A multifaceted approach to identify unknown enzymes of pyrrolizidine alkaloid biosynthesis

by Lars Hendrik Kruse

Dissertation in fulfillment of Requirements of the Doctoral Degree

Doctor rerum naturalium

of the Faculty of Mathematics and Natural Sciences at

Christian-Albrechts-Universität zu Kiel

Kiel, April 2017

Lars Hendrik Kruse

A multifaceted approach to identify unknown enzymes of pyrrolizidine alkaloid biosynthesis

Dissertation, Faculty of Mathematics and Natural Sciences
Christian-Albrechts-Universität zu Kiel

Erstgutachter: Prof. Dr. Dietrich Ober
Zweitgutachter: Prof. Dr. Axel Scheidig
Tag der mündlichen Prüfung: 20.06.2017
Zum Druck genehmigt: 20.06.2017

ABSTRACT

Pyrrolizidine alkaloids (PAs) are anti-herbivorous toxins that are found in the Asteraceae, Boraginaceae, Convolvulaceae, Fabaceae, Orchidaceae, and Poaceae. Interestingly, the PA biosynthesis evolved several times independently in the different plant lineages. The first pathway-specific enzyme, the homospermidine synthase (HSS), evolved repeatedly by gene duplication from the primary metabolism gene deoxyhypusine synthase (DHS). Remarkably, although the independent evolution resulted in very similar PA structures, the integration into the plant's pathway was different in the various lineages. Inter alia, this is shown by the different regulation of PA biosynthesis in the plant, e.g., even in the same taxonomic order, the site of PA biosynthesis varies between different species.

In this thesis, a further facet was added to the versatile regulation of the PA pathway. It was known that comfrey (*Symphytum officinale* L.) produces PAs in the root. This work shows that *Symphytum* activates a second site of PA biosynthesis in young leaves next to developing inflorescences to boost defense against herbivores. According to the optimal defense theory, reproductive tissues have a high value for the plant and need good protection. The finding that *Symphytum* has two sites where PAs are produced, was then used to identify further enzymes of the pathway. For a substractive approach, three tissues were chosen for transcriptome sequencing (young leaf that produces PAs, young leaf that produces no PAs, and roots that also produce PAs). Subsequently, the transcript abundance was measured *in silico* and transcripts that were only expressed in PA-producing tissues were enriched. By searching the enriched transcripts, I was able to identify diamine oxidases (DAOs) that are postulated to catalyze the second step of the PA pathway, the oxidation of homospermidine. The DAO transcripts were further analyzed on their occurrence in various tissues and cell types, protocols for the heterologous expression in *Escherichia coli* were established, and it was tried to measure their enzyme activity with the postulated substrate homospermidine. The results indicate that two different DAOs are involved in PA biosynthesis, one in the leaves and the other one in roots. To test for the involvement of DAOs, and further identified candidate genes, in PA biosynthesis *in vivo*, RNA silencing was applied. In a proof-of-concept experiment, HSS was used as a target for RNAi-mediated gene-knockdown and it was successfully shown that a reduced level of HSS transcripts leads to a reduced level of PAs. These results encourage further RNAi experiments to test the function of other identified candidates.

In summary, this work provides novel knowledge about the regulation of PA

biosynthesis in *Symphytum* and adds another, unique facet to the convergent evolution of PAs. In addition, it was shown for the first time *in vivo*, that the HSS is truly essential for PA biosynthesis. Furthermore, the data and methods that were developed during this thesis will be a valuable and helpful source for the ongoing research on PA biosynthesis and the involved enzymes.

KURZFASSUNG

Pyrrolizidinalkaloide (PAs) sind giftige Abwehrstoffe von Pflanzen und kommen unter anderem in den Familien Asteraceae, Boraginaceae, Convolvulaceae, Fabaceae, Orchidaceae und Poaceae vor. Die Biosynthese von PAs in Pflanzen ist insofern besonders interessant, da sie im Laufe der Evolution mehrfach unabhängig voneinander entstanden ist. Das Eingangsenzym des Stoffwechselweges, die Homospermidine-Synthase (HSS), ist mehrfach in verschiedenen Pflanzen durch eine Genduplikation aus dem Primärstoffwechselenzym Desoxyhypusine-Synthase (DHS) hervorgegangen. Bemerkenswerterweise führte die Evolution der PA-Biosynthese zu in Ihrer Grundstruktur fast gleichen Metaboliten, die Integration in den pflanzlichen Stoffwechsel scheint allerdings jedes mal unterschiedlich erfolgt zu sein. Dies zeigt sich unter anderem an der unterschiedlichen Lokalisation der PA-Biosynthese in den verschiedenen Spezies. Sogar innerhalb einer Ordnung (z.B. Boraginales) kann der Ort der Biosynthese stark variieren.

In dieser Arbeit konnte eine weitere Facette der vielfältigen Regulation der Synthese pflanzlicher Sekundärstoffe hinzugefügt werden. Bekannt war, dass im Beinwell (*Symphytum officinale*) PAs in der Wurzel produziert werden. Neu ist, dass die Pflanze einen weiteren Ort der Biosynthese in jungen Blättern aktivieren kann, um den Schutz von neuen Blütenständen zu verbessern. Reproduktive Gewebe sind nach der sog. *optimal-defence-theory* besonders wertvoll für die Pflanze. Dass *Symphytum* in zwei verschiedenen Pflanzengeweben PAs produziert, wurde in einem folgendem Schritt für die Identifikation von weiteren Enzymen dieses Stoffwechselweges ausgenutzt. Für einen subtraktiven Ansatz wurden drei Gewebe (junge Blätter mit, junge Blätter ohne und Wurzeln mit PA-Biosynthese) ausgewählt, deren Transkriptom sequenziert und anschließend die Transkriptabundanz bestimmt. Anschließend wurden nur jene Transkripte *in silico* angereichert, die nur in PA-produzierenden Geweben vorkommen. Dazu gehörten auch Transkripte von Diaminoxidasen (DAOs), die nach den Ergebnissen von Tracer-Fütterungs-experimenten, Inhibitorversuchen und biomimetischen Experimenten für den zweiten Schritt der PA-Biosynthese, die Oxidation von Homospermidin, postuliert werden. Daher wurden Transkripte der DAOs in weiteren Schritten genauer auf ihr Vorkommen in verschiedenen Geweben und Zellen untersucht. Außerdem wurden Protokolle für die heterologe Expression in *Escherichia coli* etabliert, um die Enzymaktivität mit dem Substrat Homospermidin zu bestimmen. Die Ergebnisse sprechen dafür, dass zwei verschiedene DAOs an der PA-Biosynthese beteiligt sind,

eine in Blättern und die andere in Wurzeln. Um die DAOs und weitere Kandidatenenzyme *in vivo* auf Ihre Funktion in der PA-Biosynthese zu testen, wurde in einer weiteren Studie RNA-Inteferenz verwendet. In einem *proof-of-concept*-Experiment wurde die HSS als Ziel für einen RNAi-vermittelten Gen-*knockdown* ausgewählt. Es konnte gezeigt werden, dass eine reduzierte HSS-Transkriptmenge auch eine reduzierte PA-Menge in den Mutanten zur Folge hat. Dies sind vielversprechende Ergebnisse, die für weitere solcher Experimente mit anderen Kandidatenenzymen sprechen.

Diese Arbeit bietet neue Erkenntnisse über die Regulation der PA-Biosynthese in *Sympyhtum officinale* und fügt dem Aspekt der konvergenten Evolution der PAs eine weitere, einzigartige Facette hinzu. Außerdem konnte zum ersten Mal *in vivo* gezeigt werden, dass die HSS tatsächlich essentiell für die PA-Biosynthese ist.

CONTENTS

1. INTRODUCTION....1
1.1. The Diversity and Evolution of Plant Toxins....1
1.2. Pyrrolizidine Alkaloids and their Toxicity....3
1.3. The Biosynthesis of Pyrrolizidine Alkaloids....6
1.4. PAs in the Boraginaceae and Comfrey....7
2. OBJECTIVES AND OUTLINE OF THIS WORK....9
3. CHAPTER 1....11
3.1. ABSTRACT....12
3.2. INTRODUCTION....13
3.3. RESULTS....15
3.3.1 Levels of *hss* Transcript and Protein Expression are Dependent on Leaf Position in Relation to the Inflorescence....15
3.3.2 Immunolocalization of HSS in Leaves of *S. officinale*....17
3.3.3 Capacity of *S. officinale* Leaves to Synthesize PAs....18
3.3.4 Leaves Boost PA Levels in Inflorescences....20
3.4. DISCUSSION....23
3.5. EXPERIMENTAL....25
3.5.1 Plant Material....25
3.5.2 RNA Isolation and Quantification of Transcripts....25
3.5.3 Protein-Blot Analysis of *S. officinale* Leaves....27
3.5.4 Immunohistochemical Staining of HSS in Leaf Cross Sections....27
3.5.5 Radioactive Tracer-Feeding Experiments and Product Analysis....28
3.5.6 Quantification of total PAs in inflorescences....29
3.5.7 Accession Numbers....29
3.6. Acknowledgments....29
3.7. Supplementary material....30
4. CHAPTER 2....37
4.1. ABSTRACT....38
4.2. INTRODUCTION....39
4.3. RESULTS....42
4.3.1 Study design, extraction of total RNA and generation of RNA-Seq data.42
4.3.2 Optimization and evaluation of transcriptome assembly....43
4.3.3 *In silico* substractive approach....47
4.3.4 Identification of candidate genes for postulated steps of PA biosynthesis 48
4.3.5 Transcription levels of DAO genes in PA producing tissues....50
4.3.6 Sequence similarity and functional amino acids....52
4.3.7 Heterologous expression of diamine oxidases....54
4.4. DISCUSSION....55
4.4.1 Different scenarios of PA pathway evolution require different subsets for candidate gene identification....55
4.4.2 Two putative homospermidine oxidases in different organs?....56
4.4.3 Is an enzyme involved in cyclization of the necine base?....57
4.4.4 Candidate sequences for other steps of the PA biosynthesis pathway...58
4.4.5 Summary and Outlook....59
4.5. EXPERIMENTAL....60
4.5.1 Plant material....60

4.5.2 RNA extraction and quantitative real time PCR (qRT-PCR).....60
4.5.3 Laser-assisted microdissection.....62
4.5.4 High throughput sequencing (Illumina®) and bioinformatic analysis.....62
4.5.5 Identification of candidate sequences.....63
4.5.6 Cloning of putative DAO encoding sequences.....63
4.5.7 Heterologous expression and detection of biochemical activity of putative diamine oxidases.....67
4.6. Acknowledgments.....68
4.7. Supplementary material.....69
5. CHAPTER 3.....73
5.1. ABSTRACT.....74
5.2. INTRODUCTION.....75
5.3. RESULTS.....77
5.3.1 Establishment of RNAi knock down mutants.....77
5.3.2 HSS is functionally relevant for PA biosynthesis.....78
5.4. DISCUSSION.....83
5.4.1 PA content of hairy roots of *S. officinale*.....83
5.4.2 RNAi mediated down-regulation of HSS leads to reduced PA levels.....84
5.4.3 Perspectives for future studies.....84
5.4.4 Summary.....85
5.5. EXPERIMENTAL.....87
5.5.1 Plant material.....87
5.5.2 Cloning of RNAi constructs.....87
5.5.3 Transformation of *Agrobacterium rhizogenes*.....87
5.5.4 Generation of transgenic hairy root mutants.....88
5.5.5 Hairy root growth and experimental setup.....89
5.5.6 Quantification of PAs.....89
5.5.7 Identification of the quantified PAs by GC-MS.....90
5.5.8 RNA extraction and quantitative real time PCR (qRT-PCR).....91
5.5.9 Statistical analyses.....93
5.6. Acknowledgments.....93
5.7. Supplementary material.....94
6. CONCLUSIONS AND GENERAL PERSPECTIVES.....95
7. OUTLOOK.....99
7.1. PA biosynthesis in comfrey (*Symphytum officinale*).....99
7.2. Heterologous expression of DAOs and biochemical characterization.....100
7.3. Localization of HSS and DAOs by fusion with fluorescent proteins.....101
7.4. Generation of hairy root DAO knockdown mutants.....102
CONTRIBUTIONS.....103
DANKSAGUNG & ACKNOWLEDGEMENTS.....106
REFERENCES.....109
LEBENSLAUF.....127

1. INTRODUCTION

1.1. The Diversity and Evolution of Plant Toxins

Plants produce an exceptional diversity of specialized metabolites to cope with various conditions in their specific environment. Specialized metabolites are organic compounds with low molecular weight that are found only in a specific plant lineage (Fraenkel, 1959). Until today, about 250,000 plant species are recognized and the number of specialized compounds identified is estimated to be more than 200,000 (Pichersky and Gang, 2000; Dixon and Strack, 2003; Pichersky and Lewinsohn, 2011). Those metabolites are traditionally referred as secondary metabolites because of their often undescribed function and the former belief to be only waste products (Fraenkel, 1959). But today, most of these compounds are known to be products of highly specialized traits that evolved during plant's adaption to their specific ecological niche and thus are termed as specialized metabolites. Co-evolution with insects is thought to be one large factor for the rise of chemo-diversity in plants, for example the interaction with herbivores or pollinators (Ehrlich and Raven, 1964; Romeo et al., 1996). Especially the continuous selective pressure of herbivores on plants led to diversity, complexity and optimization of toxic secondary metabolites (Berenbaum and Feeny, 1981; Becerra, 2003; Benderoth et al., 2006; Berenbaum and Zangerl, 2006; Becerra, 2007). But questions raised how it was possible that such a diversity of metabolites evolved and how this number of different enzymes

emerged. It was hypothesized that gene duplication might be a mechanism for the evolution of new enzymes with new functions (Ohno, 1970; Ohta, 1988). In general, gene duplicates are the result of duplications of a region of genomic DNA that contains a gene. They can arise through different mechanisms, like unequal crossing over, retroposition, chromosome duplication, or whole genome duplication (Zhang, 2003). From an evolutionary perspective, different models have been developed that describe the fate of two gene copies after the duplication event (Innan and Kondrashov, 2010). (1) One fate of gene duplicates (that is believed to be the most common) is nonfunctionalization (Conant and Wolfe, 2008). In this scenario, two gene copies are redundant after the duplication and one of the two copies is no longer under functional constraints. Subsequently it can accumulate mutations until deleterious mutations result in a loss-of-function, a process called pseudogenisation (Conant and Wolfe, 2008). (2) In the case that accumulated mutations led to a new function in one of the two copies, while the other retains the ancestral function, this is called neofunctionalisation (Conant and Wolfe, 2008). (3) If the ancestor of the two gene copies already had more than one function and after the duplication, each of the two copies specializes by selection on one of the ancestral functions, this mechanism is called subfunctionalization (Conant and Wolfe, 2008). The models describing the evolution by gene duplication are under constant discussion and several theories try to explain the dynamics after gene duplications, e.g., paralog interferences, dosage-balance hypothesis, duplication-degeneration-complementation, and escape-from-adaptive-conflict (Stoltzfus, 1999; Papp et al., 2003; He and Zhang, 2005; Innan and Kondrashov, 2010; Kaltenegger and Ober, 2015). Today it is believed, that the evolution of new enzymes in plants that lead to the synthesis of new metabolites, is to a large extent driven by gene duplication (Pichersky and Gang, 2000; Conant and Wolfe, 2008; Innan and Kondrashov, 2010; Rensing, 2014; Panchy et al., 2016). These specialized enzymes have often been recruited from enzymes of the primary metabolism but also genes of secondary metabolism are described to give rise to new enzymes after duplication (Pichersky and Gang, 2000; Ober, 2005). Interestingly, many of those duplicated genes belong to only a few major enzyme families i.e. *N*- and *O*-methyltransferases, cytochrome-P450-dependent monooxygenases, 2-oxoglutarate-dependent dioxygenases, a/b -hydrolases and terpene synthases (Bohlmann et al., 1998; Hashimoto and Yamada, 2003). This can be interpreted as those enzymes act as a building block for new functions that are recruited as adaptations to new ecological conditions (Ober, 2005; Pichersky and

Lewinsohn, 2011). New functions are often transformations and rearrangements of existing structures, which can be observed in all living organisms from enzymes to whole organs, e.g., the skeleton of the vertebrates. Convergent evolution seems to be relatively common in plants and many examples exist where plants evolved molecules that were already present in other lineages, or evolved molecules that are different but fulfill the same function in another plant (Pichersky and Lewinsohn, 2011). Examples for different molecules with the same functions are red colored pigments (anthocyanins in strawberry, betains in pitaya fruit, and carotenoids in tomato) or yellow colored pigments in flowers (xanthocyanin in moss rose, aurone glucosides in the flower of snapdragon, and xanthophylls in tomato flowers) (reviewed in Pichersky and Lewinsohn, 2011). In other cases, the same molecules evolved independently in different lineages. A prominent example are the phenolic compounds stilbenes that are found in distant taxa like *Pinus*, *Arachis*, and *Vitis*, where the key enzyme stilbene synthase (STS) evolved repeatedly from the chalcone synthase (CHS) (Tropf et al., 1994; reviewed in Pichersky and Lewinsohn, 2011). Another example for a repeated evolution of the same molecules are the pyrrolizidine alkaloids, where the initial enzyme homospermidine synthase (HSS) evolved several times from the primary metabolism gene deoxyhypusine synthase (DHS) (Ober and Hartmann, 1999a; Anke et al., 2004; Reimann et al., 2004; Nurhayati and Ober, 2005; Nurhayati et al., 2009; Kaltenegger et al., 2013; Irmer et al., 2015). The evolution of PA biosynthesis is described in more detail in the following sections.

1.2. Pyrrolizidine Alkaloids and their Toxicity

Alkaloids are one important group of often toxic secondary metabolites and are optimized for the defense against herbivores (De Luca and Laflamme, 2001). They form a large group of secondary metabolites with more than 12,000 different structures (Mithöfer and Boland, 2012). Alkaloids have a huge variety of pharmaceutical activities that resulted from the continuous optimization during evolution to act at specific targets, i.e., receptors of animals (Zenk and Juenger, 2007). Therefore, they are widely used as stimulants, narcotics and poisons, e.g., caffeine, morphine and strychnine. A diverse and widely distributed group of alkaloids are the pyrrolizidine alkaloids (PAs). Pyrrolizidine alkaloids occur in a number of plant lineages, for example, Boraginaceae, Asteraceae, Convolvulaceae, Fabaceae, and also in the monocots, namely in families of Poaceae and Orchidaceae (Hartmann and

Witte, 1995; Koulman et al., 2008). The backbone of all PAs is composed of a *N*-containing double-ring system, the necine base, which precursor is 1-hydroxy-methylpyrrolizidine (Hartmann and Witte, 1995). This necine base is esterified with one or more necic acids forming different monoesters, open-chain or macrocyclic diesters (Hartmann and Witte, 1995). The variable arrangement of the different side chains of PAs lead to the large variety of more or less 400 known PAs (Stegelmeier et al., 1999; Roeder, 2000; Fu et al., 2004). Pyrrolizidine alkaloids are classified into seven different alkaloid types: senecionine, triangularine, lycopsamine, monocrotaline, phalaenopsine, and miscellaneous PAs (Fig. 1). The occurrence of many of these different PA types is plant lineage specific. Being responsible for chemical defense against herbivores, most of those PA types are known to be toxic (van Dam et al., 1995a; Hartmann and Witte, 1995; Fu et al., 2004). The toxicity of the most common 1,2 unsaturated PAs for animals is based on conversion by cytochrome P450 monooxygenases, which actual purpose is the detoxification of xenobiotics (Mattocks, 1968). Cytochrome P450 enzymes are involved in important phase I detoxification mechanisms of a broad variety of toxic substances (Heckel, 2014). In the case of PAs, these enzymes activate the toxicity by converting the PA to pyrrolic intermediates that can easily react with DNA and proteins resulting in the genotoxic, cancerogenic and hepatotoxic characteristics (Fu et al., 2004; Chen et al., 2010; Langel and Ober, 2011; Li et al., 2011; Macel, 2011). In addition to 1,2 unsaturated PAs, in Poaceae and Orchidaceae occur 1,2 saturated PA molecules. The biological function and toxicity of these PAs is not clear yet, because they cannot act via the described mechanism. Their occurrence in the reproductive tissues in orchid species suggests, that they also act as toxins or repellents for herbivores (Frölich et al., 2006). The toxicity of PAs is a serious threat to humans and livestock by different sources. Several medicinal plants that are used in traditional and conventional medicine contain PAs, e.g., *Tussilago farfara* (coltsfoot), *Petasites hybridus* (butterbur), *Borago officinales* (borage), and *Symphytum officinale* (comfrey) (Roeder, 1995). Another, even more important source for intoxication for humans or livestock are food or fodder contaminates that lead to a number of recognized cases, e.g., contaminations of hay by *Senecio jacobeae* (tansy ragwort) and a widespread poisoning in Tadjikistan by grain contaminated with *Heliotropium lasiocarpum* (Mayer and Lüthy, 1993; Wiedenfeld and Edgar, 2011; Jank and Rath, 2017). The development of new detection methods in the last period have led to increasing public interest to toxic PA contaminants in, e.g., milk, honey, herbal teas and herbal

medicines (Chen et al., 2010).

