosynthesis in opium poppy.* J Biol Chem, 2009. **284**(36): p. 24432-42.

Berkov, S., et al., *Evolution of alkaloid biosynthesis in the genus Narcissus.* Phytochemistry, 2014. **99**: p. 95-106.

Bhandarkar, J.G. and G.W. Kirby, *Structure and biosynthesis of chlidanthine.* Journal of the Chemical Society C: Organic, 1970(9): p. 1224-1227.

Abdallah, O.M., A.A. Ali, and H. Itokawa, *2-O-acetyl chlidanthine; An alkaloid from Haemanthus multiflorus.* Phytochemistry, 1989. **28**(11): p. 3248-3249.

Ziegler, J., et al., *Comparative transcript and alkaloid profiling in Papaver species identifies a short chain dehydrogenase/reductase involved in morphine biosynthesis.* Plant J, 2006. **48**(2): p. 177-92.

Unterlinner, B., R. Lenz, and T.M. Kutchan, *Molecular cloning and functional expression of codeinone reductase: the penultimate enzyme in morphine biosynthesis in the opium poppy Papaver somniferum.* Plant J, 1999. **18**(5): p. 465-75.

Lenz, R. and M.H. Zenk, *Purification and properties of codeinone reductase (NADPH) from Papaver somniferum cell cultures and differentiated plants.* Eur J Biochem, 1995. **233**(1): p. 132-9.

Chen, X. and P.J. Facchini, *Short-chain dehydrogenase/reductase catalyzing the final step of noscapine biosynthesis is localized to laticifers in opium poppy.* Plant J, 2014. **77**(2): p. 173-84.

Hagel, J.M. and P.J. Facchini, *Dioxygenases catalyze the O-demethylation steps of morphine biosynthesis in opium poppy.* Nat Chem Biol, 2010. **6**(4): p. 273-5.

Hagel, J.M. and P.J. Facchini, *Biochemistry and occurrence of o-demethylation in plant metabolism.* Front Physiol, 2010. **1**: p. 14.

Farrow, S.C. and P.J. Facchini, *Dioxygenases catalyze o-demethylation and o,o-demethylenation with widespread roles in benzylisoquinoline alkaloid metabolism in opium poppy.* J Biol Chem, 2013. **288**(40): p. 28997-9012.

Ikezawa, N., et al., *Molecular cloning and characterization of CYP719, a methylenedioxy bridge-forming enzyme that belongs to a novel P450 family, from cultured Coptis japonica cells.* J Biol Chem, 2003. **278**(40): p. 38557-65.

Ikezawa, N., K. Iwasa, and F. Sato, *Molecular cloning and characterization of methylenedioxy bridge-forming enzymes involved in stylopine biosynthesis in Eschscholzia californica.* FEBS J, 2007. **274**(4): p. 1019-35.

Diaz Chavez, M.L., et al., *Characterization of two methylenedioxy bridge-forming cytochrome P450-dependent enzymes of alkaloid formation in the Mexican prickly poppy Argemone mexicana.* Arch Biochem Biophys, 2011. **507**(1): p. 186-93.

Hagel, J.M., et al., *Characterization of a Flavoprotein Oxidase from Opium Poppy Catalyzing the Final Steps in Sanguinarine and Papaverine Biosynthesis.* Journal of Biological Chemistry, 2012. **287**(51): p. 42972-42983.

Gesell, A., et al., *Heterologous expression of two FAD-dependent oxidases with (S)-tetrahydroprotoberberine oxidase activity from Arge mone mexicana and Berberis wilsoniae in insect cells.* Planta, 2011. **233**(6): p. 1185-97.

Takemura, T., et al., *Molecular cloning and characterization of a cytochrome P450 in sanguinarine biosynthesis from Eschscholzia californica cells.* Phytochemistry, 2013. **91**: p. 100-8.

Beaudoin, G.A.W. and P.J. Facchini, *Isolation and characterization of a cDNA encoding (S)-cis-N-methylstylopine 14-hydroxylase from opium poppy, a key enzyme in sanguinarine biosynthesis.* Biochemical and Biophysical Research Communications, 2013. **431**(3): p. 597-603.

Schilmiller, A.L., E. Pichersky, and R.L. Last, *Taming the hydra of specialized metabolism: how systems biology and comparative approaches are revolutionizing plant biochemistry.* Curr Opin Plant Biol, 2012. **15**(3): p. 338-44.

Xiao, M., et al., *Transcriptome analysis based on next-generation sequencing of non-model plants producing specialized metabolites of biotechnological interest.* Journal of Biotechnology, 2013. **166**(3): p. 122-134.

Desgagne-Penix, I., et al., *Integration of deep transcriptome and proteome analyses reveals the components of alkaloid metabolism in opium poppy cell cultures.* BMC Plant Biol, 2010. **10**: p. 252.

Desgagne-Penix, I. and P.J. Facchini, *Systematic silencing of benzylisoquinoline alkaloid biosynthetic genes reveals the major route to papaverine in opium poppy.* Plant J, 2012. **72**(2): p. 331-44.

Liscombe, D.K., et al., *Targeted metabolite and transcript profiling for elucidating enzyme function: isolation of novel N-methyltransferases from three benzylisoquinoline alkaloid-producing species.* Plant J, 2009. **60**(4): p. 729-43.

Minami, H., et al., *Microbial production of plant benzylisoquinoline alkaloids.* Proc Natl Acad Sci U S A, 2008. **105**(21): p. 7393-8.

Hawkins, K.M. and C.D. Smolke, *Production of benzylisoquinoline alkaloids in Saccharomyces cerevisiae.* Nat Chem Biol, 2008. **4**(9): p. 564-73.

Nakagawa, A., et al., *Bench-top fermentative production of plant benzylisoquinoline alkaloids using a bacterial platform.* Bioeng Bugs, 2012. **3**(1): p. 49-53.

Facchini, P.J., et al., *Synthetic biosystems for the production of high-value plant metabolites.* Trends Biotechnol, 2012. **30**(3): p. 127-31.

Fossati, E., et al., *Reconstitution of a 10-gene pathway for synthesis of the plant alkaloid dihydrosanguinarine in Saccharomyces cerevisiae.* Nat Commun, 2014. **5**: p. 3283.

B.8 Supplementary Figures

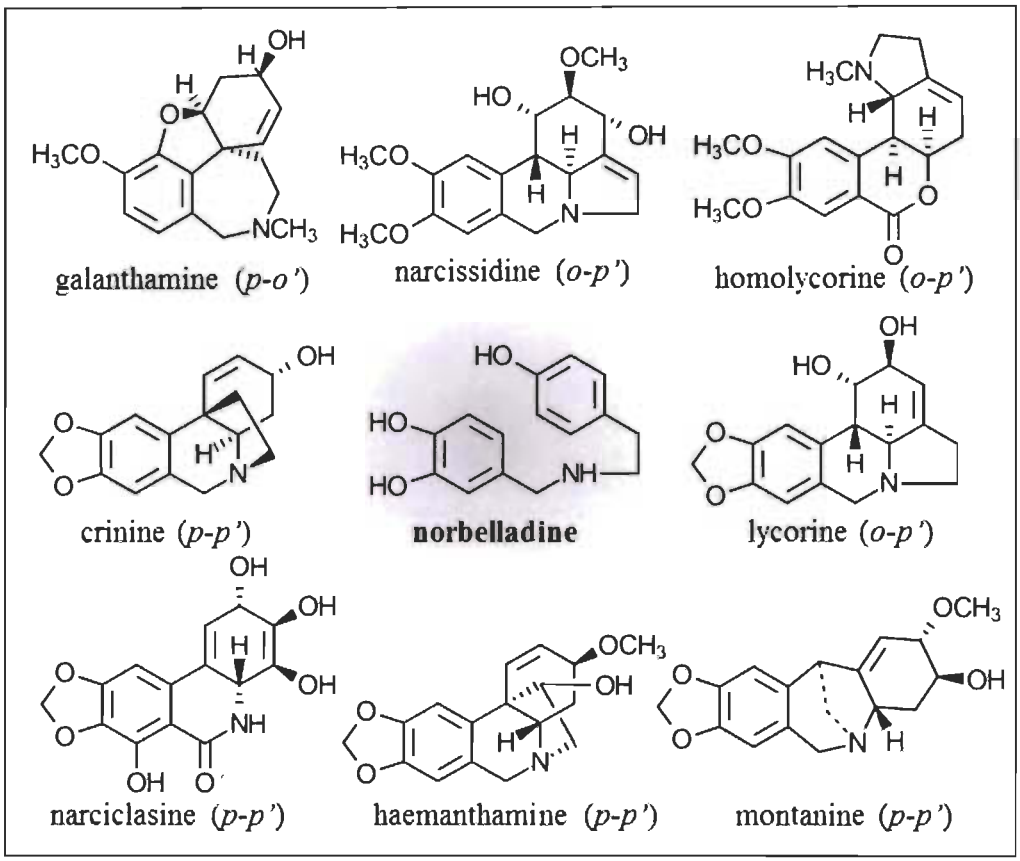


Figure S1. Examples of diverse chemical structures of AA types derived from simple skeletal structure nobelladine (shaded grey).