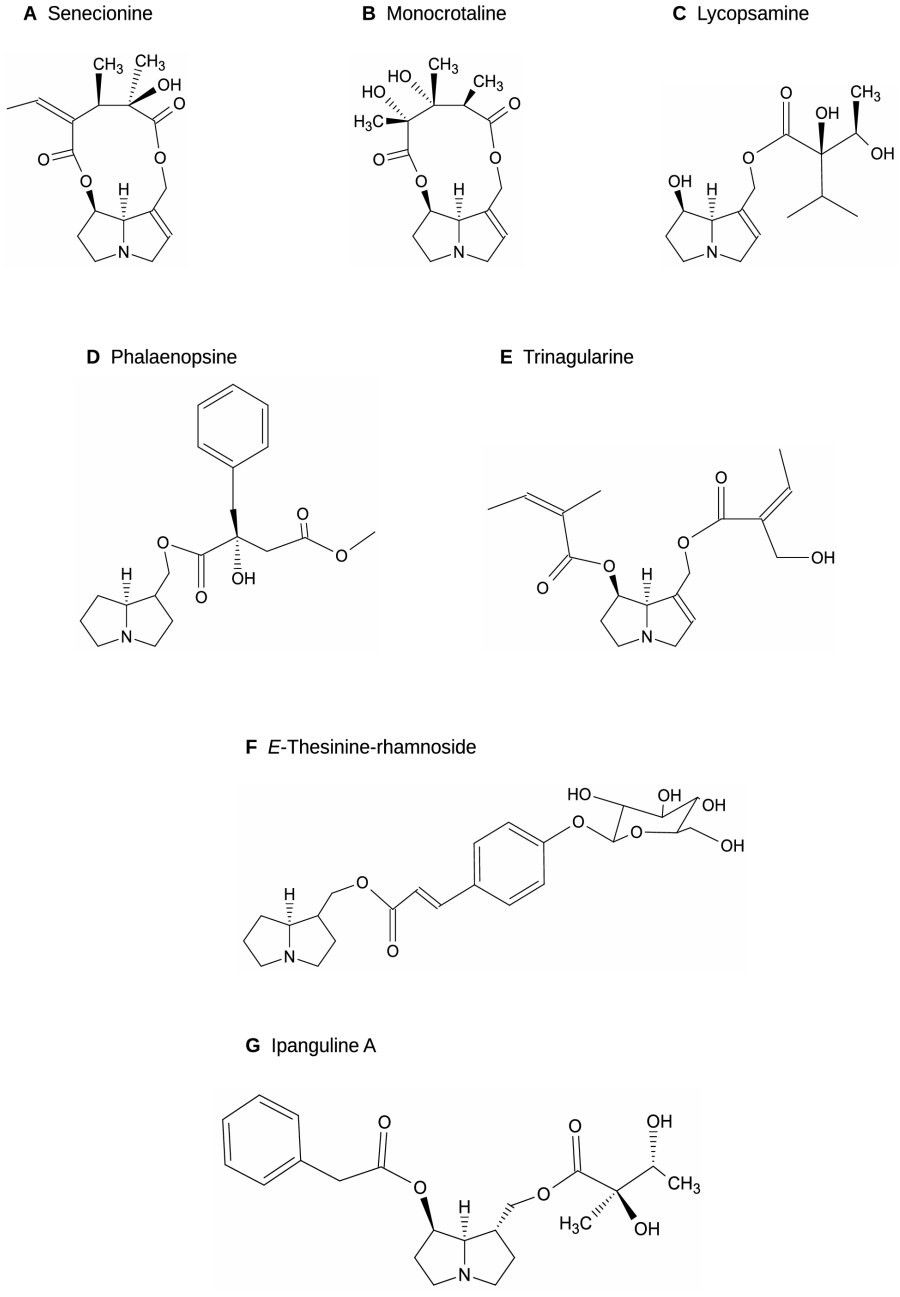


Figure 1. Pyrrolizidine alkaloids (PAs) of different types as defined by Hartmann and Witte (1995). Senecionine type (A), monocrotaline type (B), lycopsamine type (C), phalaenopsine type (D), triangularine type (E), miscellaneous PAs, (F,G).

1.3. The Biosynthesis of Pyrrolizidine Alkaloids

Although the scientific community is aware of PAs since the 1950s, surprisingly little is known about the enzymes involved in biosynthesis (Hartmann and Witte, 1995; Kaltenegger and Ober, 2015). During the early stages of "PA research" indications for general steps of the biosynthesis came mainly from radioactive tracer feeding experiments and revealed that the necine base is synthesized from putrescine generated from the aminoacids L-ornithine and L-arginine (Nowacki and Byerrum, 1962; Bottomley and Gheissman, 1964; Hughes et al., 1964; Bale and Crout, 1975). Later, tracer feeding experiments with ^{13}C, revealed that the initial substrate of the PA biosynthesis is homospermidine that is exclusively incorporated into the necine base of PAs (Khan and Robins, 1981; Grue-Sørensen and Spenser, 1982; Khan and Robins, 1985; Robins, 1989). Robins (1982) was able to synthesize trachelanthamidine, a precursor of the 1-hydroxy-methylpyrrolizidine and the base of many plant derived PAs, in a bio-mimetic experiment with pea seedling DAO. In this experiments, pea seedling diamine oxidase was incubated under physiological conditions for several days with homospermidine until an unknown intermediate (probably an imminium ion, "Metabolite X" in Frölich et al., 2007) was formed. The subsequent addition of an unspecific alcohol dehydrogenase (ADH) led to the formation of trachelanthamidine. The synthesis of the necine acids remains even more unclear. It is hypothesized that the typical necic acid of lycopsamine type PAs, the trachelantic acid, originates from the branched-chain amino acids (Weber et al., 1999). This synthesis step is presumably carried out by an acetohydroxacid synthase (AHAS). When this reaction takes place and where other necic acids are derived from is not clear yet. An overview of postulated steps of PA biosynthesis is shown in Chapter 2, Fig. 1.

The only enzyme of the PA biosynthesis that is identified until today is the homospermidine synthase (HSS) (Böttcher et al., 1993). The HSS catalyzes the transfer of a 4-aminobutyl moiety from spermidine to putrescine while forming homospermidine (Böttcher et al., 1994). The HSS originated by gene duplication from the primary metabolism enyzme deoxyhypusine synthase (DHS) and was independently recruited in different plant lineages (Ober and Hartmann, 1999a; Anke et al., 2004; Reimann et al., 2004; Nurhayati and Ober, 2005; Kaltenegger et al., 2013; Irmer et al., 2015). The DHS is an essential enzyme in all eukaryotes and

catalyzes the activation of the eukaryotic initiation factor 5A (eIF5A) which is essential for cell division and growth (Park et al., 2010). The enzymes following homospermidine are unknown but it is hypothesized that a diamine oxidase catalyzes the oxidation of homospermidine allowing the formation of the necine base (Robins, 1982; Kelly and Robins, 1988; Robins, 1989).

1.4. PAs in the Boraginaceae and Comfrey

The family of the Boragninaceae contains the largest number of PA producing plants and PAs are commonly found in almost every related species (Hartmann and Witte, 1995; El-Shazly and Wink, 2014). Typical PAs are monoesters, open chain diesters, and less often macrocyclic diesters like lycopsamine, myoscorpine, and senkirkine, respectively (Hartmann and Witte, 1995). They are stored as *N*-oxides or mixtures with their respective free base (Hartmann and Witte, 1995; Hartmann and Ober, 2000; El-Shazly and Wink, 2014). In hairy root cultures of *Cynoglossum officinale* and *Symphytum officinale* PAs were mostly found as *N*-oxides (Frölich et al., 2007). It was also suggested that PAs might be induceable by the elicitors methyl jasmonate, quercetine, and salicylic acid in hairy roots cultures of *Echium rauwolfii* (Abd El-Mawla, 2010). Sievert et al. (2015) found no evidence for induction of PA biosynthesis in *Heliotropium indicum* (Heliotropiaceae, a closely related family of Boraginaceae), *Cynoglossum officinale*, and *Symphytum officinale* by treatment with methyl jasmonate. In general, PAs are regarded as non induceable and are constitutively produced while the plant grows and develops inflorescences (Hartmann and Witte, 1995; Wittstock and Gershenzon, 2002; Hartmann and Ober, 2008).

In Boraginaceae occurs a huge variation of PA concentrations between different tissues (El-Shazly and Wink, 2014). Most of the time, developing young plant organs, like young leaves or inflorescences, as well as roots, the predominant site of PA biosynthesis, show the highest concentrations of PAs in many species (Catalfamo et al., 1982; van Dam et al., 1994; van Dam et al., 1995b; Frölich et al., 2007; Kruse et al., 2017). A good example for a plant with such a pattern is comfrey (Sympyhtum officinale). *Symphytum officinale* is a perennial, herbal plant and is with coarse, hairy leaves and stem a typical representative of the Boraginaceae. It was shown that *Symphytum* produces PAs in the root by immunolocalisation of HSS and radioactive tracer-feeding experiments (Frölich et al., 2007; Niemüller et al., 2012). The highest concentrations of PAs in *S. officinale* are typically found in the rhizome and roots

(Chizzola, 1993; Couet et al., 1996). Regarding the different PA structures that are found in the plant, *Symphytum* shows a relative complex mixture of retronecine O^9-monoesters and O^7,O^9-diesters compared to other species of the Boraginaceae (Frölich et al., 2007). The toxic PAs could bear a serious health risk for users as comfrey is widely used as a medicinal plant for the treatment of various issues (Staiger, 2012).

2. OBJECTIVES AND OUTLINE OF THIS WORK

Little is known about the enzymes involved in the PA biosynthetic pathway compared to the high number of known PA structures, identified intermediates and the deep knowledge about the first pathway specific enzyme HSS. Therefore, the main question of this thesis is: which further enzymes are involved in the PA biosynthesis? Three consecutive approaches were conducted to answer this question, and are described in the three chapters of this thesis.

The general hypothesis behind these approaches was that genes that are involved in steps of the PA biosynthesis are exclusively expressed in tissues that produce PAs. Therefore, the first aim was the localization of PA biosynthesis in the PA producing plant *Symphytum officinale* (comfrey). *Symphytum* has been chosen as PA-containing model plant because of knowledge about occurring PAs, one identified site of PA biosynthesis (root), and the number of identified transcripts. The identification and characterization of a further site of PA biosynthesis in *Symphytum* is described in **Chapter 1** and was recently published in *Plant Physiology* (Kruse et al., 2017). **Chapter 2** outlines how the localization of PA biosynthesis in specific tissues of *Symphytum* was used for an *in silico* approach to identify candidate genes coding for enzymes involved in different steps of the PA biosynthesis pathway. In this chapter, a particular focus was the assembly of a *de novo* transcriptome, the identification of candidate genes by differential expression analysis on transcript level, and the establishment of a heterologous expression system and *in vitro* enzyme assay for

copper-containing diamine oxidases (DAO). Subsequently, a method for RNAi mediated gene silencing in hairy roots of *S. officinale* was developed to allow direct testing of candidate genes by gene knockdown and is described in detail in **Chapter 3**. Here, the special interest was the implementation of a proof-of-concept experiment to test the capabilities and limitations of the method. This was addressed by generating HSS knockdown hairy root mutants, following quantification of transcripts, and identification and quantification of PA structures.

In the last part of this thesis, an overall conclusion based on the three complementary approaches draws and suggests follow-up experiments that could answer remaining open questions. All chapters contain co-authored material and the according author contributions are listed at the end of the thesis.

3. CHAPTER 1

Second site of Pyrrolizidine Alkaloid biosynthesis in Comfrey boosts Plant Defense in Floral Stage

Lars H. Kruse, Thomas Stegemann, Christian Sievert, and Dietrich Ober

Botanisches Institut, CAU Kiel, Am Botanischen Garten 1-9, 24118 Kiel

Detailed author contributions are listed at the end of the thesis.

3.1. ABSTRACT

Pyrrolizidine alkaloids (PAs) are toxic secondary metabolites that are found in several, distantly related families of the angiosperms. The first specific step in PA biosynthesis is catalyzed by homospermidine synthase (HSS), which has been recruited several times independently by duplication of the gene encoding deoxyhypusine synthase (DHS), an enzyme involved in the post-translational activation of the eukaryotic initiation factor 5A. HSS shows highly diverse spatiotemporal gene expression in various PA-producing species. In *Symphytum officinale* (Boraginaceae), PAs are reported to be synthesized in the roots, with HSS being localized in cells of the root endodermis. Here, we show that *S. officinale* plants activate a second site of HSS expression when inflorescences start to develop. HSS has been localized in the bundle sheath cells of specific leaves. Tracer feeding experiments have confirmed that these young leaves not only express HSS, but the whole PA biosynthetic route. This second site of PA biosynthesis results in drastically increased PA levels within the inflorescences. The boost of PA biosynthesis is proposed to guarantee optimal protection especially of the reproductive structures.

3.2. INTRODUCTION

Plant secondary metabolism is highly diverse on a chemical and regulatory level (Grotewold, 2005; Pichersky et al., 2006; Weng et al., 2012; Cordell, 2013). It is often characterized by a complex interplay of various cell types and enzymes (Facchini and St-Pierre, 2005; Ziegler and Facchini, 2008). Well-characterized examples of the involvement of various cell types in the biosynthesis of alkaloids can be found in the benzylisoquinoline alkaloid biosynthesis of opium poppy (*Papaver somniferum*; Bird et al., 2003), in the tropane alkaloid biosynthesis of *Hyoscyamus niger* and *Atropa belladonna* (Nakajima et al., 1999; Suzuki et al., 1999a; Suzuki et al., 1999b), and in the monoterpenoid indole alkaloid biosynthesis of *Catharanthus roseus* (St-Pierre et al., 1999; Burlat et al., 2004). Moreover, the transport from the site of synthesis to the site of accumulation requires various cell types as described for nicotine in tobacco (Neumann, 1985; Morita et al., 2009) or pyrrolizidine alkaloids (PAs) in *Senecio* (Sander and Hartmann, 1989).

PAs are toxic compounds for the chemical defense of plants and show a scattered occurrence in many distantly related angiosperms but have in common homospermidine synthase (HSS), which is the first pathway-specific enzyme of PA biosynthesis in all these lineages. HSS uses spermidine and putrescine as substrates to catalyze the formation of homospermidine, an intermediate that has been shown to be incorporated exclusively into the necine base moiety, the characteristic bicyclic ring system of PAs (Böttcher et al., 1993; Böttcher et al., 1994). HSS has evolved by gene duplication of deoxyhypusine synthase (DHS), an enzyme involved in the post-translational activation of the eukaryotic initiation factor 5A (Ober and Hartmann, 1999a; 1999b). Recent studies have shown that this duplication has occurred several times independently during angiosperm evolution in lineages that are able to produce PAs (Anke et al., 2004; Reimann et al., 2004; Kaltenegger et al., 2013; Irmer et al., 2015). This observation suggests that the whole pathway of PA biosynthesis also evolved several times requiring repeated integration into the metabolism of the plant. Expression analyses of HSS in PA-producing plants of various lineages show that this integration resulted in different cellular organizations of PA biosynthesis in different species. In the Asteraceae family, for example, in which independent duplications have been described in two lineages, HSS expression has been detected in *Senecio* (ragwort, tribe Senecioneae) only in cells of the endodermis and

adjacent root parenchyma, cells that lie opposite to the phloem, whereas, in *Eupatorium cannabinum* (hemp-agrimony, tribe Eupatorieae), HSS expression has been found in all cells of the root cortex, but not in the endodermis (Moll et al., 2002; Anke et al., 2004). In orchids (*Phalaenopsis* species), HSS has been localized in meristematic cells of the root apical meristem and in young flower buds (Anke et al., 2008). Within Boraginales, a single duplication event resulted in the evolution of HSS. However, in this lineage, the expression pattern of HSS has been shown to be multifaceted, suggesting independent scenarios of optimization and integration of the gene duplicate into the metabolism of the plant. Within the Heliotropiaceae (*Heliotropium indicum*, Indian heliotrope), HSS expression is found in the lower leaf epidermis, whereas in *Cynoglossum officinale* (houndstongue, Boraginaceae), the endodermis and pericycle of the root express HSS (Niemüller et al., 2012). In *Symphytum officinale* (Comfrey, Boraginaceae), HSS has been localized exclusively in cells of the endodermis (Niemüller et al., 2012) supporting previous studies by Frölich et al. (2007) who have identified the roots of S. *officinale* as the site of PA biosynthesis by tracer experiments with radiolabeled putrescine. Studies by Niemüller et al. (2012) using protein blot analyses suggested that HSS expression occurs during a restricted window of leaf development in *S. officinale*, an observation that warrants a closer look at the PA biosynthetic capacity of leaves of this plant species.

In this study, we show that *S. officinale* synthesizes PAs not only in the roots, but also in young leaves subtending an inflorescence with unopened flower buds. We further show that HSS is expressed in specific cells of the leaf during certain stages of inflorescence development coinciding with PA production in the leaf. This second auxiliary site of PA biosynthesis during flower development is discussed in comparison with a similar mechanism previously observed in the orchid *Phalaenopis amabilis* in which young flower buds produce PAs to boosts PA levels in the mature flower in order to guarantee an efficient supply of deterrent PAs for the developing reproductive structures (Anke et al., 2008).

3.3. RESULTS

3.3.1 Levels of *hss* Transcript and Protein Expression are Dependent on Leaf Position in Relation to the Inflorescence

Analyzing three species of Boraginales, Niemüller et al. (2012) has shown that HSS expression is species-specific. In *S. officinale*, in which the whole PA biosynthesis is localized in the roots, their study has indicated that HSS is also expressed in leaves of a specific developmental stage, *viz*., leaves that are directly beyond a terminal inflorescence with closed flower buds.

In order to correlate leaf position with *hss* expression in *S. officinale*, we harvested leaves from various positions in the shoot system of a plant just opening its first flower buds (Fig. 1A). Analyses with reverse transcription (RT)-PCR and quantitative reverse transcription (qRT)-PCR showed that the highest levels of *hss* transcript were detectable in young leaves directly next to inflorescences with buds and flowers that had just opened (Fig. 1B and Suppl. Fig. 2). The *hss* transcript level differed between the youngest and the oldest leaf by almost a factor of 30 (Suppl. Fig. 2). To test whether *hss* transcript levels were correlated with the expression of HSS protein, protein extracts of leaves from various positions of the shoot system from another plant individual were analyzed by a protein gel blot with an antibody specific for HSS of *S. officinale* (Niemüller et al., 2012). Like the *hss* transcript level, *hss* protein expression is strongest in young leaves directly subtending an inflorescence (leaf 1 and 2) and fades away with increasing distance of the leaves to the flower (Fig. 1C). Of note, in leaves that have no axillary branch developing in their axils, HSS protein is not detectable (leaves 9, and 10, blue in Fig. 1A), whereas leaves that possess axillary branches show distinct HSS expression (leaves 3, 4, 6, 8, red in Fig. 1A). Congenital fusions of an axillary branch with the main axis are common for *S. officinalis* (Kotelnikova et al., 2011) and result in a lateral bud on the axis distant from its subtending leaf (e.g. leaves 5 and 7, purple in Fig. 1A).

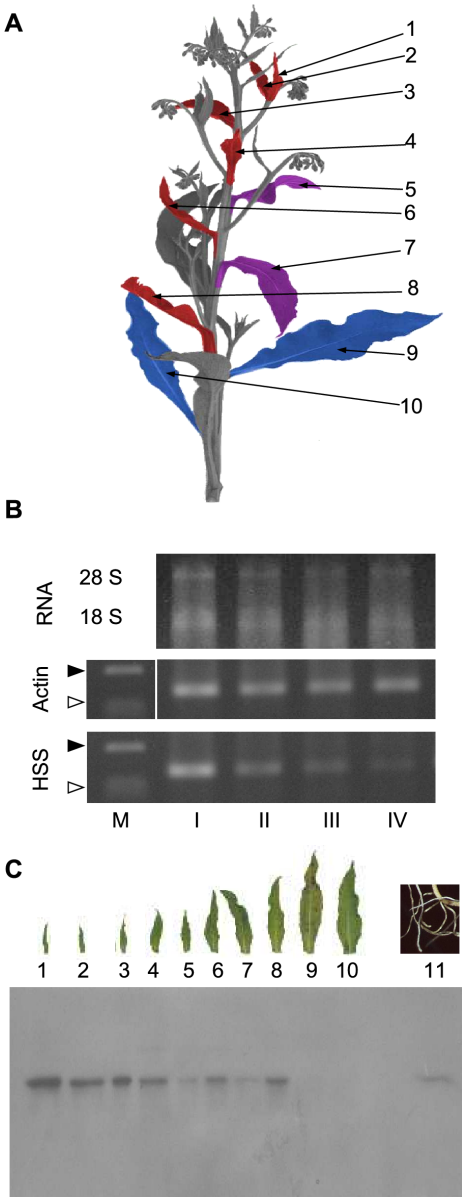


Figure 1. Transcript and protein blot analyses of leaves at various developmental stages. A, Habitus of a *S. officinale* shoot. Numbers indicate leaves that were analyzed by protein blot analysis. The color code indicates the level of HSS expression according to the protein blot analysis given in C (red, high expression; purple, medium expression; blue, no expression). B, Semi quantitative RT-PCR shows the highest expression of HSS transcripts in leaf type I. Leaf types I to IV comply with leaves 1, 4, 8, and 10 in Fig. 1A, respectively, but were harvested from another plant individual as independent proof. RNA integrity was confirmed by detection of two distinct bands of 28 S and 18 S rRNA. Actin served as the reference gene. Closed arrow heads show marker bands (M) at 150 bp, open arrows indicate marker bands (M) at 100 bp. C, Protein blot analysis of protein extracts from leaves numbered according to A. Crude protein extracts (10 µg protein per lane) were separated via SDS-PAGE and blotted onto a polyvinylidene difluoride (PDVF) membrane. Affinity-purified HSS-specific antibody was used for detection in combination with a secondary goat-anti-rabbit antibody conjugated with horseradish peroxidase. A protein extract of a root of *S. officinale* served as a positive control (sample 11).

3.3.2 Immunolocalization of HSS in Leaves of *S. officinale*

Leaves that tested positive for HSS expression were used to identify, by immunolocalization, the cells expressing HSS. HSS expression was found to be restricted to bundle sheath cells that formed a layer of compactly arranged parenchyma around the vascular bundle (Fig. 2C and Fig. 2D). The labeled layer had a thickness of one to two cells and was separated from the epidermis by at least a single row of unlabeled mesophyll cells. Specificity of the label was tested by pre-incubation of the antibody with soluble, heterologously expressed HSS, DHS, or BSA before labeling. Only pre-incubation with HSS resulted in a decrease of signal, whereas pre-incubation with DHS or BSA had no effect (Suppl. Fig. 3).

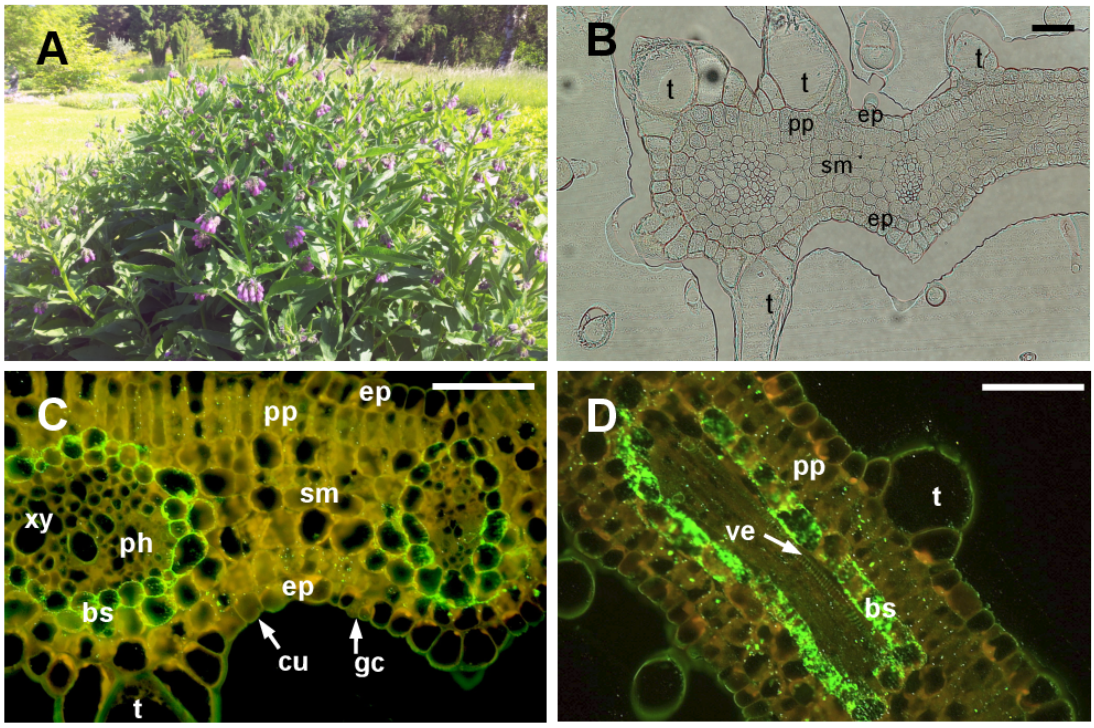


Figure 2. Immunolabeling of HSS in cross-sections of a young leaf subtending an inflorescence with unopened flower buds of *S. officinale*. A, Plant habitus. B, Brightfield microscopy of a leaf cross-section. C, Detail of B. D, Leaf cross-section with a vascular bundle cut longitudinally. In C and D, HSS was immunolabeled by incubation with an affinity-purified HSS-specific antibody and visualized with a secondary antibody conjugated with AlexaFluor488 under UV-light. Size bars refer to 10 µm. ep, epidermis; bs, bundle sheath cells; cu, cuticle; pp, palisade parenchyma; sm, spongy mesophyll; ph, phloem; t, trichome; ve, vessel element; xy, xylem; gc, guard cell.

The close vicinity of HSS expression and the vascular bundle is also observed in the roots of *S. officinale*, in which HSS is only detectable in the endodermis that encloses the root central cylinder with the vascular tissue (Niemüller et al., 2012). Both of these tissues, namely the endodermis and the bundle sheath cells, play a key role in controlling fluxes to and from the vascular bundle and are important for maintaining the transport of nutrients and water in the plant (Steudle and Peterson, 1998; Leegood, 2008; Geldner, 2013).
Of note, nucleotide sequences encoding HSS amplified from the leaves proved to be identical to those amplified from the roots, suggesting that a single *hss* gene is responsible for the expression of HSS in the endodermis of the root and the bundle sheath of specific leaves.

3.3.3 Capacity of *S. officinale* Leaves to Synthesize PAs

By applying ^{14}C-labeled putrescine to young detached leaves expressing HSS, we tested whether leaves expressed only HSS or the complete biosynthetic machinery to produce PAs. After incubation for four days, the tracer was completely taken up by the leaves, of which 16% of the applied radioactivity was detectable in the alkaloid-containing sulfuric acid crude extract. After reduction to convert all PA *N*-oxides to their tertiary form, PAs were enriched by solid phase extraction. This alkaloid-enriched fraction containing 8% of the applied radioactivity was split into two fractions that were treated differentially to confirm the identity of the labeled compounds as PAs (Tab. 1). One fraction was analyzed directly by HPLC to detect radioactively labeled PAs; several peaks of putative PAs were found (Fig. 3A). To establish that the detected radioactive peaks were PAs that resulted from the incorporation of the tracer putrescine, the second fraction was hydrolyzed to reduce all potential PAs to their backbone, the necine base. *S. officinale* produces a PA bouquet of intermedine derivatives including O^9-monoesters, O^7,O^9-diesters, and 3'-acetyl derivatives such as 7-acetylintermedine, 7-acetyllycopsamine, 7-senecioylintermedine (echiupinine), or 7-tigloylintermedine (myoscorpine) (Hartmann and Witte, 1995; Frölich et al., 2007). The necine base of these PAs, all belonging to the so-called lycopsamine type, is retronecine (Hartmann and Witte, 1995). If the radioactive peaks represent PAs, hydrolyzation should result in a single peak that co-migrates with a retronecine standard. In contrast to the several peaks detectable in the non-hydrolyzed extracts (Fig. 3A), the hydrolyzed extract (Fig. 3B) showed only one prominent peak that had the same retention time as retronecine (Fig. 3C). Spiking the hydrolyzed sample with

retronecine resulted only in an intensification of the signal indicating that the same substance was being analyzed (Fig. 3D). As an additional proof of identity, the feeding experiment was repeated as described, but with unlabeled putrescine. The alkaloid-containing fraction was hydrolyzed and separated by thin layer chromatography (TLC) parallel to the hydrolyzed sample resulting from the tracer feeding experiment. After the localization of radioactivity, the corresponding area in the lane of the non-radioactive samples was scraped off the plate, and the compounds were eluted and analyzed by gas chromatography-mass spectrometry (GC-MS). In this spot, retronecine was detected as the major compound, giving support to the interpretation that retronecine was the radioactive compound detected after hydrolysis of the extract after tracer feeding (Suppl. Fig. 4).

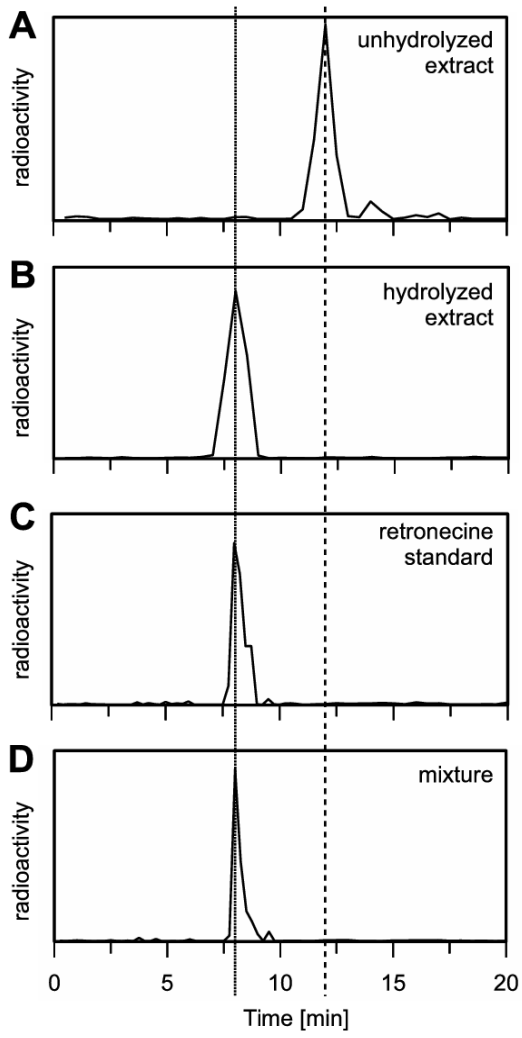


Figure 3. HPLC elution profiles of alkaloid extracts of leaves after tracer feeding. Young leaves subtending an inflorescence with unopened flower buds were fed with [14C]putrescine as tracer. Extracts were enriched for PAs by solid phase extraction (SPE) purification before separation. A, Non-hydrolyzed leaf extract. B, Hydrolyzed leaf extract. C, Radiolabeled retronecine standard. D, Mixture of hydrolyzed extract and retronecine standard.

Table 1. Tracer-feeding experiment. Recovery of applied radioactivity.

	radioactivity kBq	**% of recovered radioactivity**
Total radioactivity applied	169	100
Crude extract (sulfuric acid)	25	16
Crude extract after reduction	27	16
Extract purified via Strong Cation eXchange-Solid Phase Extraction (SCX-SPE, PA enrichment)	15	8

3.3.4 Leaves Boost PA Levels in Inflorescences

To test the consequences of the capacity of leaves to produce PAs, we analyzed the transcript and protein levels of HSS and the total PA content in leaves subtending an inflorescence and in their associated inflorescences during development. Seven different inflorescence developmental stages were defined with respect to size and flowering state (Fig. 4A). Stage I was characterized by its small size (approx. 1 cm in diameter), whereas stage II was much larger but had flower buds that were still closed. Size increased until stage V. At stage IV, the flower bud tips became purple. In stage V, the first buds started flowering, but most of the buds maintain closed, whereas by stage VI, all the flowers had open, with those that had opened first were beginning to darken. At stage VII, the flowers withered, and fruits began to develop. We observed that HSS transcript and protein in leaves were detectable from stage I onward, and that HSS expression levels reached their maximum at stage II and III, after which they decreased to a low level and remained constant to the last stage (Fig. 4B and Suppl. Fig. 6). PAs were not detectable in leaves of stage I. PA levels increased from stage II to IV to a concentration of about 20 µg/g dry weight, and reached at stage VII concentrations of about 180 µg/g dry weight (Fig. 4C), similar to that described for leaves of *S. officinale* (Mattocks, 1968; Couet et al., 1996). As in the leaves, no PAs were detectable in inflorescences at stage I. Inflorescences at stages II-V showed PA concentrations of approx. 300-400 µg/g dry weight, increasing at stage VI to a concentration of 3344 µg/g dry weight (Fig. 4C) with a total PA content of 700 µg per inflorescence (Fig. 4D). These high concentrations correspond to PA levels described for flowers of other PA-producing plant species (Muetterlein and Arnold, 1993; Couet et al., 1996; Frölich et al., 2006; Frölich et al., 2007). In the

last analyzed stage, namely stage VII, the concentration dropped to approx. 500 µg/g dry weight. The comparatively low PA concentrations in leaves compared with inflorescences suggest that PAs are efficiently translocated from the leaves to the inflorescences. Thus, PA biosynthesis in the young leaf subtending the inflorescence ensures a sufficient supply of PAs (i) to keep alkaloid levels constant as long as the inflorescences enlarge and (ii) to enable high PA levels to accumulate when the flowers open.

To test whether the inflorescence itself contributes to the high PA concentrations, we analyzed HSS expression in inflorescences of various developmental stages (Suppl. Fig. 7). Only very low levels of HSS protein were detectable at any stage of inflorescence development. This observation is in accordance with Niemüller et al. (2012) who have found low levels of HSS transcript but no protein in inflorescences. In contrast to the inflorescences, roots showed a high HSS level of approx. 40-50 ng per 20 µg of total protein, a value which is comparable with the highest HSS levels detected in the leaf at stage II with a maximum value of 39 ng. The delay between the highest levels of HSS (stage II) and the peak of PA accumulation (stage VI) results most likely from two factors: (i) the need to transport PAs from their site of synthesis in the young leaf into the inflorescence, and (ii) the finding that PA biosynthesis in species of the Boraginales seems to be less channeled than in other PA-producing species as shown by Frölich et al. (2007) using radioactive labeled putrescine as a tracer. The maximal incorporation of tracer into PAs took more than two weeks with a temporary accumulation of pathway intermediates such as homospermidine or a polar uncharacterized compound. The drastic decrease of PA levels in stage VII might have various causes. One explanation could be the abscission of the petals, which was observed at this stage. PAs in the petals might be a strategy to fend off nectar robbers such as short-tongued bumblebees of the genus *Bombus* (Utelli and Roy, 2001; Irwin and Maloof, 2002). Nectar robbers bite a hole in the petal tube to reach nectar without actually entering and pollinating the flower (Castro et al., 2008). Another explanation for the drop of PA levels could be a degradation of PAs in these tissues or the translocation of PAs from the inflorescences to other tissues of the plant, as PA *N*-oxides are highly mobile and can easily be translocated from one tissue to another (Hartmann et al., 1989). Further studies should elucidate the dynamics of PA levels in these reproductive structures.

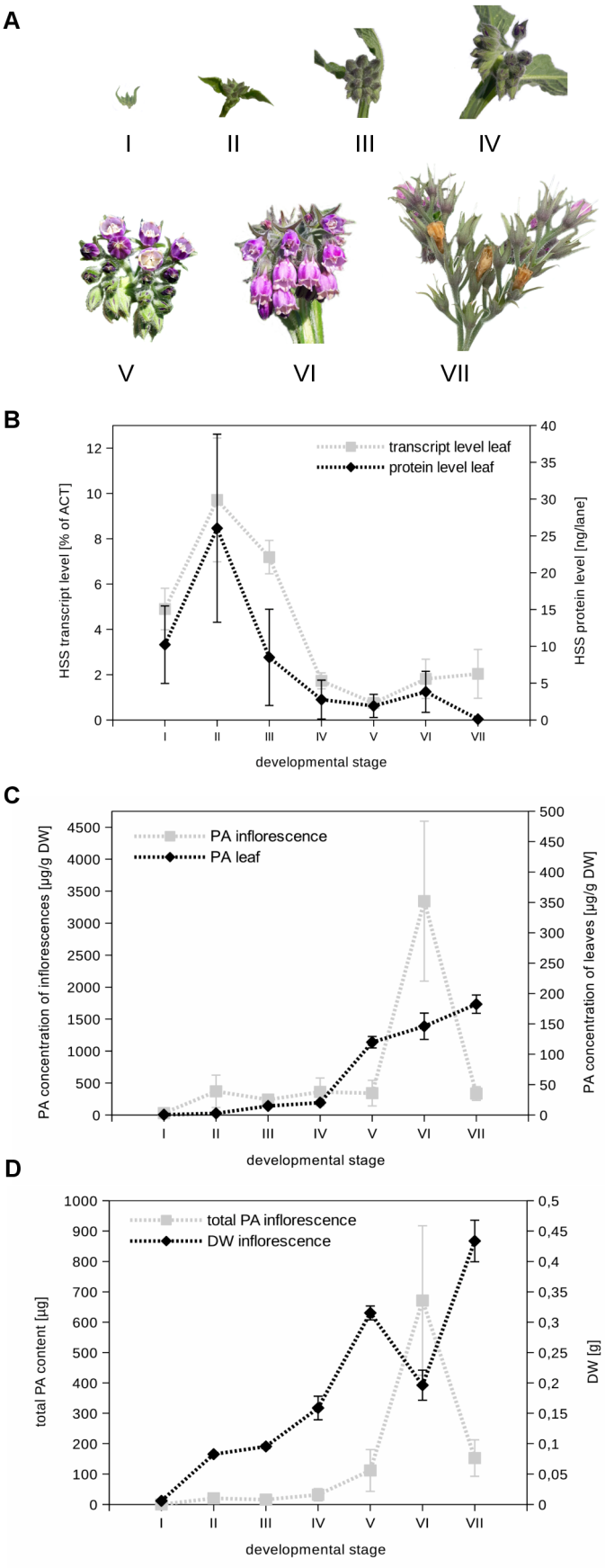


Figure 4. PA content and HSS expression in leaves and inflorescences of various developmental stages.
A, Inflorescences at various developmental stages. B, Protein and transcript levels of HSS in leaves subtending inflorescences. C, PA concentrations in the inflorescences and in the subtending leaves at the respective developmental stages. D, Total PA content and dry weight of the inflorescence. Inflorescences and corresponding subtending leaves were harvested, and inflorescences at the developmental stages I-VII were weighed. PAs from inflorescences and leaves were extracted and analyzed according to the sum parameter method (C and D). For relative transcript quantification, qRT-PCR was used with actin (ACT) as the reference gene (B). For relative protein quantification, total protein was extracted and separated via SDS-PAGE before being transferred to a PDVF membrane. Affinity-purified HSS-specific antibody was used for detection in combination with a secondary goat-anti-rabbit antibody conjugated with horseradish peroxidase. Protein level was calculated in relation to 25 ng recombinant HSS that was used as a standard (B). Error bars indicate the standard error of the mean (SEM) of three independent biological replicates for each developmental stage. Error bars of PA concentration in leaves of stage II and III are missing in C because of insufficient sample material.

3.4. DISCUSSION

In this study, we have been able to show that *S. officinale* has two PA-biosynthetic sites that differ with respect to their function, *i.e.,* the roots as a well-characterized organ of PA biosynthesis and young leaves that are in the direct vicinity of a developing inflorescence. The roots in *S. officinale* are responsible for constitutive PA production, most likely for the protection of the vegetative parts of the plant. Confirmation that the whole PA pathway is expressed in the roots has been obtained from root cultures that are able to produce PAs (Frölich et al., 2007). A link between growth and PA biosynthesis should guarantee a constant ratio between PA and biomass, a phenomenon that we have also observed for PA-producing species of other lineages (Anke et al., 2004; Anke et al., 2008; Niemüller et al., 2012). In contrast, the young leaves subtending an inflorescence with unopened flower buds characterized in this study provide an additional PA supply for the reproductive tissues within the flowers of *S. officinale*, a pattern previously observed for PA-producing *Phalaenopsis* (Orchidaceae). In *Phalaenopsis*, specific structures within small flower buds, but not of open flowers, produce PAs during flower development, boosting PA levels up to about 6 mg/g fresh weight (Anke et al., 2008). The assumption that, in *S. officinale* PA biosynthesis in the young leaf is also of functional importance is supported by the incorporation rate of labeled putrescine into PAs of almost 10% (Tab. 1). This incorporation rate is remarkable bearing in mind that putrescine is not only a precursor for PA biosynthesis, but also part of the highly dynamic polyamine pool of primary metabolism, including essential functions such as growth, development, or biotic and abiotic stress responses (Evans and Malmberg, 1989; Alcázar et al., 2010; Gill and Tuteja, 2010; Tiburcio et al., 2014). As HSS evolved independently in *Symphytum* and *Phalaenopsis* (Reimann et al., 2004; Nurhayati et al., 2009), the selection pressure that resulted from the need to protect the reproductive structures against herbivores obviously resulted in a similar regulation of PA biosynthesis in these two lineages. The protection of reproductive structures is a phenomenon that is often observed in the plant kingdom and that is in good accordance with the optimal defense theory (McKey, 1979).

The pattern of HSS expression in PA-producing angiosperms is highly variable. Nonetheless, several PA-producing plants show similarities in HSS expression. The vicinity of the vascular tissue to cells involved in PA biosynthesis is a motif not just

found in the roots and leaves of *Symphytum*, but also in other PA-producing plants such as *C. officinale* (also Boraginaceae) expressing HSS in the endodermis and the pericycle of the roots or in *Senecio* species (Asteraceae) in which HSS is expressed in cells directly next to the phloem. In *Senecio*, the phloem has been shown to be the tissue by which PAs are transported from the site of synthesis to the shoot in which they are then efficiently allocated to the inflorescences, the major site of storage (Hartmann et al., 1989; Moll et al., 2002). Moreover, a link between alkaloid biosynthesis and the vascular tissue has been described for other plant systems, e.g., for the early steps of monoterpene indole alkaloid biosynthesis (Burlat et al., 2004), for the biosynthesis of morphine in opium poppy (Bird et al., 2003; Weid et al., 2004), and for tropane alkaloid biosynthesis in solanaceous plants (Kanegae et al., 1994; Suzuki et al., 1999a).

Despite the probability that HSS, and most likely, the complete pathway of PA biosynthesis evolved several times independently in various lineages of the angiosperms (Reimann et al., 2004; Anke et al., 2008; Kaltenegger et al., 2013), the similarity between the PA structures produced is remarkable. Not only the bicyclic ring system characteristic for all PAs, namely the necine bases, are identical with respect to their structure and stereochemistry between unrelated plant lineages, but also the complete PA molecules (Hartmann and Witte, 1995). Similar selection pressures probably forced the evolution of similar, almost identical traits (Pichersky et al., 2006). In this study, we have been able to show that the strategy of using a second auxiliary site for PA biosynthesis to protect floral structures also evolved independently in two distantly related species, extending the aspect of convergent evolution from molecule structures to the regulation of the whole pathway. Further work will be needed to analyze the evolutionary mechanisms underlying the repeated recruitment, optimization, and integration of PA biosynthesis into the metabolism of the plant.

3.5. EXPERIMENTAL

3.5.1 Plant Material

Symphytum officinale was grown in pots with a mixture of TKS2 (Floragard) and lava granulate at a ratio of 3:1 in the Botanical Gardens Kiel from April to September.

3.5.2 RNA Isolation and Quantification of Transcripts

For transcript analyses of leaves in relation to their position in the shoot system, plant samples were pulverized in liquid nitrogen with mortar and pestle before total RNA was extracted with Trizol® (Invitrogen, Life Technologies) according to the manufacturer's protocol, but with two additional phase separation steps involving the addition of 200 µl chloroform to the aqueous phase followed by phase separation via centrifugation. Finally, RNA was washed twice with ice-cold ethanol (75%, v/v, in water). A subsequent lithium chloride (2 M final concentration) precipitation (16 h, 4°C) and two additional washing steps with ice-cold ethanol were then performed to remove possible inhibitors of the following reverse transcription reactions (RT). RNA was dissolved in RNAse-free water, and RNA integrity and purity were tested by agarose gel electrophoresis and by 260/280 nm and 260/230 nm ratio measurements by using a NanoDrop 1000 UV/VIS spectrometer. One microgram of total RNA was used as a template for reverse transcription with an Oligo(dT)17 primer (Suppl. Tab. 1) and RevertAid® Premium Reverse Transcriptase (ThermoScientific). For the transcript quantification of HSS in leaves subtending inflorescences at various developmental stages, a slightly modified protocol for total RNA isolation was employed. Instead of washes with ethanol, RNA was diluted 1:1 with 100 % ethanol and loaded on spin columns from the Direct-zol™ RNA MiniPrep Kit (Zymo Research). Samples were further processed as recommended by the manufacturer including an optional on-column Dnase I digestion. For transcript quantification, 1.5 µg total RNA was used for reverse transcription as previously described. To test for contaminating genomic DNA, control reactions were prepared without reverse transcriptase and used as a template in control PCRs (Suppl. Fig. 5).

For semiquantitative PCR (sqPCR) GoTaq-DNA polymerase (Promega) with deoxyribonucleotide triphosphates(dNTPs, 0.2 mM each) and primers (0.2 µM each)

was used, and aliquots were taken after 20, 25, 30, and 35 cycles to ensure sampling before PCR product formation reached saturation (Suppl. Fig. 1). Products were analyzed on a 2% (w/v) agarose gel. Quantitative real-time PCR was performed in a Rotor-Gene® Q System (Qiagen) with GoTaq® qPCR Master Mix (Promega) following the manufacturer's protocol. Melting curve analyses were undertaken to distinguish specific PCR products from primer dimers or unspecific PCR products. The calculation of the relative transcript levels of HSS was carried out by comparative CT methods (2-ΔΔCt for Suppl. Fig. 2), 2-ΔCt for Fig. 4B) (Schmittgen and Livak, 2008). Actin served as the reference gene to normalize expression levels. For both sqPCR and qRT-PCR, an annealing temperature of 60°C was used. Primer pairs were: P1 and P2 (for HSS) and P3 and P4 (for actin, Suppl. Tab. 1). The actin-specific primers resulted from an actin-encoding sequence amplified with a pair of degenerate primers (P5/P6) designed according to an alignment of the actin encoding sequences of *Arabidopsis thaliana* as previously described (Sievert et al., 2015). For the cloning of PCR products, the pGEM T-easy vector (Promega) was employed according to the manufacturer's protocol followed by transformation into chemically competent *E.coli* TOP10 cells (Invitrogen) for vector propagation. Plasmids were sequenced at MWG Eurofins to confirm the identity of the amplification products. Sequence data from this article have been deposited in the EMBL/GenBank data libraries under the following accession number: LT631489, *Symphytum officinale* partial mRNA for actin (*act* gene).