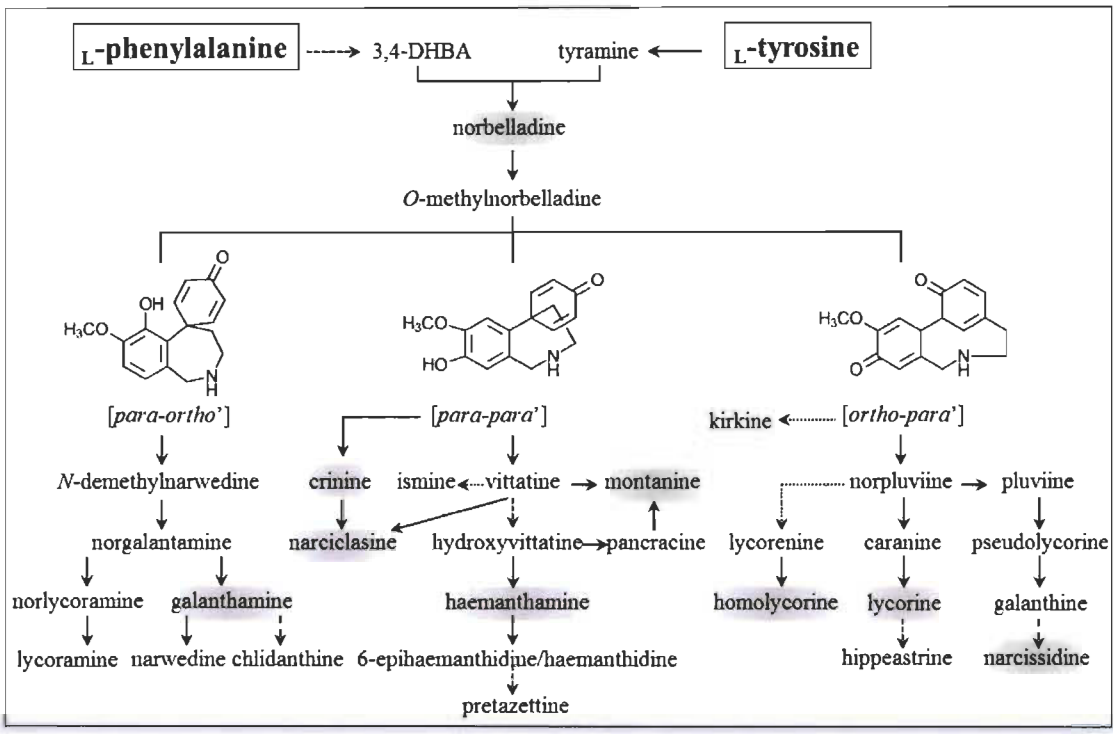


Figure S2. Proposed AA biosynthetic pathways. Three major subgroups of AAs (*para-ortho', para-para', ortho-para'*) are represented along with the nine proposed AA-types: norbelladine-, galanthamine-, crinine-, narciclasine-, haemanthamine-, montanine-, homolycorine-, lycorine-, and narcissidine-type (shaded gray). Broken arrow indicates more than one step.

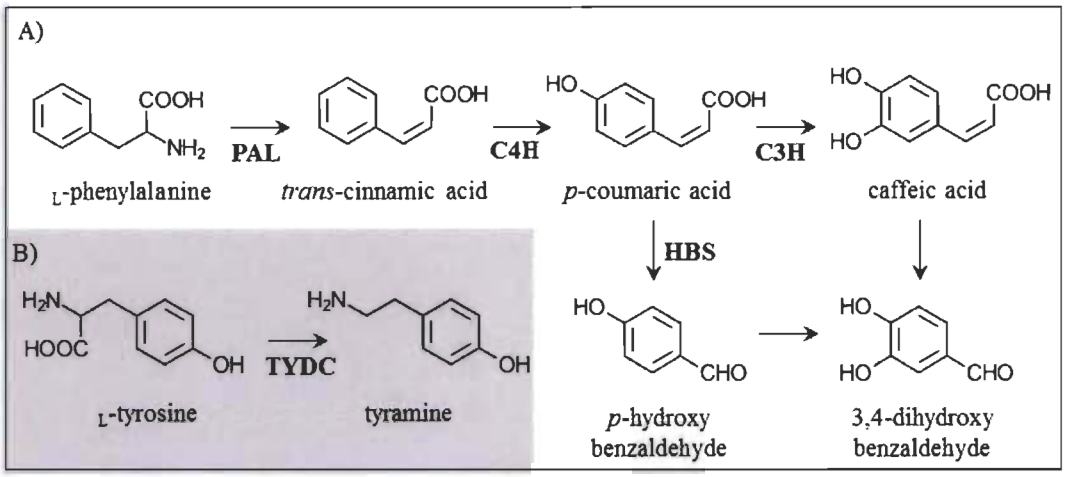


Figure S3. Proposed pathways to AA precursors. A) 3,4-dihydroxybenzaldehyde (3,4-DHBA) biosynthesis depicting the two possible routes from *p*-coumaric acid to form 3,4-DHBA: the oxidative 'ferulate' and the non-oxidative 'benzoate' pathways B) Tyramine biosynthesis. Arrows without labeling reflect chemical reactions that have not been enzymatically characterized. Enzymes that have been cloned, characterized and identified are labeled in black bold. Enzyme abbreviations: PAL, phenylalanine ammonia-lyase; C4H, cinnamate 4-hydroxylase; C3H, coumarate 3-hydroxylase; HBS, 4-hydroxybenzaldehyde synthase; TYDC, tyrosine decarboxylase.

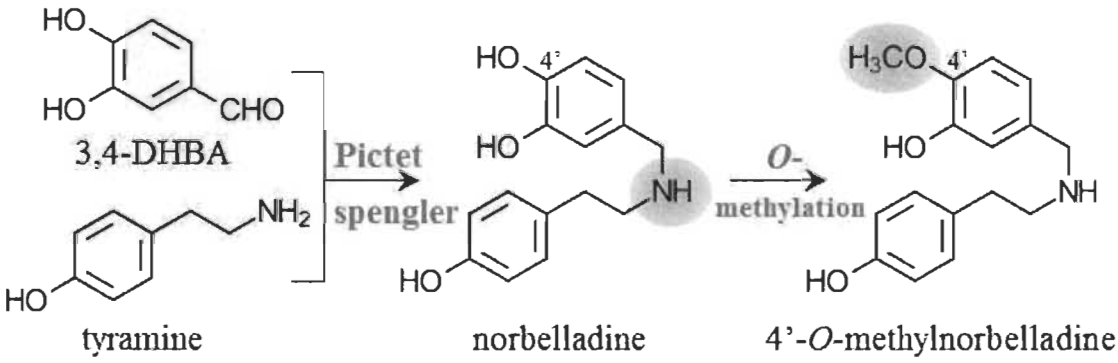


Figure S4. Biosynthesis of 4'-O-methylnorbelladine. Functional groups shaded gray highlights chemical conversion.

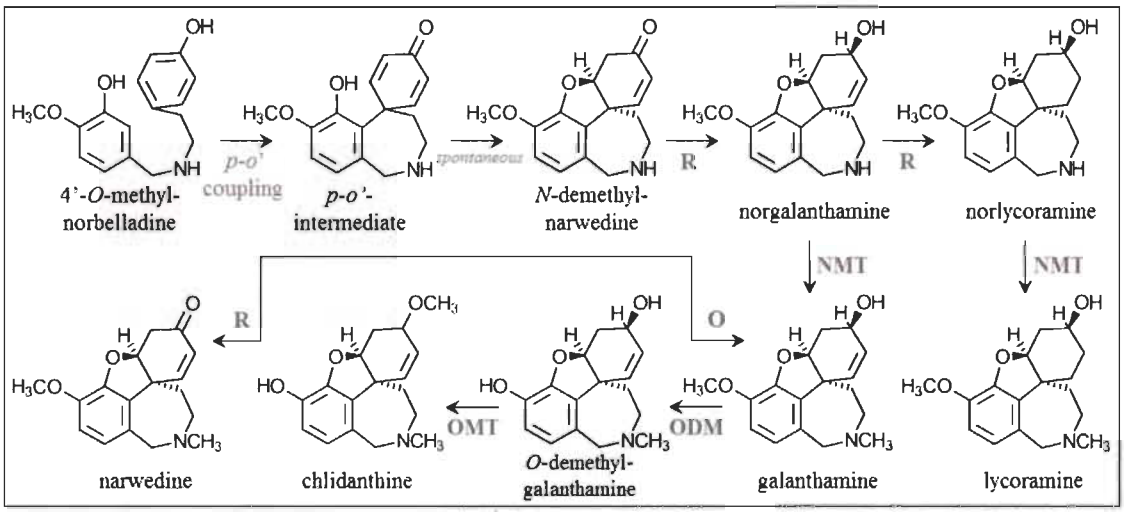


Figure S5. Biosynthesis of *p-o'* alkaloids such as galanthamine and derivatives. Abbreviations: CYP, cytochrome P450 monooxygenase; R, reduction; NMT, *N*-methylation; ODM, *O*-demethylation; OMT, *O*-methylation.