3.5.3 Protein-Blot Analysis of *S. officinale* Leaves

Samples were pulverized in liquid nitrogen with pestle and mortar. Protein was extracted with phosphate-buffered saline (PBS) supplemented with 5% (m/v) polyvinylpyrrolidone and 2.5% (w/v) sodium ascorbate to prevent protein precipitation by polyphenols. 10 to 20 µg total protein per sample were mixed with SDS loading buffer and separated by sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) followed by semidry blotting and immunodetection as described previously (Anke et al., 2008). The polyclonal antibody was affinity-purified against recombinant HSS of *S. officinale* (Niemüller et al., 2012). MultiMark Multi-Colored Standard (NOVEX), PageRuler™ Plus Prestained Protein Ladder (Thermo Scientific), and PageRuler™ Prestained Protein Ladder (Thermo Scientific) were used as protein mass standards as indicated in the figure legends. After immunodetection, PDVF membranes were stained with PageBlue™ Protein Staining Solution (Thermo Scientific) to ensure that equal protein amounts were loaded on each lane. For the relative quantification of protein levels by densitometry we used the software ImageJ (version 1.48) with 25 ng of recombinant HSS protein as the reference (Schneider et al., 2012).

3.5.4 Immunohistochemical Staining of HSS in Leaf Cross Sections

For the immunohistochemical localization of HSS in leaf cross-sections, young leaves were harvested; and the tip and the base were was cut off to be used as “reference tissue”, whereas the central part of the leaf was cut into pieces of approx. 1 cm side length before being immersed in ice-cold fixation buffer according to Anke et al. (2008). The “reference tissue” of each leaf was pulverized in liquid nitrogen, and total RNA was extracted and transcribed to cDNA as described above. Each “reference tissue” was tested by sqPCR for the presence of HSS transcripts. The fixated central regions of positively tested leaves were dehydrated in an ethanol series and embedded in Technovit 7100 (Heraeus-Kulzer). Sections of 3-4 µm thickness were cut on a microtome (HM3555S, Microm) and mounted on adhesive microscope slides (SuperFrost, Thermo Fisher Scientific). Immunodetection with HSS-specific antibodies and specificity tests with recombinant HSS and DHS of *S. officinale* were carried out as described previously (Niemüller et al., 2012).

3.5.5 Radioactive Tracer-Feeding Experiments and Product Analysis

To test the capacity of detached leaves of *S. officinale* to produce PAs, [14C]putrescine (3.95 GBq/mmol, GE Healthcare) was used as tracer. Two young leaves (approx. 3.5 cm in length) next to an inflorescence with developing flower buds were cut from the plant and transferred into a 1.5 ml reaction tube containing the tracer dissolved in 1 ml tap water (169 kBq). The leaves were incubated in a light/dark regime of 12 h/12 h (light of approx. 1,000 lx) at room temperature. Once the liquid was almost completely taken up, the same volume of tap water without tracer was added. After 4 days, the leaves were frozen in liquid nitrogen and stored at -50°C. Both leaves were pulverized in liquid nitrogen with mortar and pestle and each extracted in 1.5 ml 0.05 M H_2SO_4 by being vortexed for 3 min at room temperature. After centrifugation (10 min, 5000 x g) zinc dust was added in excess to the supernatants, which were then stirred for 3 hours for the complete reduction of alkaloid *N*-oxides. As SCX-SPE purification has been shown to be suitable for the enrichment of PAs (Colegate et al., 2005), the samples were further centrifuged (10 min, 5000 x g) and then applied to SCX-SPE cartridges (Phenomenex, Aschaffenburg, Germany) equilibrated with 6 ml methanol and 6 ml of 0.05 M H_2SO_4. After each column had been washed with 12 ml H_2O and 12 ml methanol, the alkaloids were eluted by applying 3 x 6 ml of methanol containing 5% (v/v) NH_4OH. All elution fractions were combined, the solvent vaporized, and the residue dissolved in 2 ml methanol. The supernatant was transferred in equal parts into two vials, dried, and stored at -20°C. For hydrolysis, one of the samples was incubated in 2 N sodium hydroxide for 3 h at 60°C, and the solvent was vaporized. For HPLC analyses, the hydrolyzed and the non-hydrolyzed samples were dissolved in 50 µl methanol. Each step of the extraction process was monitored via scintillation counting (Tri-Carb LSC, Perkin Elmer) to calculate the ratio of incorporated [14C]putrescine.

Radio-HPLC measurements were performed on a Merck-Hitachi L-6200 instrument with solvent A (100 mM phosphate buffer pH 7.5) and solvent B (acetonitrile) at a ratio of 85:15 with a flow rate of 1 ml per minute. Aliquots of 20 µl of sample were injected, and fractions were collected every 30 s for comparison of the hydrolyzed and the non-hydrolyzed extract and every 15 s for the comparison of the hydrolyzed extract with retronecine and with hydrolyzed and non-hydrolyzed monocrotaline standards (Roth, Karlsruhe, Germany). The radioactivity of the fractions was quantified by scintillation counting. Radiolabeled retronecine was obtained by the hydrolysis of [14C]senecionine (Lindigkeit et al., 1997).

For thin-layer chromatography, Silicagel G-25 TLC plates (Merck, Darmstadt) were developed in a mobile phase containing ethyl acetate, isopropyl alcohol, and ammonium hydroxide (25%, v/v) (45:35:20). Radioactivity was detected by using a radioactivity thin-layer-chromatography detector (RITA, Raytest, Straubenhardt. Germany). Compounds scraped from the TLC plate were eluted with the mobile phase, dried, and dissolved in 50 µl methanol for GC-MS analysis. GC-MS data were obtained with a Shimadzu GC-2010 chromatograph equipped with a 15 m Optima-1 MS capillary column coupled to a quadrupole mass spectrometer (Fisons MD800). Electron-impact mass spectra were recorded at 70 eV. GC conditions were: injector 250°C, temperature program 60°C for 3 min, 60 to 300°C at 6°C/min, carrier gas helium 1 ml/min. GC-MS data for retronecine were: Retention index (R_i, Kovats) 1478 (on Optima-1 MS), mass-to-charge ratio (*m/z)* 155 [M^+]; For the mass spectrometry spectrum, mass-to-charge ratio and (relative intensity): 80(100), 111(74), 155(45), 94(37), 68(34), 93(33), 82(31), 112(30), 67(29), 106(25).

3.5.6 Quantification of total PAs in inflorescences

Total PAs were extracted from lyophilized inflorescences, hydrolyzed, and reduced to obtain the necine base retronecine as described for the tracer feeding experiment. For quantification of the necine base, the sum parameter method involving silylation of the necine base followed by GC-MS was applied according to (Kempf et al., 2008). Retronecine from hydrolyzed monocrotaline served as the external standard.

3.5.7 Accession Numbers

Sequence data from this article have been deposited in the GenBank/EMBL data libraries under accession number LT631489, comfrey partial mRNA for actin (act gene).

3.6. Acknowledgments

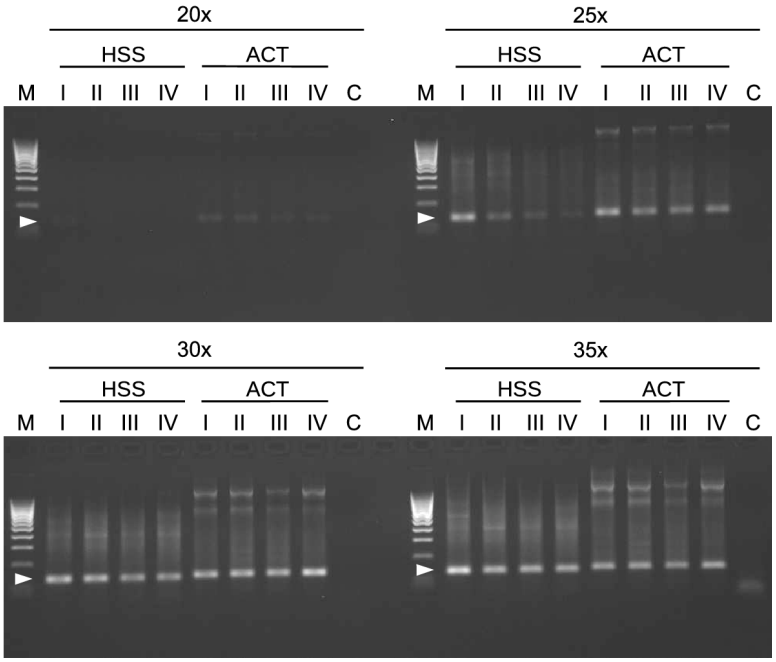
We thank Dr. Dorothee Langel for help in designing the radioactive tracer feeding experiments and Jan Baur, Margret Doose, Brigitte Schemmerling, and Karina Thöle for support in the laboratory. We are grateful to Dr. Elisabeth Kaltenegger, Dr. Jessica Garzke, and Annika Jonathas for helpful comments on the manuscript.

3.7. Supplementary material

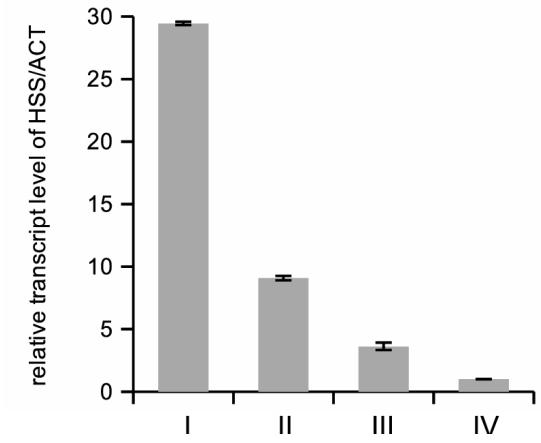
Supplementary Table 1. Primer sequences.

primer	sequence
Oligo(dT) 17	5'-GTCGACTCGAGAATTCTTTTTTTTTTTTTTTT-3'
P1	5'-AGTGCTATGGACAATGAATCAGTGA-3'
P2	5'-AGCAAAATCAGCGCCTCCA-3'
P3	5'-CAAGGCTAACAGGGAGAAAATGAC-3'
P4	5'-ATCACCAGAATCCAGCACAATACC-3'
P5	5'-WSNAAYTGGGAYGAYATGGA-3'
P6	5'-TCRBHYTTNGTDATCCACA-3'

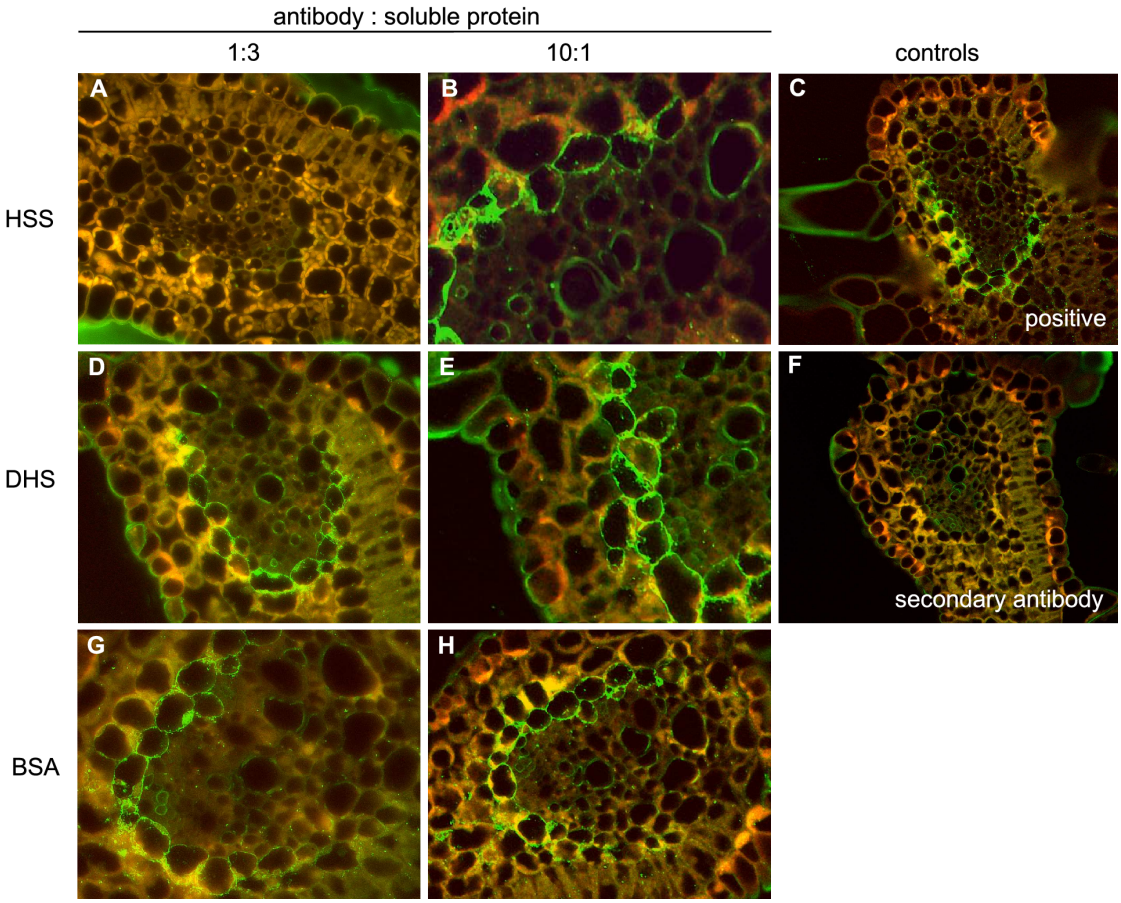
For degenerate primers P5 and P6, the following code is used: B = C + G + T, D = A + G + T, H = A + T + C, N = A + T + C + G, R = A + G, S = C + G, W = A + T, and Y = T + C.



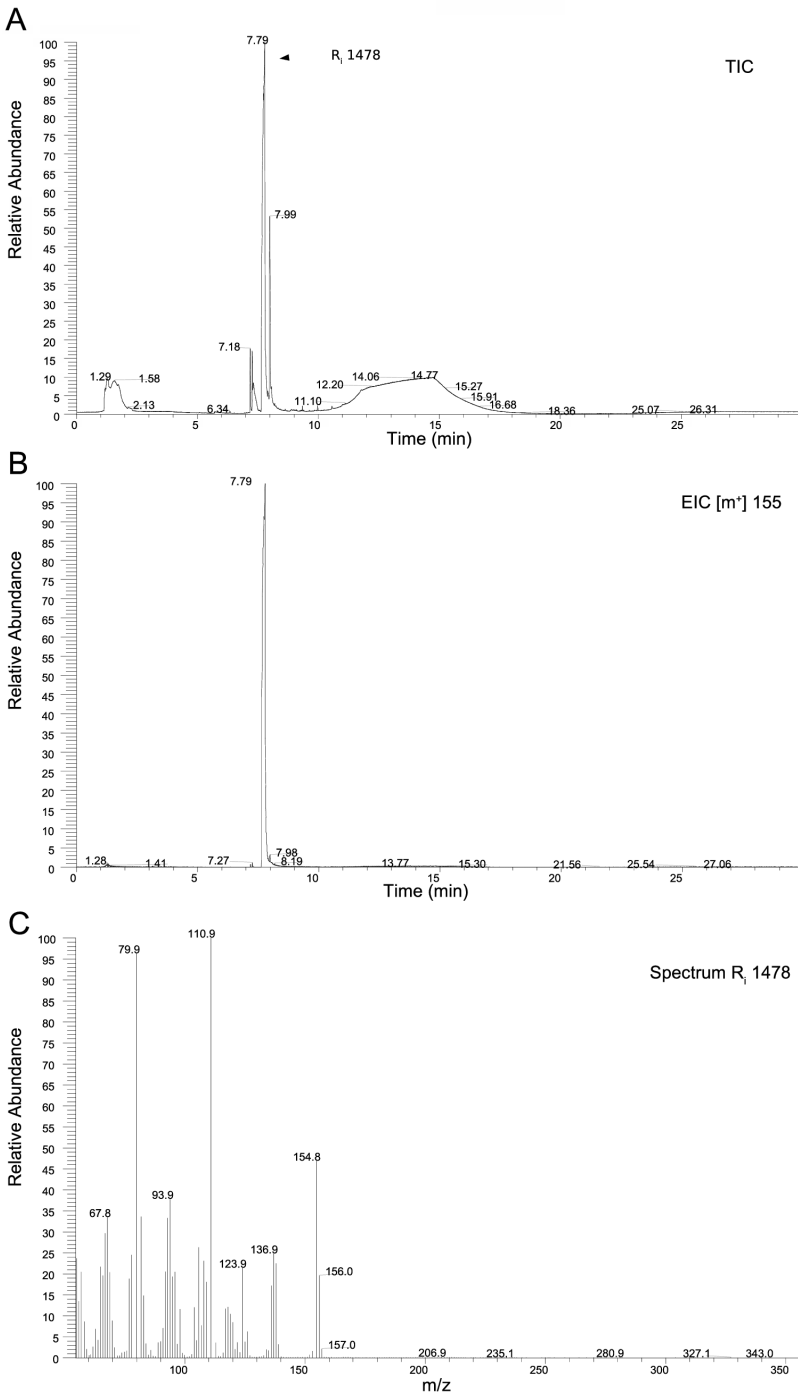
Supplementary Figure S1. PCR cycle optimization. Levels of HSS- and actin (ACT)-ecodng transcripts in the various leaf types were estimated by semi-quantitative RT-PCR. To ensure quantification during the exponential phase, amplification aliquots were analyzed after 20, 25, 30, and 35 cycles by gel electrophoresis. Arrowheads indicate PCR products with the expected size between 100 and 150 bp. Roman numerals refer to leaf types given in Fig. 1C. M, DNA ladder; C, no template control.



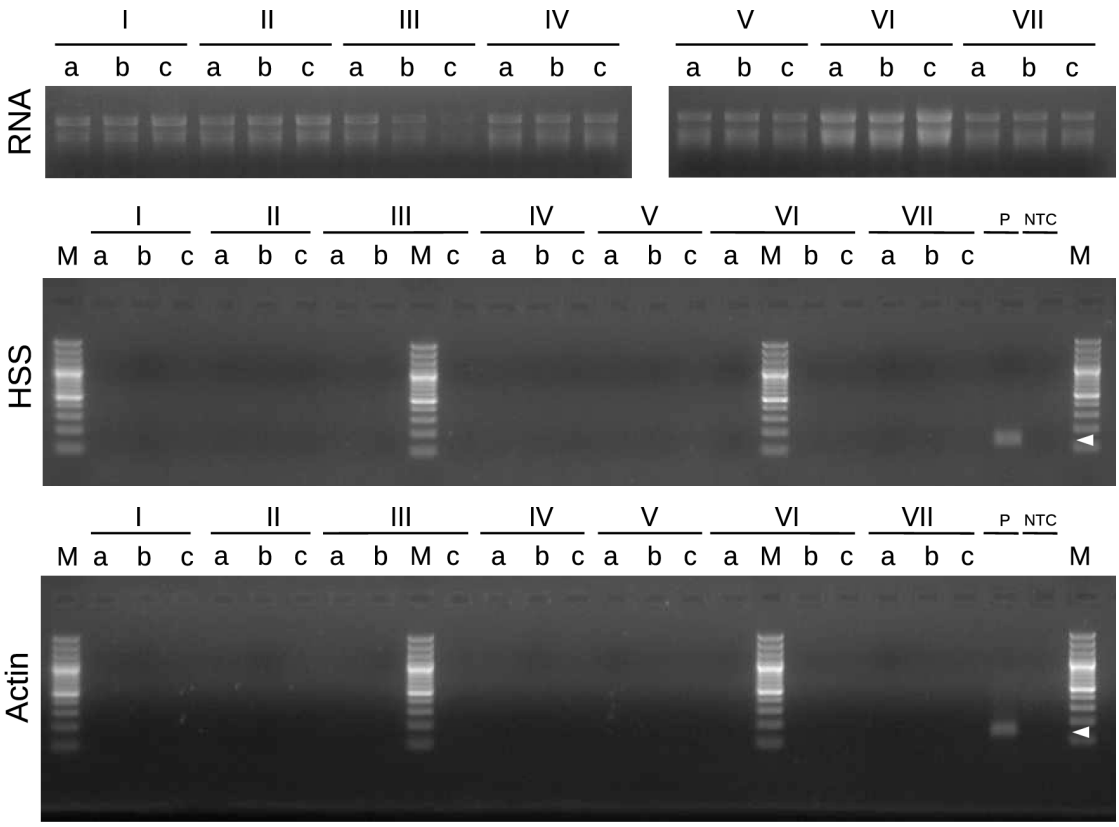
Supplementary Figure 2. Relative HSS transcript levels with actin (ACT) as reference gene. HSS transcript quantification in leaves subtending inflorescences of various developmental stages by using the same cDNA as for qRT-PCR given in Fig. 1B. The HSS transcript level of leaf type IV was set to one. Error bars indicate the standard deviation of three technical replicates (n=1).