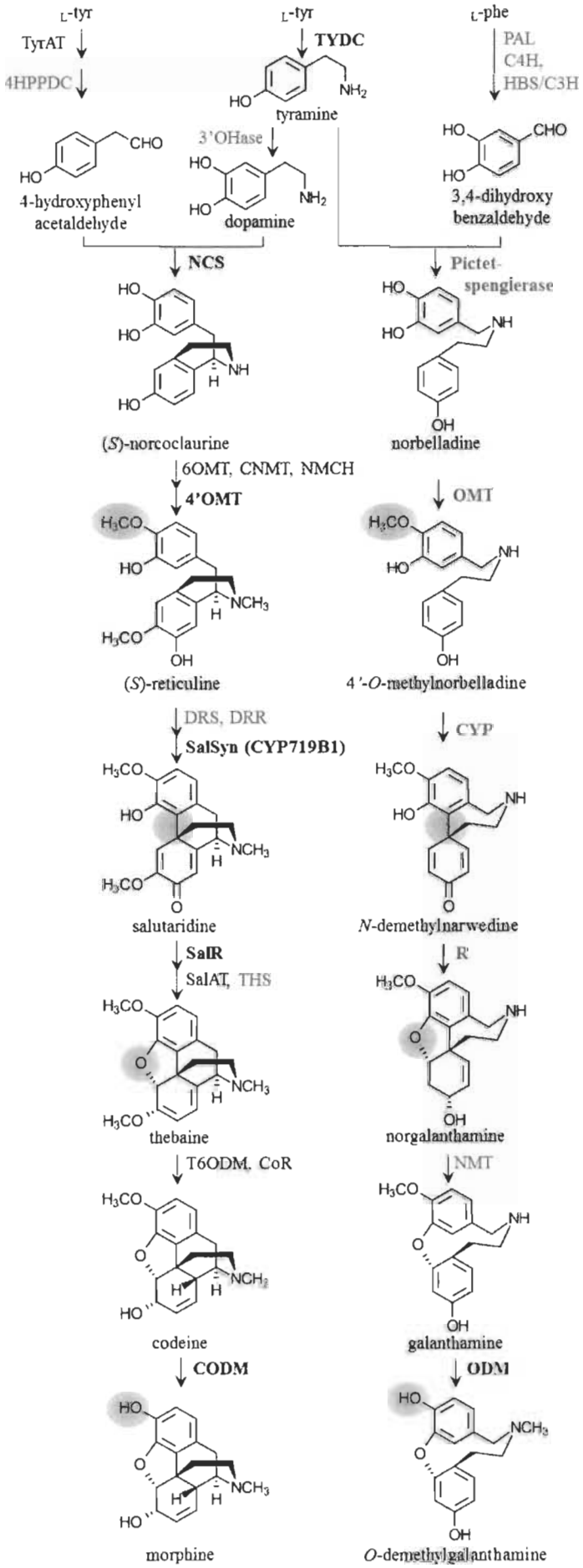

L-tyr
TyrAT
4HPPDC
L-tyr
TYDC
tyramine
L-phe
PAL
C4H,
HBS/C3H
4-hydroxyphenyl acetaldehyde
3'OHase
dopamine
3,4-dihydroxy benzaldehyde
NCS
Pictet-spengierase
(S)-norcoclaurine
norbelladine
6OMT, CNMT, NMCH
4'OMT
OMT
(S)-reticuline
4'-O-methylnorbelladine
DRS, DRR
SalSyn (CYP719B1)
CYP
salutaridine
N-demethylnarwedine
SalR
SalAT, THS
R'
thebaine
norgalanthamine
T6ODM, CoR
NMT
codeine
galanthamine
CODM
ODM
morphine
O-demethylgalanthamine

Figure S6: Analogy of reactions between the BIA (left) and AA (right) biosynthetic pathways. Broken arrows represent more than one reaction. Enzymes for which corresponding genes have been isolated from *P. somniferum* are shown in black. Enzymes for which corresponding genes have not been isolated or proposed enzymes are shown in gray. Functional groups shaded gray highlights analogous chemical conversions. Abbreviations: TyrAT, tyrosine aminotransferase; 4HPPDC, 4-hydroxyphenylpuruvate decarboxylase; TYDC, tyrosine/DOPA decarboxylase; 3OHase, tyrosine/tyramine 3-hydroxylase; PAL, phenylalanine ammonia-lyase; C4H, cinnamate hydroxylase; HBS, hydroxybenzyl synthase; C3H, coumarate hydroxylase; NCS, norcoclaurine synthase; 6OMT, norcoclaurine 6-*O*-methyltransferase; CNMT, coclaurine *N*-methyltransferase; NMCH, *N*-methylcoclaurine 3'-hydroxylase; 4'OMT, 3'-hydroxyl-*N*-methylcoclaurine 4'-*O*methyltransferase; SalSyn, salutaridine synthase; SalR, salutaridine reductase; SalAT, salutaridinol 7-*O*-acetyltransferase; T6ODM, thebaine 6-O-demethylase; CoR, codeinone reductase; CODM, codeine *O*-demethylase; OMT, *O*-methyltransferase; CYP, cytochrome P450; R, reductase; NMT, *N*-methyltransferase; ODM, *O*-demethylase.

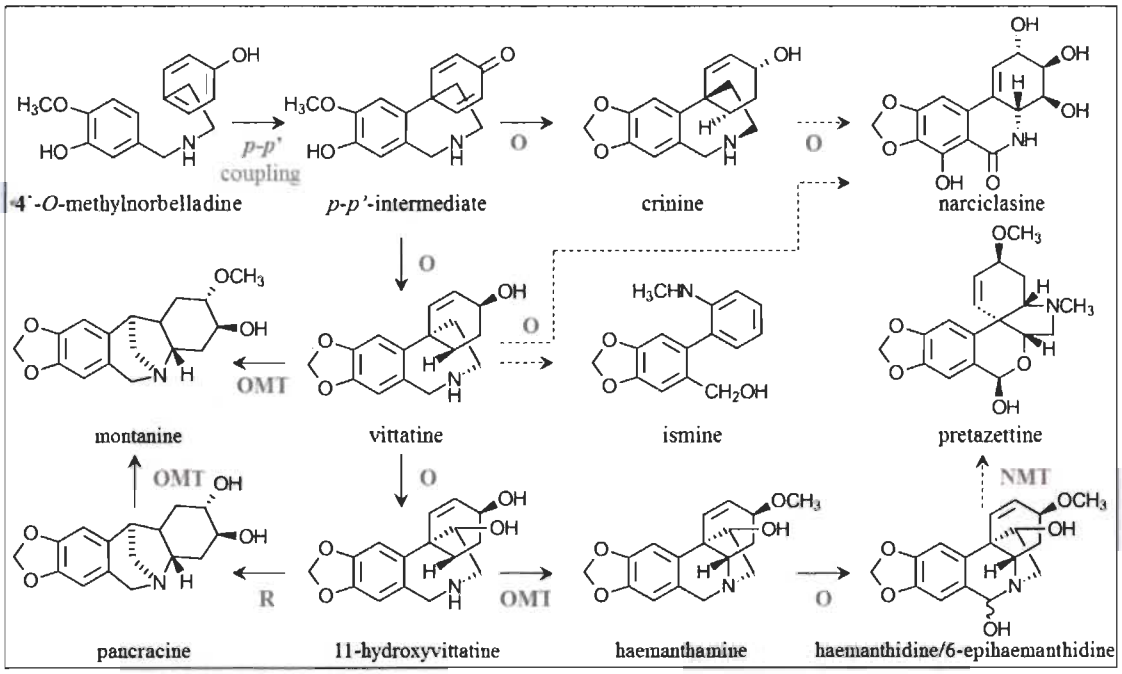


Figure S7. Biosynthesis of *p-p'* alkaloids such as crinine, haemanthamine, narciclasine and montanine. Broken arrow represents more than one biochemical reaction. Abbreviations: R, reduction; O, oxidation; NMT, *N*-methylation; OMT, *O*-methylation.