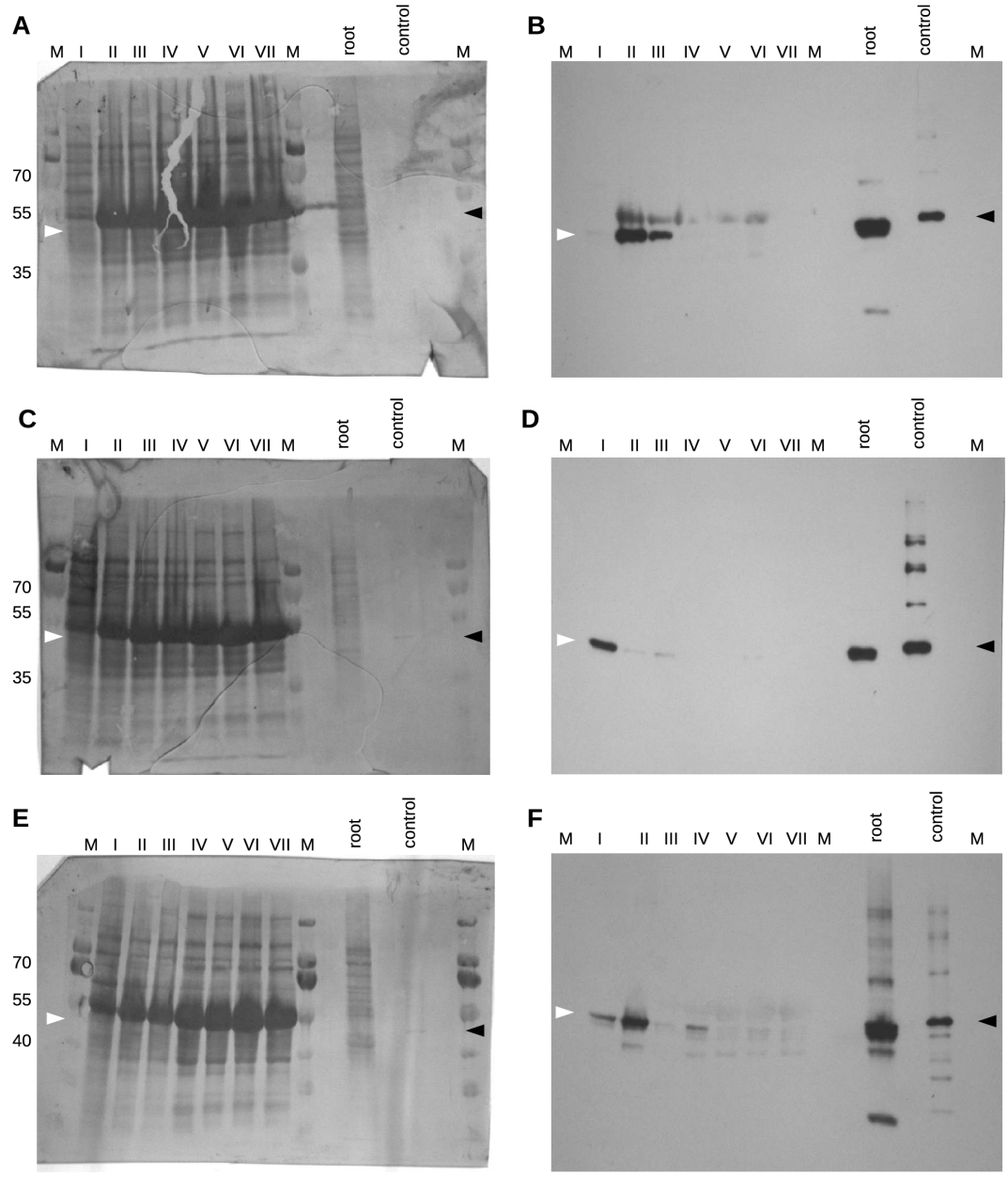
Supplementary Figure 3. Specificity tests of antibody labeling by pre-incubation of the HSS-specific antibody with soluble protein. Leaf cross-sections were labeled with HSS-specific antibody after pre-incubation with HSS (A, B), DHS (D, E), and bovine serum albumin (BSA) (G, H) at a molar ratio of antibody to added protein of 10:1 (B, E H) and 1:3 (A, D, G). A section labeled with HSS-specific antibody without pre-incubation (C) and a section labeled without primary antibody incubation to show unspecific binding of secondary antibody (F) served as controls. Only pre-incubation with soluble HSS resulted in a significant reduction of signal.



Supplementary Figure 4. MS-identification of the product accumulating in the leaf after feeding with putrescine. A, Total ion chromatogram of the spot co-migrating on TLC with the radioactive signal resulting from the [14C]putrescine feeding experiment. The largest peak shows a retention index (Kovats) at 1478. The two minor peaks with similar retention times show a typical fragmentation pattern of softeners (plasticizers). B, Extracted ion chromatogram of the same sample as in A for the [M+] of retronecine and its stereoisomer heliotridine with an m/z of 155. The peak is identical to the peak in A. C, Mass spectrum of the largest peak in A with the retention index of 1478. The peak was identified via the NIST database as being retronecine.

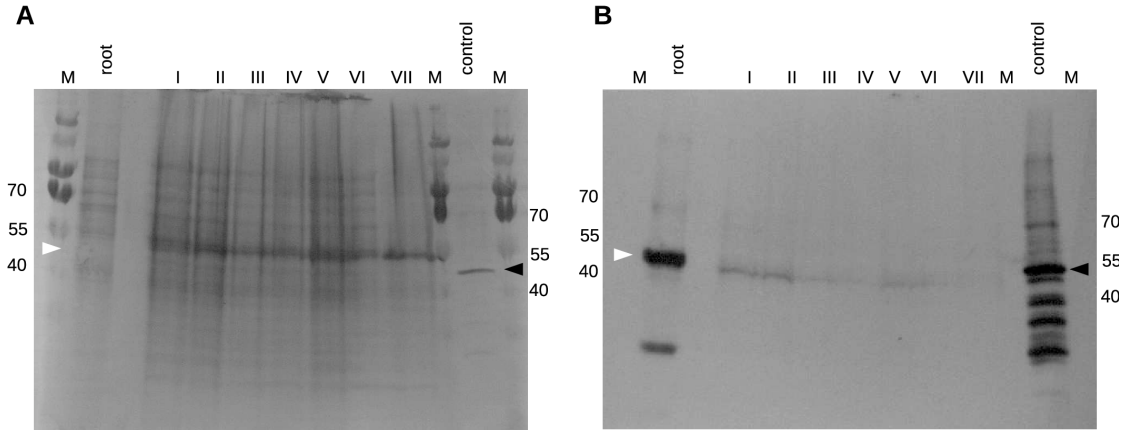


Supplementary Figure 5. Tests for contaminating gDNA in cDNA samples. Agarose gel electrophoresis of total RNA and PCR products resulting from amplification of RNA used for reverse transcription (no RT-control). Contamination of RNA with gDNA should result in fragments as shown in the positive control amplified with cDNA as the template in P, as the amplified sequence does not contain introns. Arrowheads indicate PCR products with the expected size between 100 and 150 bp. Roman numerals refer to the leaf stages from Fig. 4. The three biological replicates are indicated with a, b, and c. M, DNA ladder; NTC, no template control.



Supplementary Figure 6. Protein blot analyses of leaves subtending inflorescences of various developmental stages. Leaves were sampled in triplicate and protein extracts were blotted on separate membranes. A, C, and E, PDVF membranes stained with colloidal Coomassie G-250 dye to test for comparably loaded lanes. B, D, and F, X-ray film of the membranes corresponding to those given in A, C, and E, respectively, after ECL detection of horseradish-peroxidase generated bio-luminescence for quantification of HSS as illustrated in Fig. 4B. Crude protein extracts of leaves and the root (20 µg protein per lane) were separated via SDS-PAGE and blotted onto a PDVF membrane. Affinity-purified HSS-specific antibody was used for detection in combination with a secondary goat-anti-rabbit antibody conjugated to horseradish peroxidase. Recombinant HSS protein (25 ng) served as positive a control. The observed size difference between heterologous protein and the native protein is caused by His-tag fusion. As the antibody was generated against the protein heterologously expressed in *E. coli,* unspecific binding to contaminating *E. coli* proteins in lane "control" is expected. The apparent additional signal in root extracts was described earlier (Niemueller et al., 2012) and is caused by unspecific protein cleavage by proteases of soil bacteria. M, protein

mass standard (A-D: PageRuler™ Plus Prestained Protein Ladder, E-F: PageRuler™ Prestained Protein Ladder); I-VII represent the various developmental stages defined in Fig. 4A; numbers indicate the mass of the corresponding bands of a protein standard in kDa; white arrow head, indicates size of native HSS protein; black arrow head, size of recombinant HSS (with His-tag).



Supplementary Figure 7. Protein blot analyses of various developmental stages of inflorescences. A, PDVF membrane stained with colloidal Coomassie G-250 dye. B, ECL detection of horseradish-peroxidase generated bio-luminescence on an X-ray film. Protein extracts of inflorescences and roots collected from three independent plant individuals were pooled and separated via SDS-PAGE (20 µg protein per lane) and blotted onto a PDVF membrane. Affinity-purified HSS-specific antibody was used for detection in combination with a secondary goat-anti-rabbit antibody conjugated to horseradish peroxidase. Recombinant HSS protein (25 ng), which was heterologously expressed in *E. coli,* served as positive a control. The additional signal in root extract is attributable to unspecific protein cleavage (Niemüller et al., 2012). All analyzed developmental stages of inflorescences show only low protein levels of HSS. M, protein mass standard (PageRuler™ Prestained Protein Ladder); I-VII represents the various developmental stages defined in Fig. 4A; numbers indicate the mass of the corresponding bands of protein standard in kDa; white arrow heads indicate the size of native HSS and black arrow heads that of recombinant HSS (with His-tag).

4. CHAPTER 2

Two genes for the same function? - An *in silico* Substractive Approach for the Identification of Enzymes from Specialized Metabolism

Lars H. Kruse, Annika Engelhardt, Yu-Chen Wu, and Dietrich Ober

Botanisches Institut, CAU Kiel, Am Botanischen Garten 1-9, 24118 Kiel

Detailed author contributions are listed at the end of the thesis.

4.1. ABSTRACT

Pyrrolizidine alkaloids (PAs) are toxic secondary metabolites in several, not closely related plant families. So far, in all studied PA containing plants, the biosynthesis of PAs is restricted to specific cells and developmental stages of the plant. In *Symphytum officinale* (Comfrey, Boraginaceae) PA biosynthesis can be found in roots and young leaves subtending a young inflorescence. To this day, homospermidine synthase (HSS) is the only known enzyme of PA biosynthesis. It catalyzes the formation of the first pathway specific precursor homospermidine. HSS expression was found in bundle sheath cells of young leaves and in endodermis cells of the root. Under the assumption that also the genes of the whole PA biosynthesis pathway are differentially expressed, we sequenced transcripts of alkaloid and non-alkaloid producing tissues, assembled a *de novo* transcriptome and quantified the tissue specific expression of the sequences. Contigs that were up-regulated in PA producing tissues were extracted and compared to sequences that were up-regulated in the PA producing lower leaf epidermis of *Heliotropium indicum (*Indian heliotrope, Heliotropiaceae, Boraginales). We were able to identify candidate genes of several steps of the core biosynthesis of the PAs by blast guided annotation. Inter alia, diamine oxidases (DAO) that are hypothetically involved in the second step of PA biosynthesis, the oxidation of homospermidine and spontaneous cyclisation to a necine base precursor, were identified. We found two enzymes that are co-expressed with *hss* in young PA producing leaves (SoDAO1) and endodermis cells (SoDAO5), respectively, suggesting that these two enzymes were recruited for an identical function in different plant tissues. In addition, we were able to show that the new opportunities of next generation sequencing are a fast and convenient approach to identify candidate enzymes for pathways that can not be studied in model organisms. A data set of candidate genes that were generated in this study will be a valuable source for future experiments enlightening PA biosynthesis in Boraginales.

4.2. INTRODUCTION

Plants are adapted to their respective habitat by the production of numerous chemical compounds that allow them to cope with diverse biotic and abiotic stressors in their environment (Pichersky and Gang, 2000; Hartmann, 2007; Pichersky and Lewinsohn, 2011). These specialized metabolites, also called secondary metabolites, are synthesized by biosynthetic pathways that are characterized by a complex interplay of many specialized enzymes. The enzymes that are required for the production, transport, and storage of specialized metabolites account for approximately 10-20 % of the genes in plant genomes (Bevan et al., 1998; Somerville and Somerville, 1999; Pichersky and Gang, 2000; Pichersky and Lewinsohn, 2011). The evolution of those genes and of the resulting diversity in metabolites is believed to be mainly driven by gene duplication and successive divergence (Ohno, 1970; Ober, 2005; Pichersky et al., 2006; Innan and Kondrashov, 2010; Koch, 2014; Rensing, 2014; Panchy et al., 2016).

Alkaloids are one of the most important groups of specialized metabolites that are produced by approximately 20 % of plant species and are involved in defense against herbivores (Facchini, 2001; Facchini and St-Pierre, 2005). One diverse group of alkaloids are the pyrrolizidine alkaloids (PAs) (Hartmann and Ober, 2000). The occurrence of PA structures is scattered within the plant kingdom. The repeated independent recruitment of the fist pathway specific enzyme (homospermidine synthase, HSS), suggests that the complete PA biosynthesis evolved several times independently, making PAs an attractive model for studying convergent evolution of pathways (Ober and Hartmann, 1999a; Hartmann and Ober, 2000; Langel et al., 2011; Kaltenegger et al., 2013).

The biosynthesis of PAs can be divided in different parts: (1) the synthesis of the necine base, (2) the synthesis of the necic acid moiety, (3) the conjugation of the backbone, and (4) the individual modification of side chains in the resulting ester. The necine base synthesis starts with homospermidine, which is synthesized by HSS (step 1, Fig. 1) (Böttcher et al., 1993; Böttcher et al., 1994; Ober and Hartmann, 1999a). Homospermidine is believed to be oxidized by a diamine oxidase (DAO EC 1.4.3.22, step 2, Fig. 1) to allow spontaneous cyclisation to the necine base precusor 4,4'-iminobisbutanal (Robins, 1982; Kelly and Robins, 1988; Böttcher et al., 1993; Graser and Hartmann, 1997). Following steps of the necine base synthesis, necic

acid synthesis, and conjugation of the backbone are unknown so far. It is proposed that PAs are *N*-oxigenized by flavin-dependent monooxygenases (FMO, Step 10) for translocation and storage of PAs in the plant vacuole (Hartmann and Zimmer, 1986; Chang and Hartmann, 1998).

The site of biosynthesis of alkaloids is found to be class-specific (Facchini, 2001; Facchini and St-Pierre, 2005). Also, PAs show a remarkable variability regarding their site of synthesis. For almost every species or family, which is analyzed in more detail, a new site of synthesis on organ, tissue, or cellular level was identified (Moll et al., 2002; Anke et al., 2008; Niemüller et al., 2012; Irmer et al., 2015; Kruse et al., 2017). For the identification of enzymes involved in secondary pathways, several studies used the differential regulation of the participating enzymes (Gershenzon et al., 1992; Lange et al., 2000; Murata and Luca, 2005; Murata et al., 2008). Recently, Sievert et al. (2015) used laser-capture-microdissection, suppression subtractive hybridization (SSH), and subsequent 454 sequencing successfully for the identification of candidate enzymes in the PA producing plant *H. indicum* (Indian heliotrope, Heliotropiaceae). We modified the *in vitro* approach of Sievert et al. (2015) to an *in silico* method for the identification of candidate genes for PA biosynthesis in *S. officinale*. This approach results in the identification of putative homospermidine oxidases (copper-containing diamine oxidase, DAO) that might catalyze the second step of the pathway to the necine base. Additionally, we were able to identify other promising candidate enzymes for various steps of the PA biosynthesis by comparing PA producing with non-producing tissues and as well as by comparison of expressed contigs with *H. indicum*. We show that the combination of a substractive approach, NGS sequencing, and knowledge about the physiology of a pathway can be a powerful tool to identify related genes.

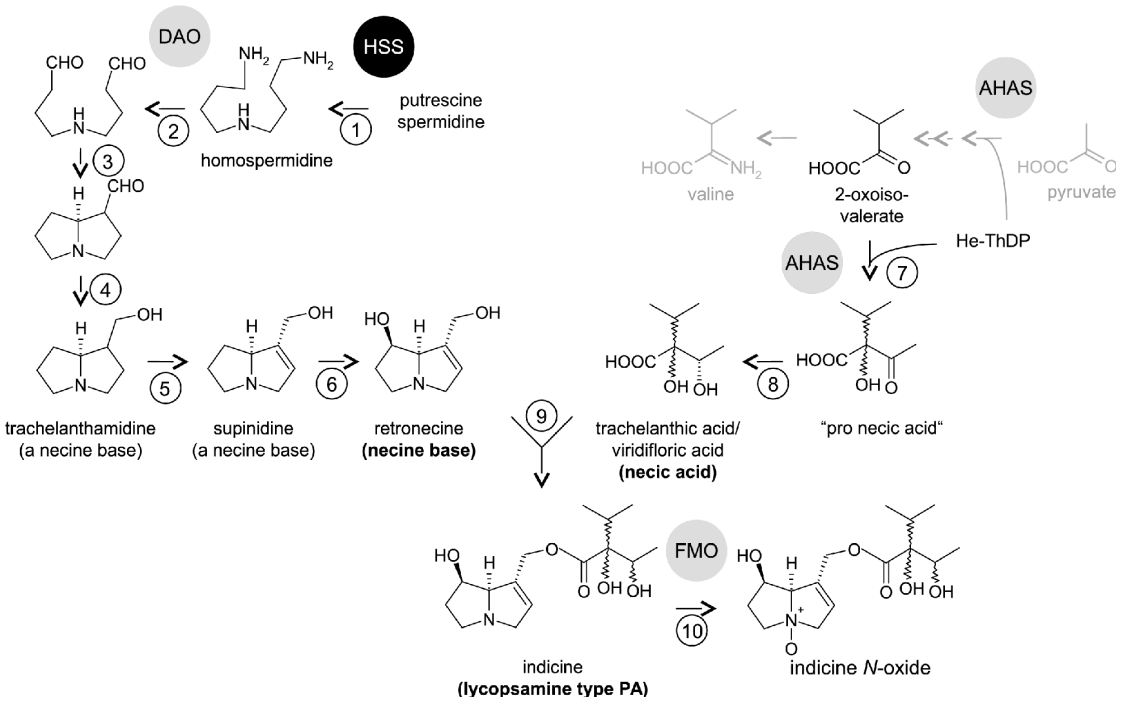


Figure 1. Postulated pyrrolizidine alkaloid biosynthesis. The proposed biosynthesis of the necine base is shown on the left side, that of the C7-necic acid of lycopsamine type-PAs on the right side. Indicine *N*-oxide is shown as example for a lycopsamine type-PA. Steps for which an enzyme is postulated, are indicated by a number. Enzymes for which an experimental evidence exist, are highlighted by a black circle (established finding) or a grey circle (indications) (Robins, 1982; Kelly and Robins, 1988; Böttcher et al., 1993; Chang and Hartmann, 1998; Weber et al., 1999).

4.3. RESULTS

4.3.1 Study design, extraction of total RNA and generation of RNA-Seq data

PA biosynthesis in *S. officinale* is restricted to two distinct locations in the plant, namely young primary roots and young leaves next to inflorescences (Niemüller et al., 2012; Kruse et al., 2017). In this study, we used this phenomenon to identify genes, which are only present in tissues that are capable of PA biosynthesis. Therefore, we used (1) young leaves from plants without inflorescences (PA -, Fig. 2A), (2) young leaves in close proximity to inflorescences (PA +, Fig. 2B), and (3) roots (root PA +, Fig. 2C) as the predominant site of PA biosynthesis in *S. officinale* (2007). From five leaf samples, three from plants without and two from plants with inflorescences, total RNA was extracted, an aliquot was reverse transcribed and the resulting cDNA was analyzed with semi quantitative PCR (sqPCR) for transcripts of *hss* and *actin* (Fig. 2D) to ensure that all samples meet the requirements for “leaf PA +” and “leaf PA -” that were formulated above. We chose a sample with (leaf sample 5, PA +) and a sample without (leaf sample 2, PA -) detectable HSS transcripts for sequencing (Fig. 2D). Root samples were harvested from plants in axenic and green house culture and analyzed for transcripts of HSS and actin (root PA +, see above). Both samples showed detectable amounts of HSS transcripts but the sample from the greenhouse (root sample 7, Fig. 2D) showed a stronger signal and was regarded, with respect to transcription level, to be more similar to the leaf samples (root PA +). After the selection process, the already extracted RNA was further purified via RNA Clean-up columns and RNA integrity was controlled by RNA gelelectrophoresis, which indicated undegraded RNA (Fig. 2E). Before sequencing, the sequencing company (BGI, Shenzen, China) performed a second quality check with an Agilent 2100 Bioanalyzer. All samples passed with RNA integrity numbers (RIN) of 8.0 to 8.6 and 28S/18S ratios of 1.5 to 1.7 and confirmed sufficient RNA quality. Illumina® sequencing resulted in approximately 17 million adapter-free 90 bp paired-end reads per sample, found to be suitable for assembly by FastQC analysis (Andrews, 2010).

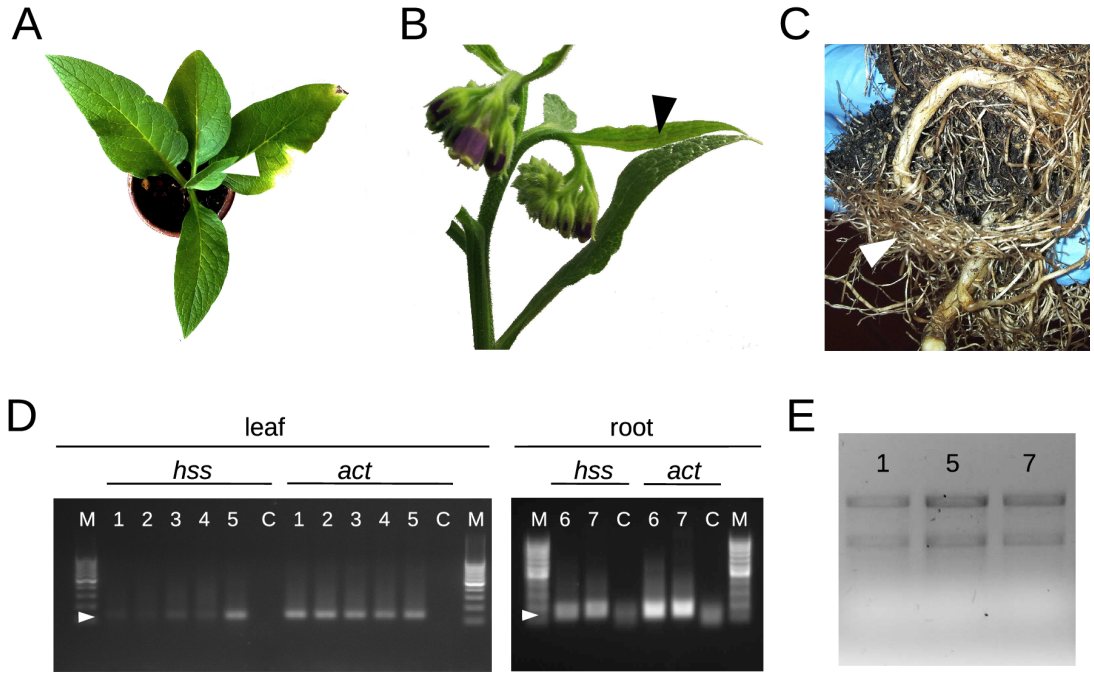


Figure 2. Selection of samples for transcriptome analysis. A, *S. officinale* plant growing as rosette during vegetative phase (1st year). Leaf that was used as sample for transcriptome analysis is indicated with a white arrow (leaf PA -, lane 1 in panel D). B, Part of a *S. officinale* plant with developed inflorescences (2nd year). Black arrow indicates leaf which was chosen for transcriptome analysis (leaf PA +, lane 5 in panel D). C, Roots of a *S. officinale* plant. Young, white primary roots were chosen for sequencing and are indicated by a white arrow (root PA +, lane 7 in panel D). D, Semi-quantitative PCR analysis of HSS and actin transcripts in different leaf and root samples from which three were chosen for NGS sequencing. E, Formaldehyde RNA-gelelectrophoresis of samples 1 (leaf PA -), 5 (leaf PA +), and 7 (root PA +). Distinct bands for 28S and 18S RNA indicating non-degraded RNA suitable for NGS sequencing. C, no template control.