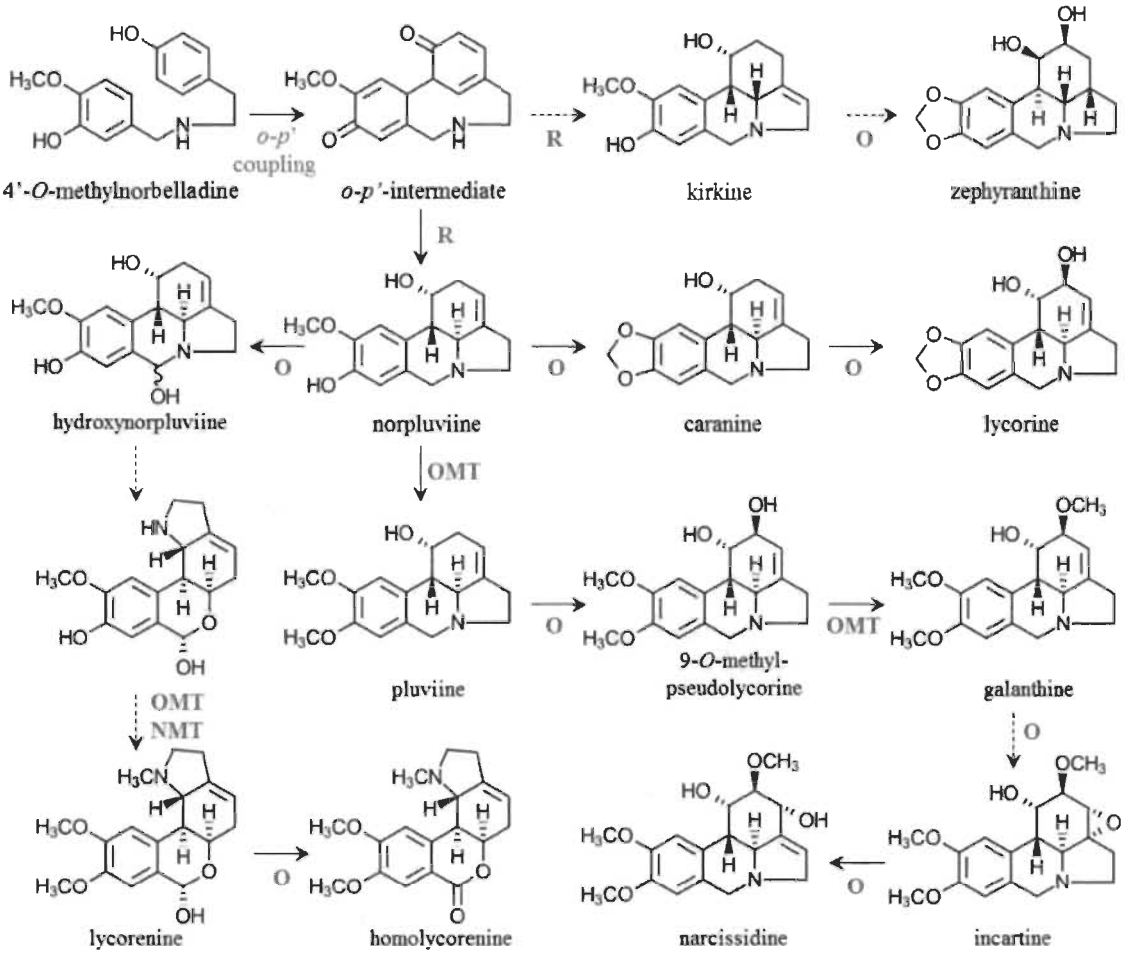


Figure S8. Biosynthesis of *o-p'* alkaloids such as lycorine, homolycorine and narcissidine. Broken arrow represents more than one biochemical reaction. Abbreviations: R, reduction; O, oxidation; NMT, *N*-methylation; OMT, *O*-methylation.

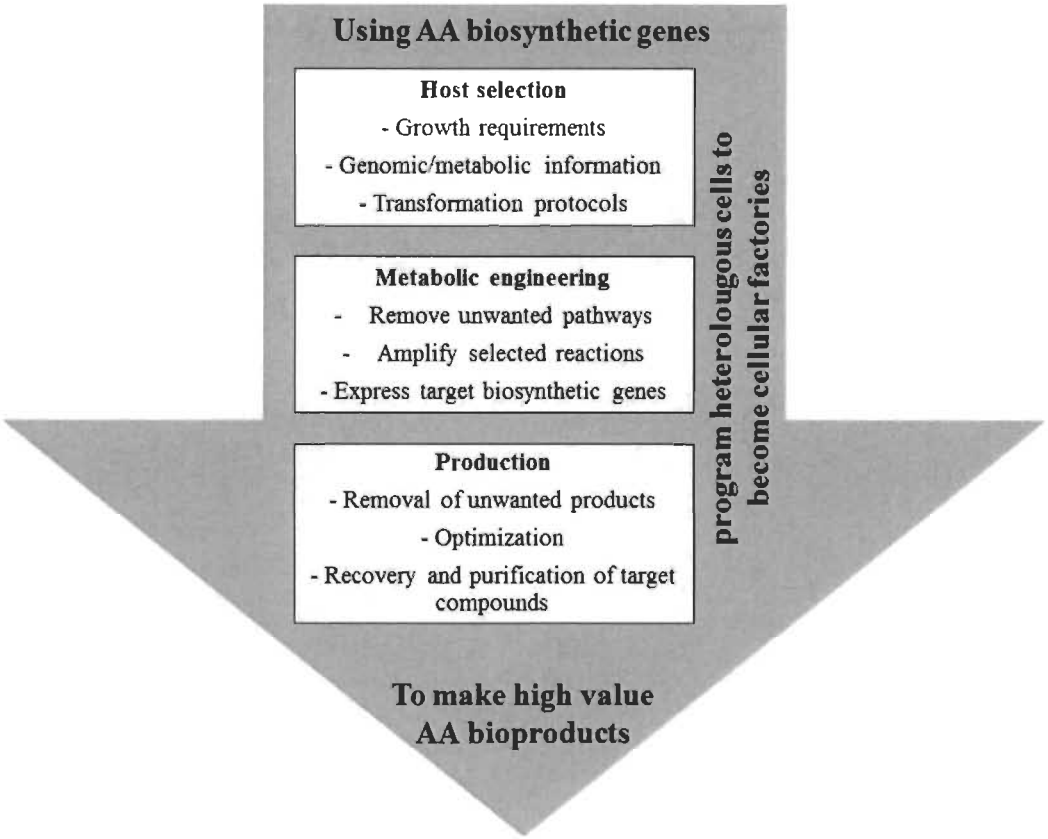


Figure S9. Representation of the strategy for the metabolic engineering of microorganism for AA drug production.

ANNEX C

SUPPLEMENTARY DATA FOR CHAPTER II

Table C.1 Illumina sequencing output and assembly overview

Read Trimming and Clipping of Adapters	
Raw Paired Reads[a]	73081603
Surviving Paired Reads[b]	66054792
Surviving Paired Reads[c] (%)	90.385
Normalization	
Surviving Paired Reads after normalization[d]	10523999
Surviving Paired Reads after normalization[e] (%)	15.932
Trinity *de novo* Assembly	
Nb. Transcripts[f]	195347
Nb. Components[f]	98332
Total Transcripts Length (bp)	148818643
Max. Transcript Length (bp)	13507
Min. Transcript Length (bp)	224
Median Transcript Length (bp)	478
Mean Transcript Length (bp)	761
N50 (bp)[g]	1150
BLAST annotation and filtered annotated components	
Nb. Transcripts	11708
Nb. Components	4288
Total Transcripts Length (bp)	17131946
Max. Transcript Length (bp)	9315
Min. Transcript Length (bp)	297
Median Transcript Length (bp)	1308
Mean Transcript Length (bp)	1463
N50 (bp)[h]	1768

[a] Number of Paired Reads obtained from the sequencer

[b] Number of Remaining Paired Reads after the trimming step

[c] Percentage of Surviving Paired Reads / Raw Paired Reads

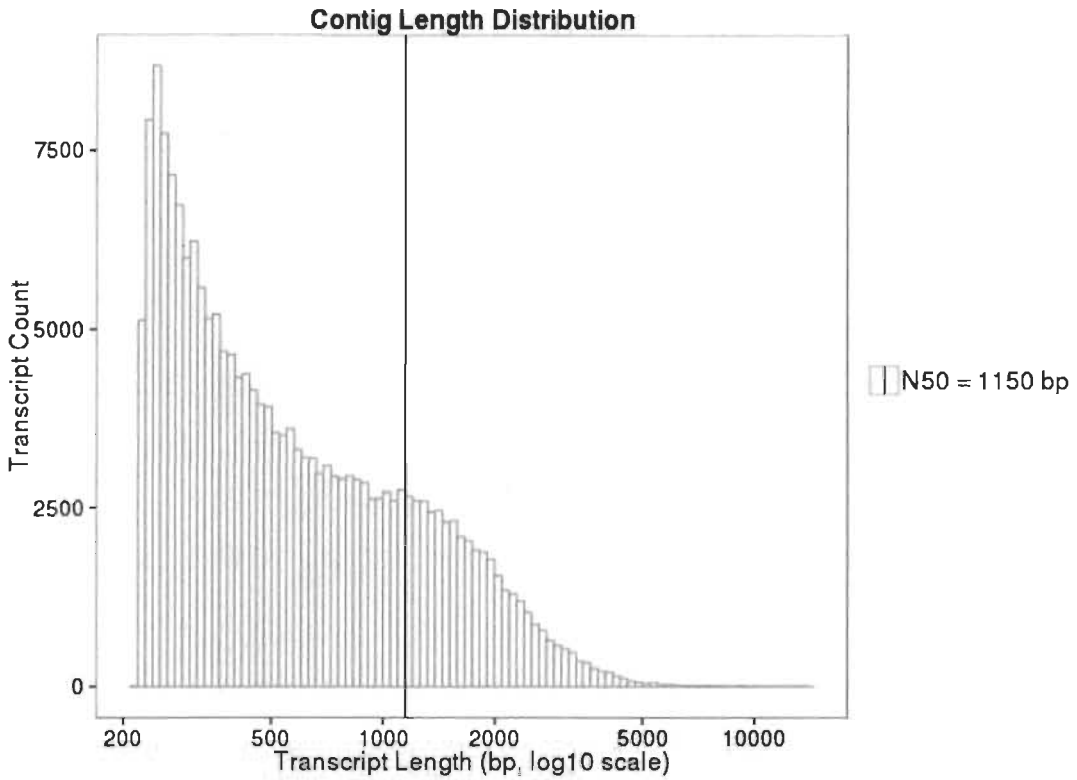
[d] Number of remaining Paired Reads after the normalization step