4.3.2 Optimization and evaluation of transcriptome assembly

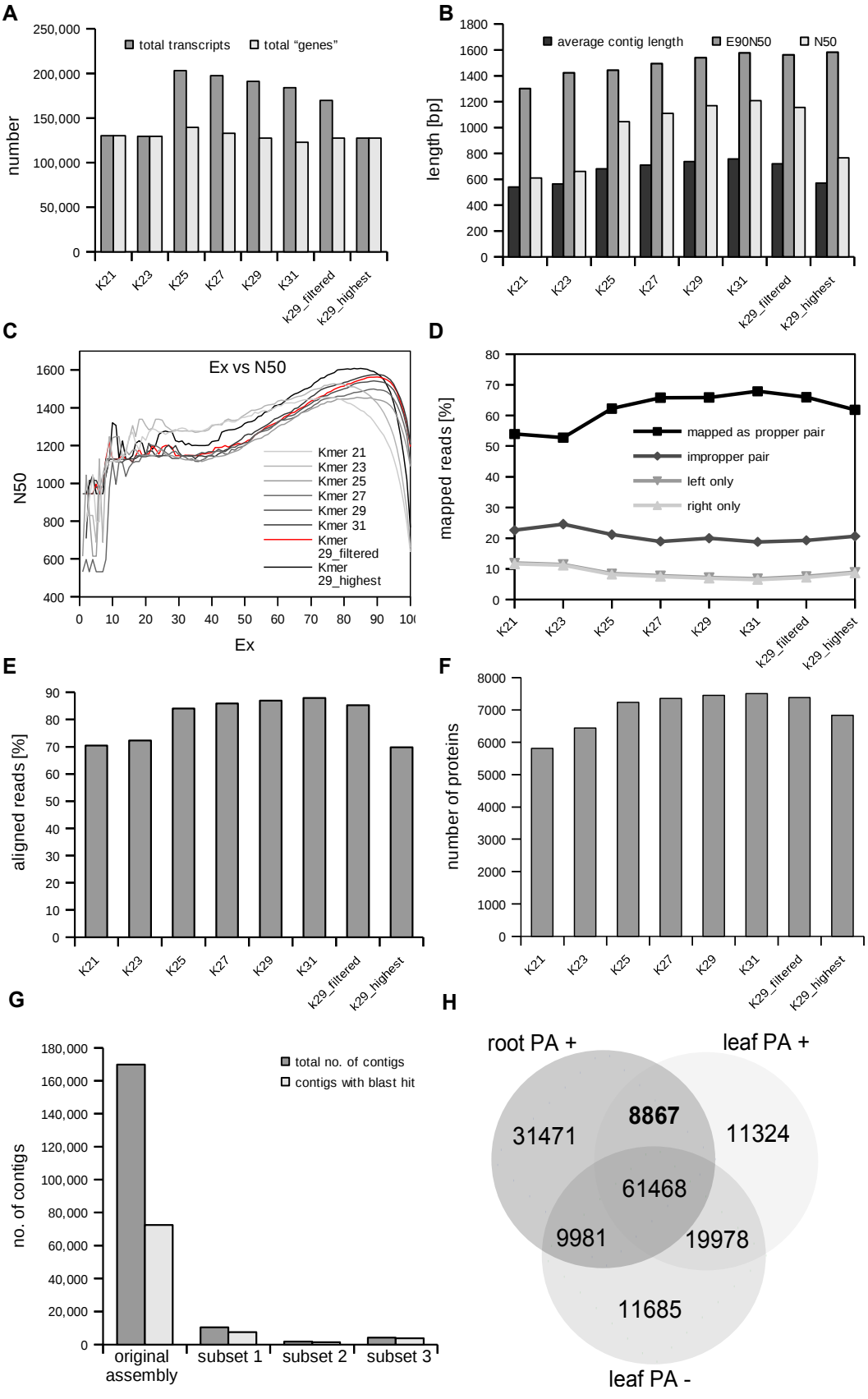
To generate a most optimal transcriptome assembly for our analyses, different assembler parameters were chosen and compared as follows: the kmer length was varied from 21 to 31 and basic assembly metrics were calculated for each assembly separately (Fig. 3). Contigs that share high sequence similarity are grouped into so called “genes” by the Trinity assembler. The term “gene” is believed to give a better representation of the number of real genes that are assembled than single isoforms (Fig. 3A). The highest number of “genes” and transcripts is assembled when using the default kmer of 25 (Fig. 3A). A lower kmer value reduces the number of transcripts closer to the number of “genes”, whereas a higher value reduces both, the number of transcripts and “genes” (Fig. 3A). In order to reduce the redundancy of highly similar isoforms and chimeras that do not represent real or new transcripts, we

tested two approaches published earlier by Yang and Smith (2013). The first approach was to filter out the lowest expressed isoform of each “gene” and the second to keep only the highest expressed isoform of each gene. When both approaches were applied to the kmer 29 assembly, the first approach reduced the overall number of transcript while mostly discarding chimeras and lowly expressed short sequences. The second approach reduced the number of transcripts to the number of genes more strictly. This effect is visible when comparing average contig lengths, N50 and E90N50 contigs (Fig. 3B). Whilst the average contig length only gives the statistical mean of length of all transcripts, the length of N50 and E90N50 contig gave more relevant information about the quality of assemblies. The N50 indicates that half of all assembled bases are found in contigs with at least this length and is mainly used for genome assemblies. The E90N50 is a better value to measure transcriptome assembly quality. Here, the contig length is calculated based on the top most highly expressed transcripts that represent 90% of the total normalized expression data (Haas, 2016). The E90N50 excludes short low expressed transcripts and therefore is significantly longer than the original N50 contig. kmer 31 has the highest E90N50 value, which indicates that in general longer reads are present in this assembly than in others (Fig. 3B). When examining the ExN50 profiles (average length of the top most expressed transcripts that represent “x”% of the total normalized expression data) in more detail, it is also observed that a longer kmer improves overall assembly quality as indicated by the peak of longest contigs shifts towards 90 % with higher kmer values (Fig. 3C) (Haas, 2016). We tested for the completeness of the assemblies in order to show how many reads are actually included in the assembly by mapping the original reads back to the assembled transcripts. Obviously, higher kmer values provide a better sequence assembly by having a higher proportion of proper paired reads that align and higher overall percentage of aligned reads (Fig. 3D and Fig. 3E). To compare the quantity of full-length contigs in each assembly, all transcripts were aligned against a database of known proteins and the number of proteins were counted that are covered by at least 80 %. For this analysis we used a database that contained only protein sequences from UniProtKB/Swiss-Prot that belong to green plants (Viridiplantae, *Taxonomy ID:* 33090). At kmer of 31, the highest number of proteins are covered by full-length transcripts and full-length coverage increases with rising kmer values (Fig. 3F).

Accounting for all measures, a kmer of 31 seems to be best fitting for a transcriptome assembly from our given data-set when using the Trinity assembler. We were able to

compare sequences that had been identified earlier from *S. officinale,* e.g., HSS, DHS and several diamine oxidases, by classic PCR and cloning techniques to the assembled sequences. We found that closely related paralogs (e.g. of diamine oxidases) are better resolved in the filtered kmer 29 assembly while maintaining only slight declines in overall assembly qualitiy. Therefore, we chose this assembly for further analysis, despite the fact that the kmer 31 assembly shows better assembly metrics.

Figure 3. Comparison of assemblies and contig subsets. The effect of kmer variation was analyzed with respect to different parameters that were calculated by using scripts included in the trinity package. A, Total number of assembled transcripts and "genes". Genes are defined by the trinity assembler as a cluster of isoforms. B, Average contig length, E90N50 contig length, and N50 contig length for assembly. C, N50 contig length was calculated based on a proportion of contigs with a given expression. D, mapping of original reads against the different assemblies as proportion of total reads. E, total number of proper aligned reads as proportion of the total read number. F, contigs of each assembly were blasted against a protein database. The total number of proteins that are covered by at least 80 percent by the contigs is given. G, the number of contigs in subsets (1-4) that were generated using different comparisons against the original assembly. H, Venn diagram of the three sequenced samples (leaf PA +, leaf PA -, and root PA +). Number of shared contigs for each section are given. The number of contigs that are shared between the two PA producing sites is given in bold letters. →


A
total transcripts
total "genes"
number
250,000
200,000
150,000
100,000
50,000
0
K21
K23
K25
K27
K29
K31
k29_filtered
k29_highest
B
average contig length
E90N50
N50
length [bp]
1800
1600
1400
1200
1000
800
600
400
200
0
C
Ex vs N50
N50
Ex
Kmer 21
Kmer 23
Kmer 25
Kmer 27
Kmer 29
Kmer 31
Kmer 29_filtered
Kmer 29_highest
D
mapped reads [%]
mapped as propper pair
improper pair
left only
right only
E
aligned reads [%]
F
number of proteins
G
no. of contigs
total no. of contigs
contigs with blast hit
original assembly
subset 1
subset 2
subset 3
H
root PA +
leaf PA +
leaf PA -
31471
8867
11324
61468
9981
19978
11685

4.3.3 *In silico* substractive approach

Supression substractive hybridisation (SSH) can successful identify genes belonging to cell specific biosynthetic routes (Diatchenko et al., 1999; Sievert et al., 2015). Sievert et al. (2015) generated cDNA samples from epidermal cells of *H. indicum* characterized by the absence or presence of PA biosynthesis, and used these for generation of a SSH library. The retained enriched transcripts were sequenced by the 454 technique and used for sequence assemblies. We adapted this strategy to an *in silico* approach by which candidate sequences were enriched via measuring their sample specific expression.

In the final *de novo* assembly we got 169,735 contigs in total, from which 72,469 contigs resulted in significant blast hits when blasted against UniProtKB/Swiss-Prot database (Fig. 3G). To reduce the data set in a first step to sequences expressed in specific tissues, the tissue (leaf PA +, leaf PA -, and root PA +) specific expression for all assembled contigs was analyzed. The vast majority of transcripts (61,468) were detected in all tissues and, therefore, might belong to primary metabolism (Fig. 3H). The root sample had the most uniquely expressed contigs (Fig. 3H). The lowest number of shared features is found when comparing the root with one of the leaf samples; 9,981 (Leaf PA - vs root PA +) and 8,867 (root PA + vs. leaf PA +) contigs (Fig. 3H). Aiming to identify genes belonging to PA biosynthesis we tested three different approaches to enrich these specific sequences. To exclude genes from the analysis with only slight differences in their expression level, we included only contigs into the analysis that are at least 4-fold differently expressed between the samples. For a first dataset, we extracted all contigs that are higher expressed in the PA producing leaf (leaf PA +) OR in roots (root PA +) than in the non-PA producing leaf (leaf PA -). This reduced the number of contigs to 10,356 with 7,335 blast hits (subset 1, Fig. 3G). Amongst others, the HSS belonged to the highest expressed transcripts in roots and young PA producing leaves and served as a positive control. Also in the further approaches, HSS served as a proof-of-concept for our comparisons. For the second dataset, we kept only those sequences that are upregulated in PA producing leaves (leaf PA +) AND in roots in comparison to non-PA producing leaves (leaf PA -). Under the assumption that genes that belong to central steps of PA biosynthesis are present in all PA producing tissues, this approach should retain candidate genes while discarding transcripts related to primary metabolism because

they are present in all samples with similar expression levels and of leaf-specific metabolism like photosynthesis. This resulted in a set of 1,817 contigs from which 1,358 had significant blast hits (subset 2, Fig. 3G). As before, HSS was found to be one of the most highly expressed transcripts. For the third dataset, we made use of a data-set which was generated for *Heliotropium indicum* (Indian heliotrope, Boraginales) in the same manner as described in this study (unpublished data). Because *H. indicum* produces PAs only in the lower leaf epidermis, only those contigs were retained which were at least 4 times upregulated in the lower leaf epidermis in comparison to the upper epidermis. In the *H. indicum* dataset, HSS was again one of the most prominent sequences. After a tblastx search of *S. officinale* subset 1 against all *H. indicum* sequences, the number of *S. officinale* sequences which had a homolog in *H. indicum* was found to be 4184 from which 3765 had a blast hit against UniProtKB/Swiss-Prot database (subset 3, Fig. 3G).

The initial assembly and all generated subsets were functionally annotated and assigned (Supp. Fig. 1). Whereas the original assembly includes the most diverse collection of gene ontology (GO) terms, subsets 1 to 3 show a reduced diversity (Supp. Fig. 1A-D). Especially the proportion of contigs assigned to metabolic processes is increased in all three subsets. Therefore, these enriched subsets should contain promising candidate sequences for enzymes involved in PA biosynthesis. That the subsets always retained HSS, the first pathway specific enzyme of PA biosynthesis, is regarded as a good proof-of-concept that our approach is able to enrich PA pathway related transcripts.

4.3.4 Identification of candidate genes for postulated steps of PA biosynthesis

We used Blast® for annotation and expression analysis in order to identify candidate genes. Each of the three generated subsets of *S. officinale* was manually searched for transcripts coding for enzymes that could be involved in one of the postulated steps of PA biosynthesis (Robins, 1982; Kelly and Robins, 1988; Böttcher et al., 1993; Böttcher et al., 1994; Graser and Hartmann, 1997; Langel et al., 2011). We hypothesized that in *S. officinale* and *H. indicum* PA biosynthesis should involve functional similar genes because only one independent duplication event of HSS, early in the evolution of the Boraginales, suggests only one common origin of PA biosynthesis in this lineage (Reimann et al., 2004). Therefore sequences that are

found in the PA producing parts of *S. officinale* and have closely related homologs in *H. indicum* lower leaf epidermis are more promising candidates than genes that are only present in one of these plant tissues. Sequences coding for diamine oxidases, alcohol dehydrogenases, acetohydroxyacid synthases, cytochrome P450 monooxygenases, flavonol synthases/flavanone 3-hydroxylases, flavin-dependent monooxygenases, MATE transporters, and for *S*-norcoclaurine synthases were identified as candidates using this strategy (Tab. 1).

Table 1. Identified contigs of candidate genes for PA biosynthesis. For each candidate the postulated steps of PA biosynthesis in Fig. 1 and and the subset (Fig. 3G) in which they were found are given. Relative expression of contigs is given for the non-PA-producing leaf (leaf -), the PA-producing leaf (leaf +) and the roots (root) as TMM normalized TPM values. ADH, alcohol dehydrogenase; AHAS, acetohydroxyacid synthase; CYP450, cytochrome P450 monooxygenase; DAO, diamine oxidase; FLS, flavonol synthase/flavanone 3-hydroxylase; FMO, flavine-dependent monooxygenase; MATE, mate transporter; NCS, *S*-norcoclaurine synthase.

postulated enzyme	name	step in Fig. 1	contigs	Subsets	Leaf PA-	Leaf PA+	Root PA+
HSS	SoHSS	1	TRINITY_DN44574_c2_g3_i1	1, 2, 3	8.7	238.8	335.7
ADH		4	TRINITY_DN31934_c0_g1_i2	1, 3	6.9	4.9	49.9
			TRINITY_DN35702_c1_g1_i3	1, 2, 3	2.6	4.1	28.0
			TRINITY_DN34716_c0_g2_i1	1, 2	8.7	10.9	91.3
AHAS	SoCSU2	7	TRINITY_DN45393_c0_g1_i1	1, 2, 3	0.0	0.7	167.8
CYP450		6	TRINITY_DN16267_c0_g1_i1	1, 2, 3	3.8	4.3	78.9
			TRINITY_DN45478_c0_g1_i2	1, 2	0.1	0.2	18.3
			TRINITY_DN48837_c0_g4_i1	1, 2	0.1	0.2	12.5
			TRINITY_DN32775_c0_g1_i1	1	0.0	0.0	18.5
			TRINITY_DN40382_c0_g1_i1	1, 2, 3	9.9	11.7	92.3
DAO	SoDAO1	2	TRINITY_DN46391_c0_g1_i4	1	8.0	56.2	0.8
	SoDAO2a		TRINITY_DN47237_c0_g1_i11	1	1.9	1.0	14.4
	SoDAO2b		TRINITY_DN47237_c0_g1_i5	1, 2	0.0	1.5	8.7
	SoDAO5a		TRINITY_DN46391_c0_g1_i5	1	0.0	0.0	70.5
	SoDAO5b		TRINITY_DN46391_c0_g1_i3	1, 3	0.0	0.0	78.9
FLS		6	TRINITY_DN2790_c0_g2_i1	1	0.0	0.0	19.0
FMO		10	TRINITY_DN47802_c2_g1_i1	1, 2	0.3	0.3	9.9
			TRINITY_DN39746_c0_g1_i1	1	0.0	0.0	10.8
			lTRINITY_DN36990_c1_g2_i1	3,	32.3	56.2	0.2
			TRINITY_DN43090_c0_g1_i1	1, 3	7.6	4.6	31.5
MATE		-	TRINITY_DN48867_c1_g1_i2	1, 2	0.0	0.2	14.6
			TRINITY_DN25491_c0_g1_i1	1, 2, 3	5.0	5.8	50.8
NCS		3	TRINITY_DN43935_c1_g2_i1	1, 3	0.0	0.0	53.2
			TRINITY_DN33358_c0_g1_i1	1	0.0	7.0	20.0

4.3.5 Transcription levels of DAO genes in PA producing tissues

In the first pathway specific step of the biosynthesis of PAs, the formation of homospermidine is catalyzed by HSS (Böttcher et al., 1993). Therefore, we focused on the second step of necine base synthesis, the oxidation of homospermidine and the potential self-cyclisation to necine base precusors, because substrates of later steps of the pathway are not commercially available. Analyzing the transcriptome data of *S. officinale,* we found five contigs coding for diamine oxidases (SoDAO1, 2, 3, 4, and 5) from which three show up in one of the three subsets as a result of the differential approach (SoDAO1, 2 and 5; Tab. 1). Analyses with quantitative real-time PCR (qRT-PCR) show a unique expression pattern of each enzyme (Fig. 4). First, we analyzed PA producing leaves (leaf PA +) in comparison to non-PA producing leaves (leaf PA -). SoDAO1 is co-expressed with HSS in young PA producing leaves (leaf PA +, Fig. 4A). Also SoDAO2 shows an expression pattern which seems to be linked to HSS expression (Fig.4A). SoDAO3-4 show an expression pattern, which seems not to be correlated with HSS expression in leaves (Fig. 4A). SoDAO5, which showed up in subset 1 and 3, is not expressed in leaves (Fig. 4A). In a more detailed analysis of DAO expression in PA producing leaves, we harvested leaves subtending inflorescences at seven different time points during flower development, because a previous study indicated that PA biosynthesis varies in different developmental stages (see chapter 1, Kruse et al., 2017) and analyzed them on transcript level. We found that HSS expression reaches a maximum in stage II and declines in the following stages (Fig. 4B). In comparison to HSS expression, SoDOA1 and 2 show a pattern similar to the HSS expression with a maximum transcript level in stage II (Fig. 4B). SoDAO4 transcript level is increasing until stage III and declines to a high persistent level from stage IV to VII. SoDAO3 shows a constantly low transcript level during all stages (Fig. 4B). As before, SoDAO5 expression is not detectable in leaves. In summary, in our analysis of PA producing leaves, only SoDAO1 and SoDAO2 show co-expression with HSS.

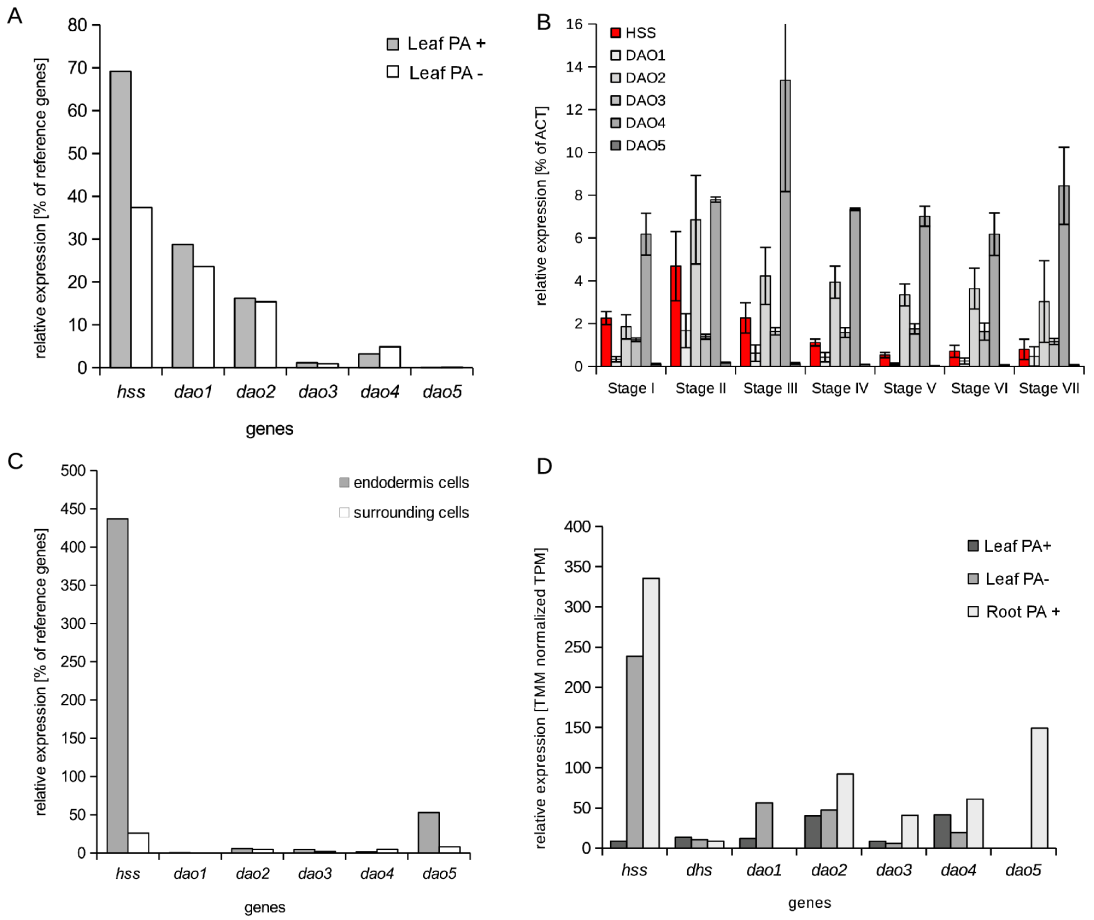


Figure 4. qRT-PCR comparison of DAO transcript levels in PA and non-PA producing tissues. A, Analysis of PA producing leaves subtending an inflorescence and more distant leaves. B, Analysis of endodermis cells that were isolated using LCM microscopy in comparison to the remaining root cells. C, expression analysis of diamine oxidases in comparison with HSS and DHS. Expression analysis was performed by mapping the original reads against all contigs using RSEM and subsequent differential expression analysis with EdgeR. Expression values are given as TMM (Trimmed Mean of M-values, Robinson and Oshlack, 2010) normalized TPM (Transcripts Per Million transcripts).