[e] Percentage of Surviving Paired Reads after normalization / Surviving Paired Reads after trimming

[f] Trinity has created a list of transcripts (contigs) representing the transcriptome isoforms. The transcripts are grouped in components loosely representing genes. Transcript names are prefixed by the component/gene name e.g. transcripts c115_g5_i1 and c115_g5_i2 are derived from the same isolated de Bruijn graph and therefore share the same component/gene number c115_g5.

[g] Corresponding contig length distribution figure N50 = 1150bp below

[h] Corresponding contig length distribution figure N50 = 1768bp below


Contig Length Distribution
Transcript Count
7500
5000
2500
0
200
500
1000
2000
5000
10000
Transcript Length (bp, log10 scale)
N50 = 1150 bp

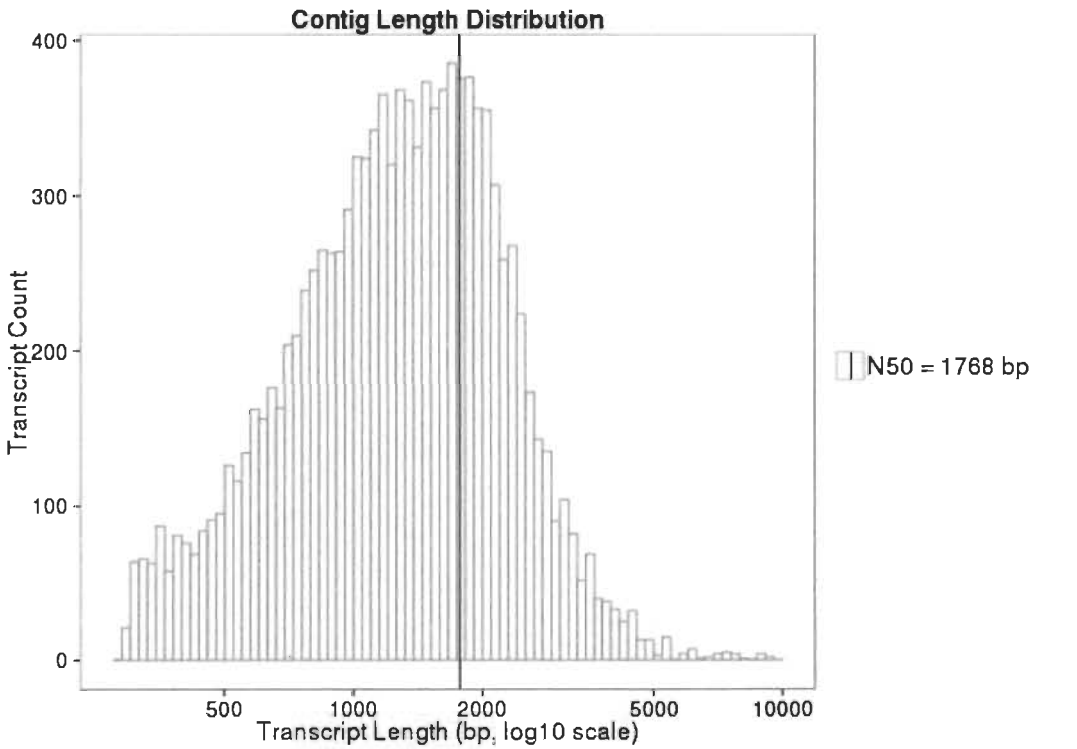

Contig Length Distribution
Transcript Count
400
300
200
100
0
500
1000
2000
5000
10000
Transcript Length (bp, log10 scale)
N50 = 1768 bp

Table C.2 Top 125 most highly expressed geen transcripts with their closest homolog species in *N. pseudonarcissus* 'King Alfred' bulb transcriptome.

Rank	FPKM	Annotation	Species name
1 to 4	5441800	186K protein, protein P1, P2 and 2A	Narcissus mosaic virus strain New Zealand
5	229410.45	metallothionein-like protein	Arachis hypogaea
6	206386	peroxidase 42-like	Eucalyptus grandis
7	199053	hypothetical protein BVRB_036970	Beta vulgaris subsp. Vulgaris
8	139000	sucrose synthase 1-like	Elaeis guineensis
9	128069	polyprotein	Artichoke latent virus
10	108077	alpha-tubulin	Moringa oleifera
11	99080	probable 2-oxoglutarate-dependent dioxygenase At3g49630	Elaeis guineensis
12	89647.8	YTH domain-containing family protein 2-like	Elaeis guineensis
13	85872	metallothionein-like protein	Hyacinthus orientalis
14	85307	no match	
15	72062.52	hypothetical protein PRUPE_ppa012400mg	Prunus persica
16	69637.96	Bifunctional 6(G)-fructosyltransferase/2,1-fructan:2,1-fructan 1-fructosyltransferase	Allium cepa
17	66209.41	vacuolar-processing enzyme	Elaeis guineensis
18	65736.76	hypothetical protein CISIN	Citrus sinensis
19	65434	(2Fe-2S)-binding protein	Sulfitobacter donghicola
20	64944.19	catalase isozyme A	Musa acuminata subsp. Malaccensis
21	64724	hypothetical protein UU56_C0010G0013	Microgenomates (Curtissbacteria) bacterium
22	60238	unnamed protein product	Coffea canephora
23	60009	ABC transporter F family member 1-like	Phoenix dactylifera
24	58301	hypothetical protein OsI_37065	Oryza sativa Indica Group
25	55312	glycine-rich RNA-binding protein GRP2A	Vitis vinifera

26	54721	glyceraldehyde 3-phosphate dehydrogenase	Daucus carota
27	54490	probable aquaporin PIP1-2	Musa acuminata subsp. Malaccensis
28	53525.66	cathepsin B-like	Musa acuminata subsp. malaccensis
29	48694	no match	
30	48584	no match	
31	46591	Aspartic proteinase nepenthesin-1 precursor, putative	Ricinus communis
32	46472	cleft lip and palate transmembrane protein 1 homolog	Phoenix dactylifera
33	46325	ATPase	Achromobacter sp.
34	44628	HMG-domain containing protein	Narcissus pseudonarcissus
35	44009	SAM-synthetase	Cicer arietinum
36	43835	enolase	Populus euphratica
37	43141	hypothetical protein FG05_05736	Fusarium graminearum
38	42965	probable aquaporin TIP1-1	Phoenix dactylifera
39	42852	granule-bound starch synthase 1, chloroplastic/amyloplastic-like	Elaeis guineensis
40	42843	adenosylhomocysteinase-like	Phoenix dactylifera
41	42691.86	no match	
42	42304	sucrose synthase 1	Elaeis guineensis
43	40806	ADP-ribosylation factor	Erythranthe guttata
44	40737	ATP synthase subunit beta, mitochondrial	Elaeis guineensis
45	40496.89	nonspecific lipid transfer protein	Vitis pseudoreticulata]
46	40320	elongation factor 2	Musa acuminata subsp. malaccensis
47	40093.81	heat shock protein 90-2	Tarenaya hassleriana
48	38591	replication-associated protein A	Peptoclostridium difficile CD44
49	34429	uncharacterized protein LOC103702207	Phoenix dactylifera
50	34251	60S acidic ribosomal protein P0-like	Musa acuminata subsp. Malaccensis
51	34104	DEAD-box ATP-dependent RNA helicase 7-like	Musa acuminata subsp. Malaccensis

52	34005	glycine-rich RNA-binding protein GRP2A	Vitis vinifera
53	33775	uncharacterized protein LOC105044022	Elaeis guineensis
54	33570	no match	
55	33203	vignain-like	Phoenix dactylifera
56	32626	UDP-glucuronate 4-epimerase 6	Vitis vinifera
57	32278	sucrose 1-fructosyltransferase	Allium cepa
58	32251	60S ribosomal protein L3	Ornithogalum longebracteatum
59	32108	protease inhibitor/seed storage/lipid transfer protein	Hyacinthus orientalis
60	31953.28	translationally-controlled tumor protein homolog	Oryza brachyantha
61	31581	ADP,ATP carrier protein 1, mitochondrial-like	Malus domestica
62	31543.34	S-norcoclaurine synthase 1-like	Phoenix dactylifera
63	31432.98	dimeric mannose specific lectin protein precursor	Narcissus hybrid cultivar 2
64	31350	probable mannose-1-phosphate guanylyltransferase 1	Elaeis guineensis
65	31335	histone H1	Zea mays
66	31269	ruBisCO large subunit-binding protein subunit beta, chloroplastic	Phoenix dactylifera
67	31214	GDSL-motif lipase	Agave americana
68	30708	MULTISPECIES: hypothetical protein	bacteria
69	30518	hypothetical protein M569_09780	Genlisea aurea
70	30159	variant of histone H1	Lilium longiflorum
71	29601.89	caffeic acid O-methyltransferase-like protein mRNA, complete cds	Narcissus tazetta cultivar Huanghua No.2
72	29577	ADP-ribosylation factor 2 (LOC109851241), mRNA	Asparagus officinalis
73	29483	oryzain alpha chain-like (LOC109841121), mRNA	Asparagus officinalis
74	29244	elongation factor 1-beta (LOC101493594), mRNA	Cicer arietinum
75	29164	ubiquitin-conjugating enzyme mRNA, complete cds	Lycoris longituba

76	29109	polyadenylate-binding protein RBP45-like (LOC109833544), mRNA	Asparagus officinalis
77	28815	pectin acetylesterase 7-like (LOC109839624), transcript variant X2, mRNA	Asparagus officinalis
78	27925	elongation factor 1-alpha (LOC100246711), mRNA	Vitis vinifera
79	27903	60S ribosomal protein L5 (LOC109837563), mRNA	Asparagus officinalis
80	27631	chitinase-like protein 1 (LOC103992403), mRNA	Musa acuminata subsp. malaccensis
81	27532	no match	
82	27364.72	heat shock protein 90 (HSP90) mRNA, complete cds	Ornithogalum longebracteatum
83	27219.95	heat shock cognate 70 kDa protein 2-like (LOC109848789), mRNA	Asparagus officinalis
84	27027	26S proteasome regulatory subunit 4 homolog B (LOC109841322), mRNA	Asparagus officinalis
85	26749	putative iron/ascorbate oxidoreductase mRNA mRNA, partial cds	Narcissus pseudonarcissus
86	26718	metallothionein-like protein type 2 (LOC110108418), mRNA	Dendrobium catenatum
87	26525.64	calnexin homolog (LOC109824517), mRNA	Asparagus officinalis
88	26417.51	uncharacterized LOC103985898 (LOC103985898), transcript variant X2, ncRNA	Musa acuminata subsp. malaccensis
89	26344	myosin-2 heavy chain-like (LOC109819892), transcript variant X2, mRNA	Asparagus officinalis
90	26327.73	xyloglucan endotransglucosylase/hydrolase protein 23 (LOC103976880), mRNA	Musa acuminata subsp. malaccensis
91	26161.89	no match	
92	25959	60S ribosomal protein L10 partial mRNA	Morus notabilis
93	25756	methyltransferase PMT26 (LOC109837966), mRNA	Asparagus officinalis
94	25549	no match	

95	25507	60S ribosomal protein L6-2-like (LOC109837539), mRNA	Asparagus officinalis
96	25268	no match	
97	25216	eukaryotic initiation factor 4A-8 (LOC109828423), mRNA	Asparagus officinalis
98	24687	MBK-2015 vittatine 11-hydroxylase mRNA, complete cds	Galanthus sp.
99	24672.26	no match	
100	24436	40S ribosomal protein S6 (LOC109833973), transcript variant X2, mRNA	Asparagus officinalis
101	24414	cellulose synthase A catalytic subunit 5 [UDP-forming] (LOC109843790), mRNA	Asparagus officinalis
102	24382	soluble starch synthase 2-2, chloroplastic/amyloplastic (LOC109833869), mRNA	Asparagus officinalis
103	24107	GTP-binding protein SAR1A (LOC109819697), mRNA	Asparagus officinalis
104	23974	heterogeneous nuclear ribonucleoprotein 1-like (LOC109360644), mRNA	Lupinus angustifolius
105	23925	no match	
106	23918.91	thiamine thiazole synthase 2, chloroplastic (LOC107488645), mRNA	Arachis duranensis
107	23541	reticulon-like protein B3 (LOC105041675), mRNA	Elaeis guineensis
108	23366	histone H2A.4 (LOC103850452), mRNA	Brassica rapa
109	23326	alpha-amylase 3, chloroplastic (LOC109820177), transcript variant X4, mRNA	Asparagus officinalis
110	23206.95	60S ribosomal protein L8 (LOC101785583), mRNA	Setaria italica
111	23188	no match	
112	23186	V-type proton ATPase subunit c1 (LOC109825900), mRNA	Asparagus officinalis
113	23100	aquaporin PIP2-6 (LOC103994286), mRNA	Musa acuminata subsp. malaccensis

114	22719	guanine nucleotide-binding protein subunit beta-like protein A (LOC105041173), mRNA	Elaeis guineensis
115	22698	putative cysteine proteinase mRNA, partial cds	Narcissus pseudonarcissus
116	22584.91	fructose-bisphosphate aldolase 1, cytoplasmic (LOC109836704), mRNA	Asparagus officinalis
117	22513	pyrophosphate-energized vacuolar membrane proton pump-like (LOC109841169), mRNA	Asparagus officinalis
118	22463	cytosolic isocitrate dehydrogenase [NADP]-like (LOC109840741), mRNA	Asparagus officinalis
119	22356	vacuolar-sorting receptor 1-like (LOC109851365), mRNA	Asparagus officinalis
120	22266	serine hydroxymethyltransferase 4-like (LOC103696913), mRNA	Phoenix dactylifera
121	22258	coatomer subunit gamma-2 (LOC109833116), mRNA	Asparagus officinalis
122	21960	annexin D2-like (LOC109833942), mRNA	Asparagus officinalis
123	21777	uncharacterized protein At1g04910-like (LOC109832548), mRNA	Asparagus officinalis
124	21747	cellulose synthase A catalytic subunit 1 [UDP-forming] (LOC109828772), mRNA	Asparagus officinalis
125	21722	norbelladine 4'-O-methyltransferase (N4OMT3)	Narcissus aff. pseudonarcissus MK-2014
sum 125 top	10638280.3		
sum total surviving reads	66054792		
% top 125	16.1052362		
% virus	8.23831222		

Table C.3 List of real time qPCR primers used in this study

Gene	Tm	Forward primers (5′-3′)	Reverse primers (5′-3′)
TYDC1	56°C	TTGGTGTCGGCGGATGAG	TGCTTCTCCTCCGTGAGG
TYDC2	56°C	GTGCCAGGCGTAACCCATTGG	CTGCCCAGAAGACAGGAATTGG
PAL1	52.6°C	GAGGAGAATCTGAAGAGC	CATCAAAGGGTAAGTGC
PAL2	56°C	CAAGCTGTCAAAAACACGGTG	AAGGGTCGTCGATGTACGTG
C4H	53°C	GTATTTGAAGATTTGCAAGC	GCTGCAACATTGATGTTCTC
C3H	56°C	CAGGTGCTTCGCCGAGTGG	CCTCACCTTCACGTAGTGGG
4CL	57°C	GGCTATTTGAATGACCCTGAGG	CCGAGCTTTTCGACGGAACTAC
HCT	53.2°C	CGACGAGGAGTATGTGAGG	CTCCAACGAAACCAGAACC
N4OMT	55°C	GTGAAGCTCGTCAGGATTGG	CAGGTTTCTGAAAGAGAGCC
CYP96T	57°C	TGCTATGGCGAGGATGAAGG	ACATGTCCCTTCACCATCTG
NorRed	57°C	TCCGGGAGCCATAAGAACGC	GGTGATGTAGGAAGCAGATGG
His	55°C	GTCTGCCCCAACAACTGGAGG	GCTTCCTAATCAGTAGCTCG