Besides young leaves, *S. officinale* produces PAs also in young roots (Frölich et al., 2007). As the endodermis was identified as the site of HSS expression in roots (Niemüller et al., 2012), we analyzed the expression of DAOs in the endodermis and the adjacent parenchyma and vascular cells. Therefore, we isolated endodermis cells by laser capture micro-dissection microscopy and quantified transcript levels of DAOs and HSS (Suppl. Fig. 4). As expected, HSS transcript levels were higher in the endodermis cells in comparison to the surrounding cells confirming that endodermis cells were quantitatively enriched by this method (Fig. 4C and Suppl. Fig. 4). Transcript levels of SoDAO2-4 are at a low level compared to HSS in both samples

(Fig. 4C). As HSS, but unlike the other DAOs, SoDAO5 shows a relatively high level of transcript in endodermis cells compared to surrounding cells (Fig. 4C). Of note, transcripts for SoDAO1 were hardly detectable in the root (Fig. 4C).

The results of the qRT-PCRs are well in accordance with the expression data generated by NGS sequencing and alignment-based quantification of transcripts (Fig. 4D). The data confirmed that HSS is highly expressed in roots and PA producing leaves. SoDAO2 to 4 are expressed in all tested tissues and do not show co-expression with HSS. SoDAO1 is expressed specifically in PA producing leaves but not in roots (Fig. 4D). In contrast, SoDAO5 is expressed only in roots (Fig. 4D). This suggests the possibility that two different DAOs were recruited for same function in the biosynthesis of PAs but in different tissues, namely SoDAO1 in young leaves and SoDAO5 in roots.

4.3.6 Sequence similarity and functional amino acids

SoDAO1 and SoDAO5 share 81 % identity on amino acid level (85 % on nucleic acid level). They are highly similar to two DAO candidates (HiDAO3, 5) that were identified in the PA producing lower leaf epidermis of *H. indicum*. Phylogenetic analyses revealed that both, SoDAO1 and 5, cluster together in a phylogeny of DAO-coding cDNAs of plant origin with three DAOs of *H. indicum* (Fig. 5, Cluster I). The phylogeny suggests a gene duplication that gave rise to SoDAO1 and 5. Most sequences from this cluster are predicted to be secretoric (Fig. 5 and Suppl. Tab. 1) which raises questions with regard to the cytosolic localization of HSS. Cluster I is well separated from another cluster (Cluster II) that contains an already characterized DAO (*N*-methylputrescine oxidase, MPO) from tobacco that is involved in the nicotine biosynthesis (Heim et al., 2007; Katoh et al., 2007; Naconsie et al., 2014). Both sequences, SoDAO1 and 5, seem to be functional and contain the typical functional amino acids of copper-containing amine oxidases: three conserved histidine residues that coordinate a copper ion in the active site of the mature protein and a conserved active site tyrosine that can be post-translationally modified into the redox co-factor topaquinone, the characteristic active site structure of copper containing DAOs (Kumar et al., 1996).

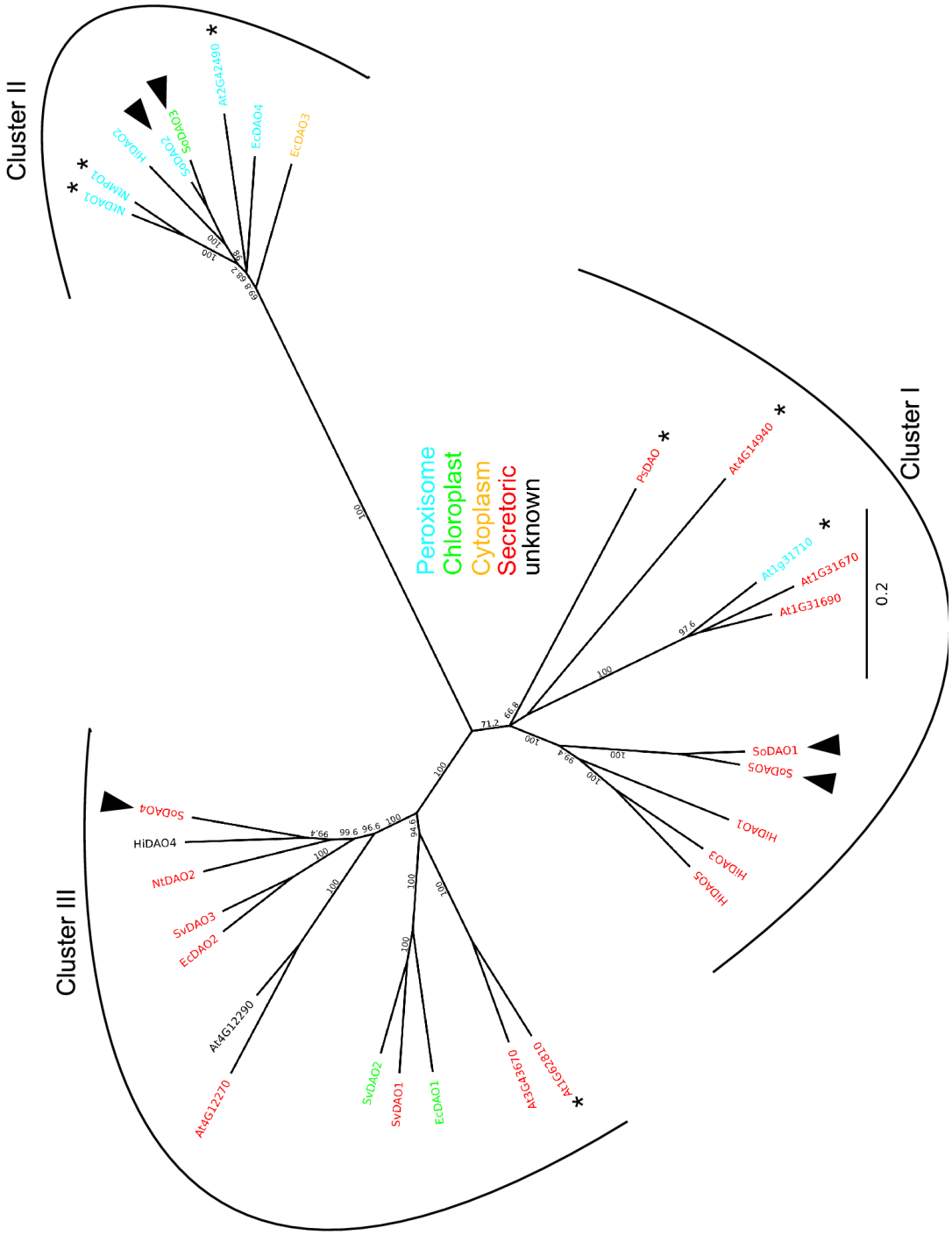


Figure 5. Neighbor-joining tree of diamine oxidase encoding sequences. Bootstrap values are based on 500 replicates. Colors indicate subcellular localization of the respective sequence. Predictions for the cell-specific localization are made by the prediction softwares *PSORT*, *TargetP*, and *Ppero*. An asterix indicates experimental evidence for the localization of the respective sequence on protein level. Arrowheads label the sequences of *S. officinale*.

4.3.7 Heterologous expression of diamine oxidases

In order to find out which of the identified DAOs is involved in PA biosynthesis we tried to heterologously express and characterize all five sequences *in vitro*. Therefore, several expression protocols in *Escherichia coli* were tested until we were able to purifiy soluble protein. Preliminary analyses of heterologously expressed and purified protein suggested the presence of truncated versions of the expressed protein. Therefore, each sequence was analyzed for potential internal Shine-Dalgarno sequences and following alternative start codons (Shine and Dalgarno, 1974). The Shine-Dalgarno sequence acts as ribosome binding site in bacterial mRNA. As this sequences does not have any effect in the plant but might occur randomly, they could act as alternative translation initiation sites in *E. coli,* leading to the expression of truncated recombinant protein (Parret et al., 2016). Using site-specific mutagenesis, we modified the predicted Shine-Dalgarno sequence to render them inactive without modifying the encoded amino acid sequence (Tab. 2). After this step, we were able to express each of the DAOs using the respective protocols in soluble form in *E. coli.* Unfortunately, only SoDAO2 showed detectable activity with the tested substrates under our assay conditions (preliminary results for putrescine: 2575 pkat/mg, and homospermidine: 9,3 pkat/mg). Optimization of the expression and assay conditions is currently running in our lab. Diamine oxidases seem to be difficult to express heterologously as protocols provided by literature are exceptionally diverse and optimized for each particular enzyme (Cai and Klinman, 1994; Schilling and Lerch, 1995; Cai et al., 1997; Koyanagi et al., 2000; Elmore et al., 2002; Heim et al., 2007; Katoh et al., 2007; McGrath et al., 2010; Rosini et al., 2012; Sun et al., 2012; Naconsie et al., 2014; Zarei et al., 2015). As most protocols involve the expression in eukaryotic hosts, we currently started to establish expression protocols for the DAOs of interest in the yeast *Pichia pastoris*.

4.4. DISCUSSION

4.4.1 Different scenarios of PA pathway evolution require different subsets for candidate gene identification

Contrary to many investigated PA producing plants with only one site of PA biosynthesis as *Senecio vernalis*, *Senecio vulgaris*, *Eupatorium cannabinum*, *Petasitis hybridus, and Heliotropium indicum*, *Symphytum officinale* has been shown to have two sites of biosynthesis: the endodermis cells of the root and the bundle sheath cells of young leaves subtending an inflorescence (Anke et al., 2004; Niemüller et al., 2012; Kruse et al., 2017). There are two likely scenarios of the evolution of a second site of PA biosynthesis in *S. officinale*: In scenario 1, the same genes that are involved in PA biosynthesis in roots are also involved in PA production in leaves. In scenario 2, genes involved in PA biosynthesis evolved independently in roots and young leaves. The two scenarios become important regarding the subsets generated in this study and the included candidate genes. In case of subset 1, contigs that are upregulated in at least one PA producing tissue, were enriched. This subset contains the highest number of contigs, which makes it more difficult to identify candidate sequences of PA biosynthesis. On the other side, this is the most resilient subset against the risk discarding sequences that are responsible for steps of the PA synthesis in only one tissue (false negatives). This approach would be suitable when genes of PA biosynthesis in roots and leaves evolved independent from each other as suggested in scenario 2. Subset 2 is the most stringent in terms of discarding contigs because contigs were only kept if they show up-regulation in both PA producing tissues. This results in a rather smaller subset, compared to subset 1, which could be beneficial for identifying genes, when PA biosynthesis is coherent in both PA producing tissues as suggested in scenario 1. In this case there is a high risk of discarding false negatives for the case that there are individual, tissue-specific genes of PA biosynthesis. Additionally, we compared enriched contigs of subset 2 with those of the PA producing *H. indicum* (Heliotropiceae), leading to an improved subset (subset 3). With this approach, the number of enriched sequences was significantly reduced and the risk of discarding false negatives was minimized, because sequences were not discarded when they were only expressed in one of the two PA producing tissues. Sequences that have a homolog in *H. indicum* could have

a higher probability to be involved in PA biosynthesis because it is known that in Boraginales, only one independent duplication event of HSS gave rise to the evolution of PA biosynthesis (Reimann et al., 2004). There is a likelihood that the core biosynthesis (step 1 to 6, Fig. 1), the synthesis of the necine base, evolved before speciation of the two lineages of the Boraginaceae and the Heliotropiaceae. Several studies showed that enzymes at the beginning of a pathway are under higher selective constraints than enzymes more downstream (Ramsay et al., 2009; Ma et al., 2010; Olson-Manning et al., 2013). Therefore, enzymes that are important for a central step of PA biosynthesis could be more conserved between species; this is apparently true for the HSS which was found to be the conserved, first pathway specific enzyme in all PA producing plants studied so far (Ober and Hartmann, 1999a; Anke et al., 2004; Reimann et al., 2004; Nurhayati and Ober, 2005; Anke et al., 2008; Nurhayati et al., 2009; Kaltenegger et al., 2013; Irmer et al., 2015). Future studies have to clarify if genes of the core biosynthesis evolved before speciation of the two lineages or have evolved convergently in *S. officinale* and *H. indicum*.

4.4.2 Two putative homospermidine oxidases in different organs?

We identified candidate sequences by searching for homologs of enzyme classes known to have the potential to catalyze similar reactions as those proposed for the PA biosynthesis. Here, we focused on the most intensely investigated part, the formation of the necine base. Tracer studies identified several intermediates following the synthesis of homospermidine, which is incorporated exclusively into the necine base (Robins, 1982; Böttcher et al., 1993; Graser and Hartmann, 1997). It is hypothesized that homospermidine is oxidized in one or two steps to 4,4'-iminobisbutanal to form an immonium ion as postulated intermediate on the path to the trachelanthamidine precursor of PAs (Robins, 1982; Kelly and Robins, 1988; Böttcher et al., 1993; Böttcher et al., 1994). This reaction is believed to be catalyzed by a diamine oxidase. Tracer experiments detected the accumulation of homospermidine, when hairy roots cultures were treated with the diamine oxidase specific inhibitor β-hydroxyethylhydrazine (Böttcher et al., 1993; Frölich et al., 2007). When the inhibitor was removed from the culture media, the accumulated homospermidine was rapidly and exclusively incorporated into the PAs (Böttcher et al., 1993; Frölich et al., 2007). We were able to identify five sequences of diamine oxidases (DAO) from *S. officinale*. Two enzymes, namely SoDAO1 and SoDAO5 showed a strong correlation with HSS expression in PA producing tissues. The other

three of the identified DAOs were not regarded as candidates due to their expression profiles or, in the case of SoDAO2, their incapability of accepting homospermidine as a substrate. Interestingly, neither of the putative homospermidine oxidases was detected in both of the PA producing tissues (leaf PA + or root PA +). SoDAO1 was exclusively found in leaves (PA +) and SoDAO5 exclusively in roots (PA +). This could indicate that different sets of genes have evolved for biosynthesis of PAs in the two different tissues. However, we found no hints for two gene copies of HSS, which suggests that the promoter of at least this gene has the capability to be active in both organs. Both DAO enzymes are predicted to have a signal peptide that directs them to the secretory pathway. Further investigations are needed to detect in which organelle the DAOs are localized and to develop a model for the subcellular localization of the complete PA pathway. HSS was found in the cytosol, but homospermidine was found to accumulate in the cell without degradation. Therefore, the intermediate homospermidine has to be protected from unspecific decay through unspecific DAOs or polyamine oxidases (Frölich et al., 2007). It is possible that homospermidine is translocated from the cytosol to another compartment where it is further processed, for example peroxisomes or peroxisome-like vesicles. Such alkaloid synthesizing vesicles are described for the benzylisoquinoline alkaloid biosynthesis (Amann et al., 1986; Facchini and Bird, 1998; Facchini and St-Pierre, 2005) and for the nicotine biosynthesis (tropane alkaloids) (Naconsie et al., 2014). The localization of PA biosynthesis in peroxisomes or peroxisome-like vesicles could be also beneficial regarding the hydrogen peroxide, which is generated during the oxidation of homospermidine. Hydrogen peroxide could efficiently be converted to harmless oxygen and water by catalases that are present in high concentrations in peroxisomes (Kaur et al., 2009).

4.4.3 Is an enzyme involved in cyclization of the necine base?

After oxidation of homospermidine, the resulting putative product 4,4'-iminobisbutanal is believed to self-cyclisize spontaneously to the unknown trachelantamedine precusor “metabolite X” (Frölich et al., 2007), which is hypothesized to be an immonium ion (Kelly and Robins, 1988; Hartmann and Witte, 1995). Preliminary results from long term incubations of homospermidine with the homospermidine unspecific SoDAO2 and subsequent LC-MS analysis of the reaction products found only traces of 4-(4-aminobutylamino)butanal and a strong signal of 4,4'-

iminobisbutanal (Suppl. Fig. 5). This raised the question if another enzyme is facilitating the cyclisation to the necine base precursor. One enzyme that prominently emerged in our analysis was a *S*-norcoclaurine synthase-like enzyme (step 3, Fig. 1, Tab. 1), which was highly upregulated in PA producing tissues of *H. indicum* and *S. officinale*. Norcoclaurine synthase was identified in the benzylisoquinoline alkaloid biosynthesis as enzyme that condenses dopamine and 4- hydroxyphenyl-acetaldehyde to give (*S*)-norcoclaurine (Luk et al., 2007; Vimolmangkang et al., 2016). This enzyme catalyzes a so called Pictet-Spengler reaction, an electrophilic-cyclization between an iminium ion and an aromatic ring (Pictet and Spengler, 1911; Cox and Cook, 1995; Larghi et al., 2005). A Norcoclaurine synthase-like enzyme might combine the terminal carbonyl groups with the secondary amino group of 4,4'-iminobisbutanal, resulting in an intramolecular cyclization of the dialdahyde. The involvement of an enzyme would explain the conserved stereochemistry of the resulting lycopsamine type PAs (Hartmann and Witte, 1995), but further experiments are needed to test the involvement of the identified sequences.

4.4.4 Candidate sequences for other steps of the PA biosynthesis pathway

Robins (1982) hypothesized that after formation of the ring system, a following reduction of the carbonyl group is catalyzed by an alcohol dehydrogenase (ADH, EC 1.1.1.1, step 4, Fig. 1) and leads to the necine base. In a biomimetic experiment by Robins (1982), a one-pot reaction with homospermidine as substrate, an unspecific DAO (of pea seedlings) and a PA unspecific liver alcohol dehydrogenase (ADH) was sufficient to produce detectable amounts of trachelanthamidine, the necine base backbone. We found sequences coding for ADHs in the enriched subsets (subset 1,2, and 3) of *S. officinale* and *H. indicum* which should be further tested (Tab. 1). Trachelanthamidine could be additionally modified by the introduction of hydroxy groups by cytochrome P450 enzymes and flavonol synthases/flavanone 3-hydroxylases (Step 5 and 6, Fig. 1) of which candidate sequences are present in the enriched datasets (Tab. 1). A further enzyme family with potential to be involved in PA biosynthesis, was an acetohydroxyacid synthase (AHAS EC 2.2.1.6, step 7, Fig. 1) which is believed to mediate the first steps in the necic acid synthesis of lycopsamine-type PAs from branched chain amino acids as shown by Weber et al. (1999). Furthermore, a MATE (multidrug and toxic compound extrusion) efflux family

protein might be responsible for the transport of PAs into the vacuole. This protein family was described to be involved in the nicotine uptake into the vacuole (Morita et al., 2009). Ehmke et al. (1987; 1988) showed that PAs are also transported to and stored in vacuoles as *N*-oxides. The coupling of necic acid und necine base could then be facilitated by O-acyltransferases (Step 9, Fig. 1) from which sequences were found in the analysis of *H. indicum.*

4.4.5 Summary and Outlook

In summary, this work identified highly promising candidate sequences for a homospermidine oxidase, the second step of PA biosynthesis, and further enzymes for other unknown steps of the pathway. The major drawback of proving the involvement of SoDAO1 and 5 in the PA biosynthesis remains the unsuccessful heterologous expression and biochemical characterization. We are currently establishing new eukaryotic expression systems (e.g., *Pichia pastoris*) to avoid difficulties in post-translational modifications that might be necessary for the activity of the enzymes. In parallel to this study, an *in vivo* test system was successfully established that allows to test transgenic hairy roots harboring RNAi constructs of sequences under study. As a proof-of-concept, we used the already characterized HSS as positive control (see chapter 3). Transgenic hairy root mutants harboring RNAi constructs of the putative homospermidine oxidases (DAOs) are currently under investigation.

4.5. EXPERIMENTAL

4.5.1 Plant material

Symphytum officinale (Rühlemann's Kräuter & Duftpflanzen, Horstedt, Germany) was grown in pots with a mixture of TKS2 (Floragard) and lava granulate in a ratio of 3:1 in the Botanical Gardens Kiel (54°20'45.2"N 10°07'05.3"E) from April to September.

4.5.2 RNA extraction and quantitative real time PCR (qRT-PCR)

Plant samples were pulverized in liquid nitrogen with mortar and pestle before total RNA was extracted with Trizol® (Invitrogen, Life Technologies) according to the manufacturer's protocol including optional steps as reported before by Kruse et al. (2017). RNA was dissolved in RNAse-free water and RNA integrity was tested by agarose gel electrophoresis and by 260/280 nm and 260/230 nm ratio measurements using a NanoDrop® ND2000 UV/VIS spectrometer. Reverse transcription was performed with RevertAid® Premium Reverse Transcriptase following Kruse et al. (2017). For each sample one control reaction was set up without reverse transcriptase (no-RT control) to control for genomic DNA contaminants. A PCR based method was used in case of laser dissected samples with low RNA amounts that do not allow no-RT controls. This method made use of a primer combination where both primers bind in different exons separated by an intron. In the case of a contamination of cDNA with genomic DNA, an additional, larger PCR product should be amplified and would have been detected by electrophoresis in a 3 % (w/v) agarose gel (Suppl. Fig. 2). qRT-PCR was performed in a Rotor-Gene® Q System (Qiagen) using GoTaq® qPCR Master Mix (Promega) following the manufacturer's protocol with an annealing temperature of 60°C. Melting curve analyses were used to exclude formation of primer dimers or unspecific PCR products; primer pairs were used as shown in Tab. 2. Each primer pair was tested for specificity against all other DAO sequences using plasmids containing the respective DAO sequence as template in standard PCR reactions (Suppl. Fig. 3). After performing PCR reactions in a BioRad® T100® Thermal Cycler a subsequent electrophoresis in a 3 % (w/v) agarose gel was used to detect formation of unspecific products. The $2^{\Delta ct}$ method was used to calculate transcripts levels (Schmittgen and Livak, 2008). Glyceraldehyde-3-

phosphate-dehydrogenase (*gapdh*) and elongation factor 1α (*ef1α*) served as reference genes to normalize expression levels.

Table 2. Sequences of Primers used for qRT-PCR.