ANNEX D

SUPPLEMENTARY DATA FOR CHAPTER III

Table D.1 Sequences for NCS and PR-10 used in phylogenetic analysis

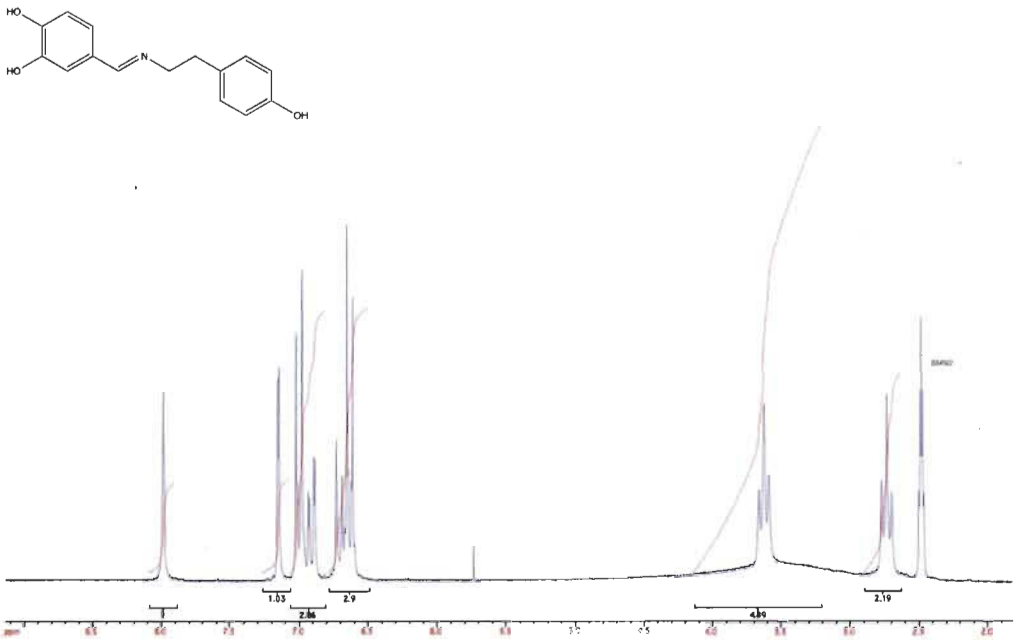
Short name	Species	Reference/Accession number
AmNCS1	*Argenome mexicana*	(Li, 2016)
CjNCS2	*Coptis japonica*	(Li, 2016)
CmNCS1	*Chelidonium majus*	(Li, 2016)
DcPR10	*Daucus carota*	(Sano, 2004)
EcNCS1	*Eschscholzia californica*	(Li, 2016)
EcNCS2	*Eschscholzia californica*	(Li, 2016)
HoPR10	*Hyacinthus orientalis*	(Liscombe, 2005)
PmPR10	*Pinus monticola*	(Liu, 2003)
StPR10	*Solanum tuberosum*	(Matton, 1989)
TfNCS1	*Thalictrum flavum*	(Li, 2016)
TfNCS2	*Thalictrum flavum*	(Li, 2016)
TfNCS3	*Thalictrum flavum*	(Li, 2016)
TfNCS4	*Thalictrum flavum*	(Li, 2016)
TfNCS5	*Thalictrum flavum*	(Li, 2016)
PsNCS1	*Papaver somniferum*	(Liscombe, MacLeod et al. 2005)
PsNCS2	*Papaver somniferum*	(Samanani and Facchini 2001)
BpPR10	*Betulla platyphylla*	(Li, 2016)
BvPR10	*Betulla verrucosa*	(Liu, 2006)
ZmPR10	*Zea mays*	(Schnable, Ware et al. 2009)
HvPR10	*Hordeum vulgare*	(Steiner-Lange, Fischer et al. 2003)
OsPR10	*Oryza sativa*	Unpublished (ACA50491)
LrPR10	*Lily regale*	Unpublished (ARX80137)
SbPR10	*Sorghum bicolor*	Unpublished (AAW83207)
PsMLP	*Papaver somniferum*	(Nessler and Burnett 1992)
HYP-1	*Hypericum perforatum*	(Michalska, Fernandes et al. 2010)
NpNBS	*Narcissus pseudonarcissus 'King Alfred'*	This study

Table D.2 List of Primer sequences used in this study

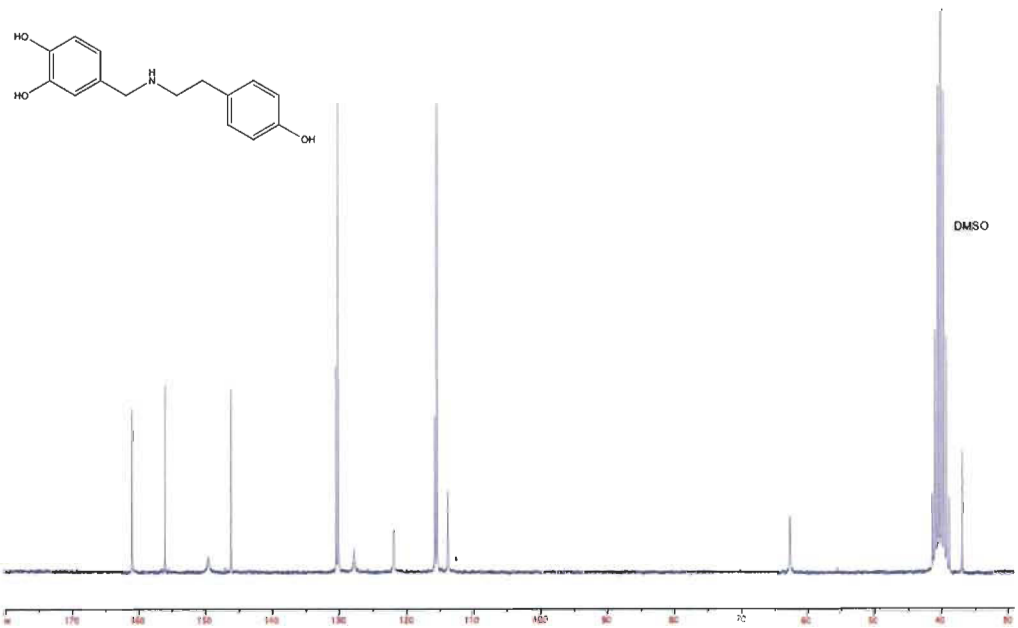
Primer name	**Primer Sequence**
*NpNBS*_GAT_F	5'-GGGGACAAGTTTGTACAAAAAAGCAGGCT AGAAGGAGATATACATATGAAGGGAAGTCTCTCC-3'
*NpNBS*_GAT_R	5'-GGGGACCACTTTGTACAAGAAAGCTGGGTA CAATAACATTAT GCTACAG-3'
*NpNBS*_qRT-PCR_F	5'-GAGTTGGAGGTTTCC TTGC-3'
*NpNBS*_qRT-PCR_F	5'-CCACCATCACCTTCCTCG-3'
*NpHIS*_ qRT-PCR_F	5'-GTCTGCCCCAACAACTGGAGG-3'
*NpHIS*_ qRT-PCR_F	5'-GCTTCCTAATCAGTAGCTCG-3'

Figure D.1 Proton and carbon NMR spectral data of newly synthesized norcraugsodine and norbelladine

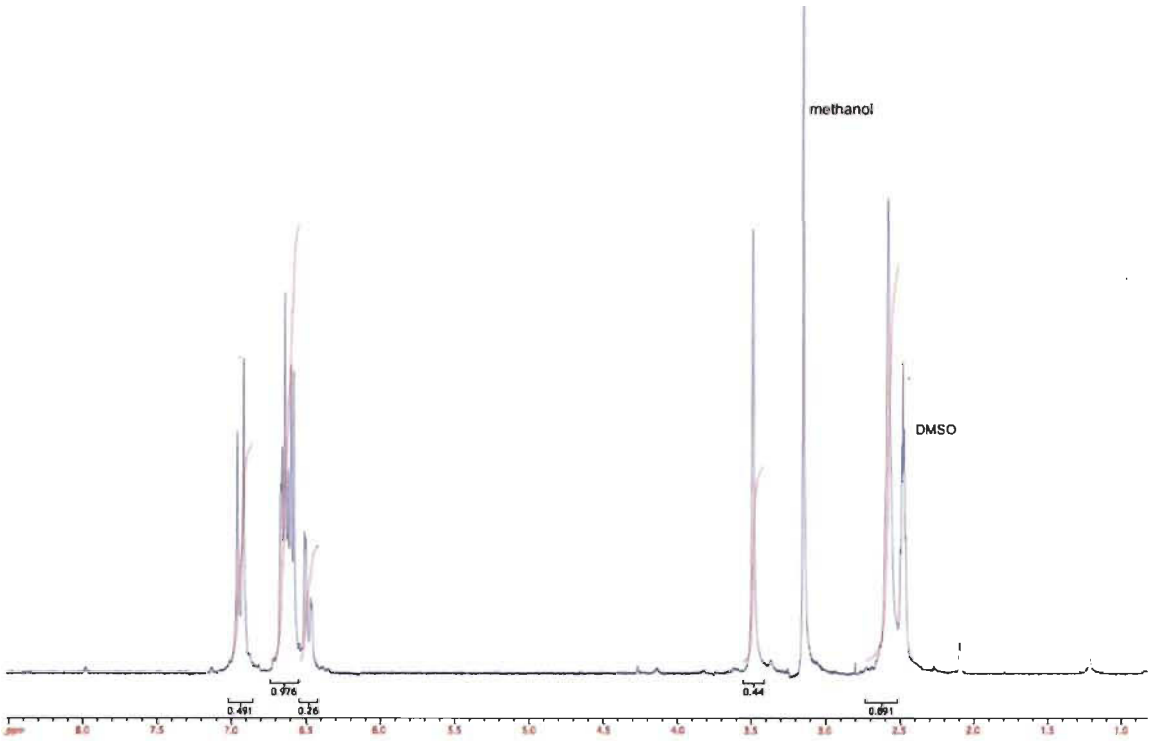
A) Norcraugsodine - Imine compound ^{1}H NMR spectrum



B) Norcraugsodine - Imine compound ^{13}C NMR spectrum



C) Norbelladine ^{1}H NMR spectrum



D) Norbelladine ^{13}C NMR spectrum

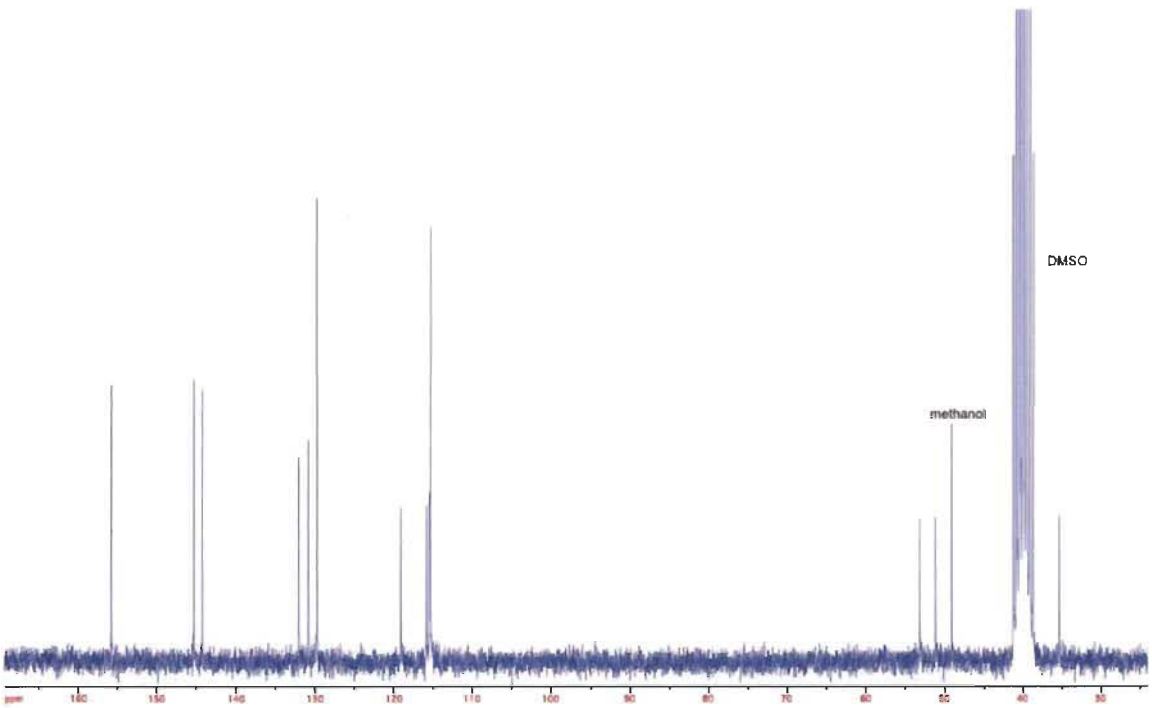
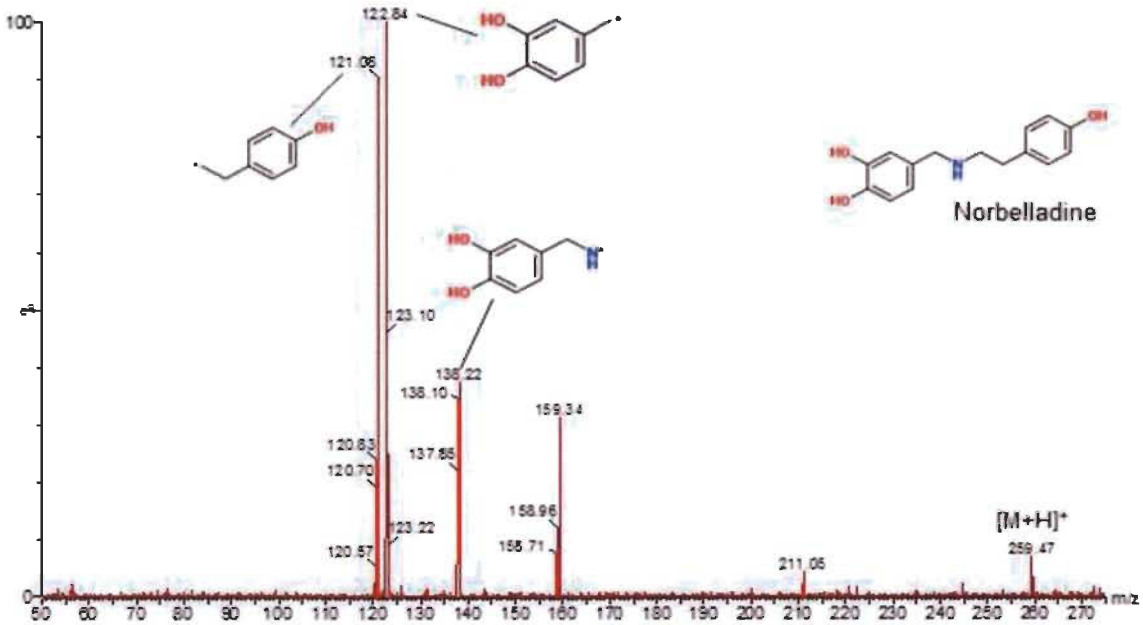
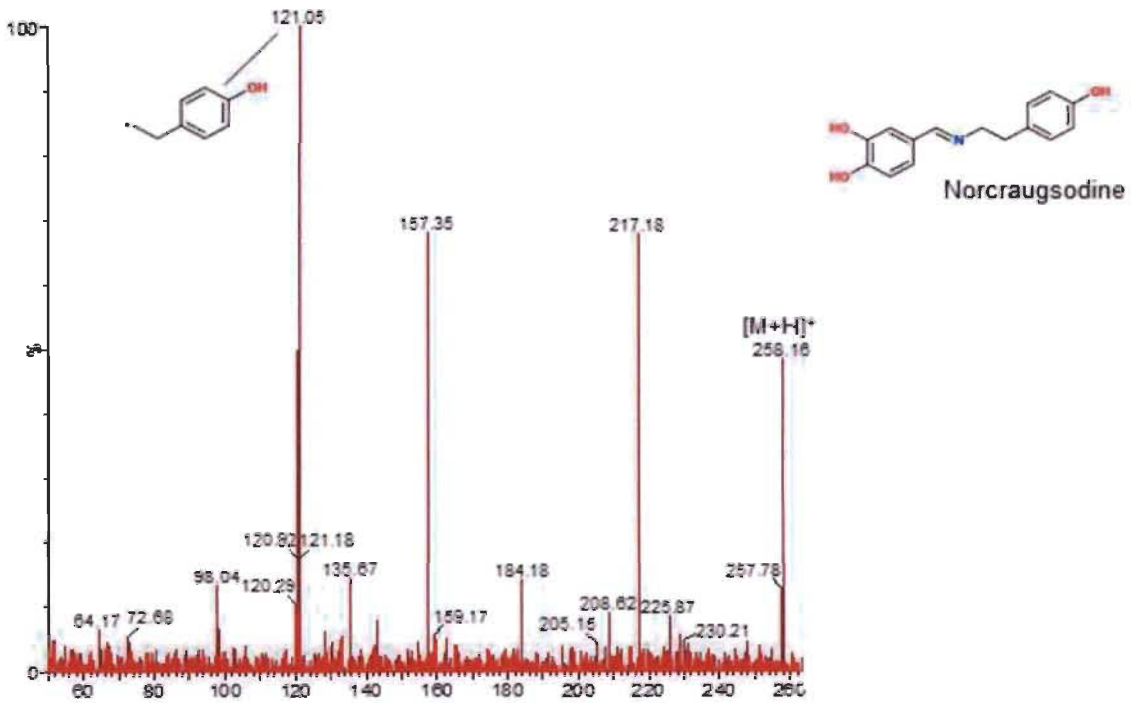


Figure D.2 Fragmentation spectra obtained from LC-MS/MS analysis of standards norbelladine and norcraugsodine followed by a table listing the parameters used for LC-MS/MS analysis.

NORBELLADINE (+ESI-MS m/z 260)



NORCRAUGSODINE (+ESI-MS m/z 258)



Compound	$[M+H]^+$	RT (min)	VC (V)	CE daughters (V)	Daughters, *m/z* (relative abundance)	Qualifier CE (V)	Qualifier transition	Quantifier CE (V)	Quantifier transition
Norbelladine	260	5.5	20	15	260 (8), 159 (30), 138 (39), 123 (100), 121 (90)	15	260→138	20	260→121
Norcraugsodine	258	5.5	20	10	258 (50), 217 (68), 157 (69), 121 (100)	15	258→157	20	258→121