	Primer	Sequence (5'-3')
PX1	CS-So-Act_for2	CAA GGC TAA CAG GGA GAA AAT GAC
PX2	CS-So-Act_rev2	ATC ACC AGA ATC CAG CAC AAT ACC
PX3	LK-RT-GAPDH-F1	AAG GCA GTC GGT AAA GTG CTT
PX4	LK-RT-GAPDH-R2	TCC TTC TCT AGC CTC ACA GTG A
PX5	SoEf1a-F2	CCA CCA CCC CAA AAT ATT CCA AG
PX6	SoEf1a-R2	CAG TCC AGG TTG GTG GAC CT
PX7	Sohss-F4	AGTGCTATGGACAATGAATCAGTGA
PX8	Sohss-R6	CAG CAA AAT CAG CGC CTC CA
PX9	LK-RT-SoDAO1-F1	TAC GCA ATC TGA AGC ACG GATC
PX10	LK-RT-SoDAO1-R1	CCT GGA TAC AAA CGG TAC CCA AT
PX11	AJ-SoDAO2-F1	TGG ATG TAC CGC CAA ATG TAT GT
PX12	AJ-SoDAO2-R1	GCT TGG CTA TCA AAC TGAGTG GAT
PX13	AJ-SoDAO3-F4	CCC TGC AAC CTG GAG AAT CTA G
PX14	AJ-SoDAO3-R4	ACC TGA ATT ATG TGC TTC CCC TG
PX15	LK-RTSoDAO4-F1	CCC ATA ATG CCA ACC GTT TCA T
PX16	LK-RTSoDAO4-R2	TCC ATG CTA GGT GGG ATC CTA A
PX17	Aj-SoDAO5-F1	AGG TTT CGA GCT AAG GCC TTA C
PX18	Aj-SoDAO5-R1	AAG GCC GGT GAG TTA TTA GCT C
PX19	AJ_SoHSS_Gen_F1	GCT CCT ACT TGT ACT GGG CAT A
PX20	AJ_SoHSS_Gen_R1	TGA ACC ATC AGT CAA AGC TGG G

4.5.3 Laser-assisted microdissection

Young white roots from 1-year-old plants were cryo-fixated without prior treatment for extraction of RNA from endodermis cells. Cryo-sectioning, laser assisted-micro-dissection (LCM), RNA extraction and cDNA synthesis were carried out as reported by Sievert et al. (2015).

4.5.4 High throughput sequencing (Illumina®) and bioinformatic analysis

RNA was extracted as described above and 2 µg of total RNA were lyophilized in RNAstable-tubes (Biomatrica, San Diego, CA, USA) using a speedvac vacuum centrifuge. Library preparation and sequencing was performed by BGI (Shenzen, China) on an Illumina® HiSeq® 2000 platform. Trinity 2.2 pipeline was used for calculation of *de novo* assembly and differential expression analysis (Grabherr et al., 2011; Haas et al., 2013). Prior to assembly, all reads were gently quality filtered and trimmed with Trimmomatic (Version 0.36, which is included in the Trinity package) with default parameters (Bolger et al., 2014; Macmanes, 2014). The resulting assembly was filtered with the script “filter_isoform_trinity_RSEM.py“ to remove lowest-covered isoforms in order to reduce redundancy and chimeric content as demonstrated by Yang & Smith, (2013). The R based program RSEM was used for alignment-based quantification of transcripts (Li and Dewey, 2011; R Core Team, 2017). We used Bioconductors EdgeR package to identify differential expressed transcripts (Robinson and Smyth, 2007; Robinson and Smyth, 2008; Robinson et al., 2010; McCarthy et al., 2012; Zhou et al., 2014). For functional annotation and gene ontology term assessment, we used the Blast2Go Basic pipeline (Conesa et al., 2005; Conesa and Götz, 2008). For blast analysis, default parameters with following changes were chosen: -e-value: 1xe-5, -max_target_seqs: 20. We used the UniProtKB/Swiss-Prot (ftp://ftp.ncbi.nlm.nih.gov/blast/db/swissprot.tar.gz; from June 12th 2016) as database.

4.5.5 Identification of candidate sequences

To identify candidate sequences, all sequences were blasted using Blast+ package (version 2.2.25, Camacho et al., 2009) against UniProtKB/Swiss-Prot database (see above) followed by extraction of sequences that showed high expression in PA producing tissues (Kruse et al., 2017). Secondly, sequences encoding enzymes with potential to catalyze postulated steps of PA biosynthesis were extracted and tissue specific expression was re-examined via qRT-PCR. Validation of the results was conducted by a comparison with candidates identified from *Heliotropium indicum.* In the case of DAOs, we first identified all sequences coding for putative DAOs, analyzed them via real time PCR and generated expression constructs for subsequent heterologous expression and biochemical characterization. Phylogenetics of DAO were calculated with the program Geneious® (Neighbor-joining algorithm, Geneious Tree Builder, Geneious® 9.1.4, Biomatters Ltd.) and subcellular localisation was predicted using PSORT (http://www.genscript.com/wolf-psort.html) and TargetP (http://www.cbs.dtu.dk/services/TargetP/) (Emanuelsson et al., 2000; Horton et al., 2007; Petersen et al., 2011). Peroxisomal targeting was assigned manually by the presence of PTS1-like motifs (Alanine-Lysine-Leucine (AKL*) and Serine-Lysine-Leucine (SKL*)) at the *C*-terminus (Kaur et al., 2009) and verified by the program Ppero (http://biocomputer.bio.cuhk.edu.hk/PP/) (Wang et al., 2017).

4.5.6 Cloning of putative DAO encoding sequences

Each DAO sequence was amplified using primers with adapters to add specific restriction sites and introduce protein tags for native purification and to facilitate the cloning of putative DAO encoding sequences into a recombinant expression vector. For each sequence, a slightly adjusted protocol was used. Generally, all sequences were cloned into a modified pET-22b (Novagene) expression vector tagged with either a C- or N-terminal Strep- or His-tag (details see Tab. 3). The multiple cloning site (MCS) of pET-22b was exchanged against the MCS of the pET-21d (Novagene) expression vector giving pET-22b mod NcoI. Each sequence was amplified with its specific primer pair (Tab. 3) using Phusion® HS II Polymerase (Thermo Fisher Scientific) in a reaction mix containing 1x Phusion HF Buffer, dNTPs (200 µM each), forward and reverse primer 0.5 µM each), Phusion HF DNA polymerase (0.02 U/µl), and diluted cDNA prepared form 1 µg total RNA (1:5) as template. Cycling conditions

were chosen as recommended by the manufacturer with primer specific annealing temperature during the first five cycles and 72 °C for the remaining cycles. Primers used are listed in Tab. 3. PCR products were analyzed via electrophoresis in a 1 % (w/v) agarose gel and the fragments of the calculated size were cut from the gel and purified using the NucleoSpin® Gel and PCR Clean-up Kit (Marcherey & Nagel) as stated by the manufacturer. The appropriate restriction enzymes (2,5 µl each, FastDigest, Thermo Fisher Scientific) were used for hydrolization of 1 µg of vector and 500 ng of insert DNA in 100 µl reactions at 37 °C. After an additional gel purification step, insert and vector were ligated with T4 DNA Ligase (Thermo Fisher Scientific) with a vector to insert ratio of 1:3 at 22 °C for 16 hours. The resulting plasmids were transformed into chemically competent *E. coli* TOP10 cells (Invitrogen) for vector propagation. Plasmids were sequenced at MWG Eurofins to confirm identity of expression constructs. Site-specific mutagenesis was conducted following the manual of the QuickChange® Site-Directed Mutagenesis Kit (Stratagene). In the case of the SoDAO5, a codon optimized sequence with adapters containing restriction sites (XbaI and XhoI) and a *C*-terminal Strep-tag for cloning into pET23a/22b was synthesized (GeneScript, Piscataway, USA) and cloned as described. To introduce the His-tag, a site-specific mutagenesis with the primers P19 and P20 was performed to remove the stop codon in front of the His-tag. The His-tag is part of the vector pET23a/22b, but was not active before due to a stop codon after the Strep-tag. The generated construct contains a *C*-terminal Strep-His-tag.

Table 3. Sequences, annealing temperature and restriction sites of primers used for cloning and site directed mutagenesis. Used restriction enzymes, generated overhang and purpose of the primer are given.

	Primer	Restriction site	Annealing	Used for
P1	YC140312_SoDAO1_Primer_f_1	BspMI; producing NcoI-overhang	58 °C	Expression of SoDAO1 with *C*-terminal Strep-tag
P2	YC140312_SoDAO1_Primer_r_1_Strep Tag	BspMI; producing XhoI overhang	58 °C	
P3	YC140310_SoDAO2_Primer_f	Eco31I, producing NcoI overhang	58 °C	Expression of SoDAO2 with *C*-terminal Strep-tag
P4	YC140310_SoDAO2_Primer_r_Strep Tag	Eco31I, producing XhoI overhang	58 °C	
P5	SoDAO3_exp_f1	Eco31I, producing NcoI overhang	66 °C	Expression of SoDAO3 with *C*-terminal Strep-tag
P6	SoDAO3_exp_r1	Eco31I, producing XhoI overhang	66 °C	
P7	SoDAO4_exp_f2	Eco31I, producing NcoI overhang	66 °C	Expression of SoDAO4 with *C*-terminal Strep-tag
P8	SoDAO4_exp_r1	Eco31I, producing XhoI overhang	66 °C	
P9	SoDAO5_opt_e_f1	NdeI	72 °C	Expression of SoDAO5 with *C*-terminal Strep-tag
P10	SoDAO5_opt_e_r1	XhoI	72 °C	
P11	SoDAO1_mut1		55 °C	Site-directed mutgenesis for removing internal Shine-Dalgarno sequence
P12	SoDAO1_mut1r		55 °C	
P13	SoDAO2_mut1		55 °C	Site-directed mutgenesis for removing internal Shine-Dalgarno sequence
P14	SoDAO2_mut1r		55 °C	
P15	SoDAO2_mut2		72 °C	Site-directed mutgenesis for removing internal Shine-Dalgarno sequence
P16	SoDAO2_mut2r		72 °C	
P17	SoDAO3_mut1		55 °C	Site-directed mutgenesis for removing internal Shine-Dalgarno sequence
P18	SoDAO3_mut1r		55 °C	
P19	SoDAO3_mut2		72 °C	Site-directed mutgenesis for removing internal Shine-Dalgarno sequence
P20	SoDAO3_mut2r		72 °C	

P21	SoDAO4_mut1_1	55 °C	Site-directed mutgenesis for removing internal Shine-Dalgarno sequence
P22	SoDAO4_mut1r_1	55 °C	
P23	SoDAO4_mut2	72 °C	Site-directed mutgenesis for removing internal Shine-Dalgarno sequence
P18	SoDAO4_mut2r	72 °C	
P19	DAO5_HISmut_f	72 °C	Site-directed mutgenesis for removal of a stop codon before the His-tag
P20	DAO5_HISmut_r	72 °C	

4.5.7 Heterologous expression and detection of biochemical activity of putative diamine oxidases

Plasmids for heterologous expression of DAOs were transformed into different *E. coli* host strains (Tab. 4). In general, protein expression was induced in mid-log phase growing *E. coli* cells containing the respective expression construct after addition of isopropyl-β-D-thiogalactopyranosid (IPTG). The cultures were incubated for 16 hours at 16 °C in LB media supplemented with 100 µg/ml ampicillin and 1 mM $CuSO_4$ or ethylenediamine tetraacetic acid copper(II) disodium salt (EDTA-Cu) while shaking at 230 rpm. When expressing SoDAO1-4 an empty pETM-22 plasmid was co-transformed and selected with 100 µg/ml kanamycin for co-expression of thioredoxin A (TrxA to enhance dilsulfide-bond formation as described by Heim et al., 2007). SoDAO5 was not co-expressed together with TrxA, because an increased amount of insoluble protein was observed. For protein extraction, the overnight grown cells were pelleted by centrifugation for 10 minutes at 5000 g and re-suspended in lysis buffer, to which lysozyme was added (1 mg/ml) and incubated for 1h at 4 °C. Subsequent sonification (50 %, intensity 3, Branson Sonifier) for 2 minutes was conducted to shear released DNA. For purification of His- oder Strep-tagged proteins, an Äkta FPLC (GE Healthcare Life Sciences) with either a HisTrap HP 1 ml cartridge (GE Healthcare Life Sciences) or a Strep-Tactin Superflow 5 ml cartridge (IBA Lifescience) was used as recommended by the manufacturers. After purification, proteins were concentrated and transferred to 100 mM Tris-HCl (pH 8) containing 100 mM NaCl using centrifugal concentrators (50 kDa Amicon® Ultra Centrifugal Filters, Millipore). Concentrated protein was stored at 4 °C (over night) or at -20 °C (longer storage) until activity measurements.

Diamine oxidase activity was measured by the consumption of the polyamine substrates. Remaining polyamine substrates were quantified by the detection of their benzoyl derivatives and was adapted from Ober et al. (2003). Briefly, putrescine and homospermidine served as the substrates. The standard reaction mixtures with a total volume of 200 µl contained 10 mM potassium phosphate buffer (pH 7,4), 5 U/µl of catalase from bovine liver and 20 µg recombinant protein while varying substrate concentrations. After defined time points, the reaction was stopped by the addition of 500 µl of 2 N sodium hydroxide, the internal standard cadaverine was added (10 µl of 10 mM solution) followed by immediate addition of 10 µl benzoyl chloride. After

incubation for 20 minutes at room temperature the reaction was stopped with 1 ml of a saturated sodium chloride solution. The benzoylated polyamines were then extracted with 2 ml diethyl ether. 1 ml of the ether phase was transferred to a new tube and was evaporated. The residue was dissolved in acetonitrile/ H_3PO_4(1.5 %) (40/60), which also served as the solvent system for HPLC separation. The detection took place at 230 nm and quantification was performed using cadevarine as internal standard. The retention times (R_f; min) for the polyamines are: putrescine, 1.1; cadaverine (internal standard), 1.2; homospermidine, 1.7.

Table 4. Expression constructs and conditions.

Gene	Expression vector	*E. coli* strain	Co-expression with	Culture conditions	Purification method
SoDAO1 SoDAO2 SoDAO3 SoDAO4	pET 22b mod NcoI	Rosetta 2	pETM-22 TrxA	• 16 h at 16 °C with 230 rpm; • 0,2 mM IPTG; • 100 µg ampicillin; • 34 µg/ml chloramphenicol; • 100 µg/ml kanamycin; • 1mM ethylenediamine • tetraacetic acid copper(II) disodium salt (EDTA-Cu) or 1mM $CuSO_4$	*C*-terminal Strep tag
SoDAO5	pET23a/22b	BL21	No co-expression	• 16 h at 16 °C; 230 rpm; • 0,2 mM IPTG; • 100 µg ampicillin; • 1mM Ethylenediamine • tetraacetic acid copper(II) disodium salt (EDTA-Cu) or 1mM $CuSO_4$	*N*-terminal His tag

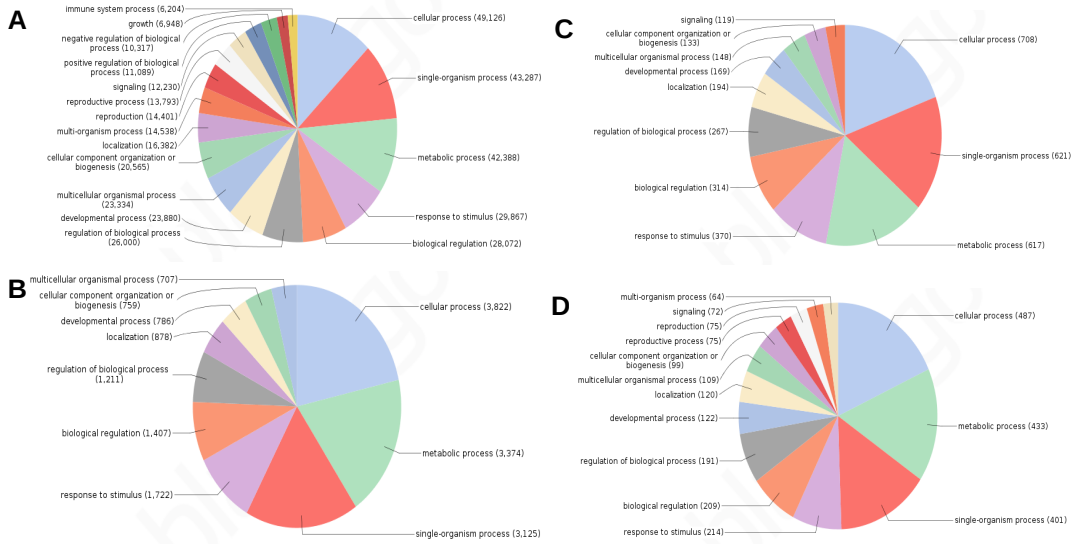
4.6. Acknowledgments

We thank Margret Doose and Brigitte Schemmerling for support in the labratory. We thank PD Dr. Christoph Plieth, Dr. Ulrich Girreser, and Thomas Stegemann for help with the enzyme activity assays and identification of the reaction product. We are grateful to Dr. Jessica Garzke for providing helpful comments on the manuscript.

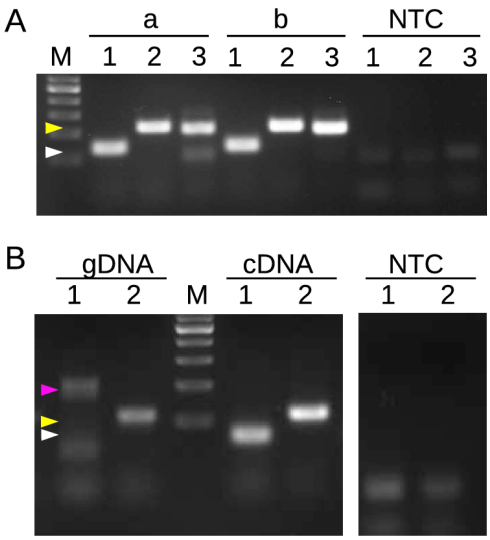
4.7. Supplementary material

Supplementary Table 1. Subcellular localization of diamine oxidases of *S. officinale*. Sequences were analyzed with different programs. vacu, vacuole; mito, mitochondrion; golg, golgi apparatus; chlo, chloroplast; cyto, cytosol; plas, plastid; extr., extracellular; E.R., endoplasmatic reticulum; pero, peroxisome; cysk, cytoskeleton; SP, signal peptide prediction score; cTP, chloroplast target peptide; RC, reliability class (1-5, the lower the better).

Gene	PSORT	TargetP	PTS1 like motif	Position of predicted cleavage site
SoDAO1	vacu: 5, mito: 2, golg: 2, chlo: 1, cyto: 1, plas: 1, extr: 1, E.R.: 1	Secretoric SP: 0.934 RC: 1	no	27
SoDAO2	pero: 13, cyto: 1	Chloroplast cTP: 0.789 RC: 4	Yes, AKL* Ppero-Score: 1.0624701 Perhaps yes	45
SoDAO3	cyto: 9, mito: 2, chlo: 1, pero: 1, cysk: 1	Chloroplast cTP: 0.732 RC: 3	no	24
SoDAO4	vacu: 3, mito: 2, extr: 2, E.R.: 2, golg: 2, chlo: 1, cyto: 1, plas: 1	Secretoric SP: 0.792 RC: 2	no	25
SoDAO5	vacu: 4, mito: 2, extr: 2, golg: 2, chlo: 1, cyto: 1, plas: 1, E.R.: 1	Secretoric SP: 0.710 RC: 2	no	23

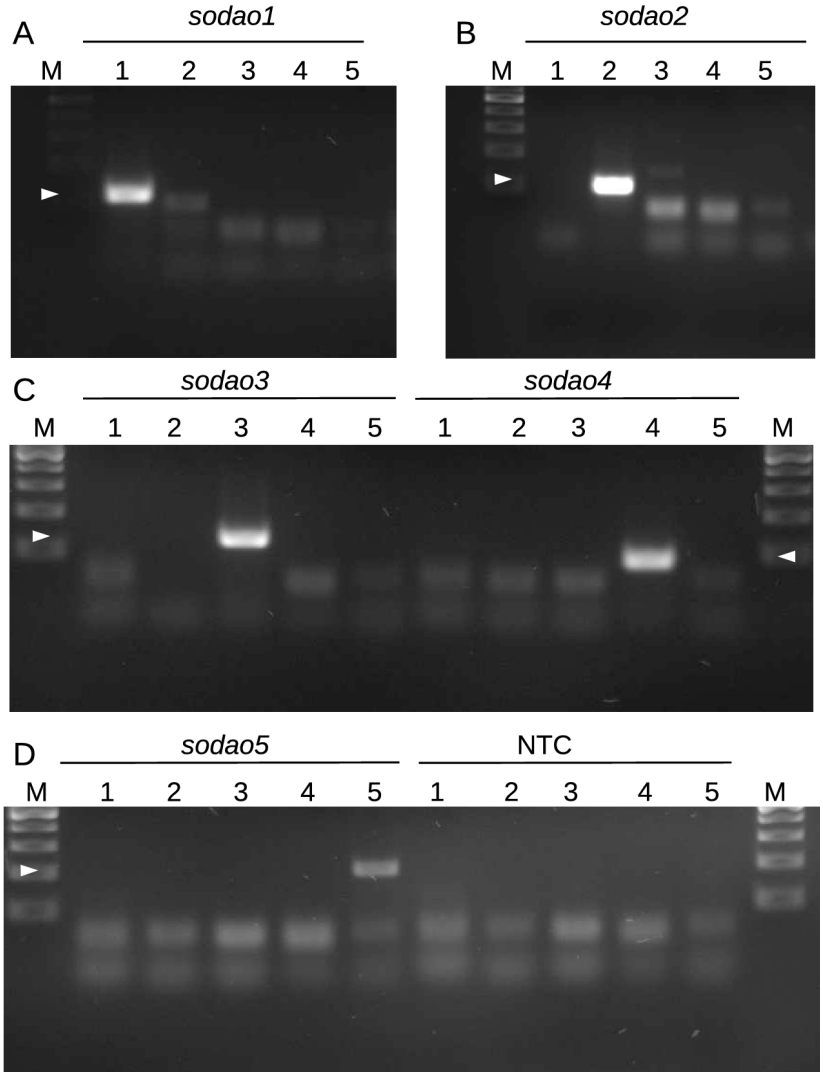


Supplementary figure 1. Gene ontology term assignment to biological processes. GO terms were assigned with Blast2Go for each generated subset. A, the original assembly. B, subset 1 that contains all contigs that are up-regulated either in the PA producing leaf or roots in comparison to the non-PA-producing leaf. C, subset 2 contains contigs that are upregulated in the PA producing leaf and the root in comparison to the non PA producing leaf. D, subset 3 contains contigs that are up-regulated in the PA producing leaf or the root in comparison to the non-PA producing leaf and that have a homolog in *H. indicum* that is upregulated in the lower leaf epidermis.

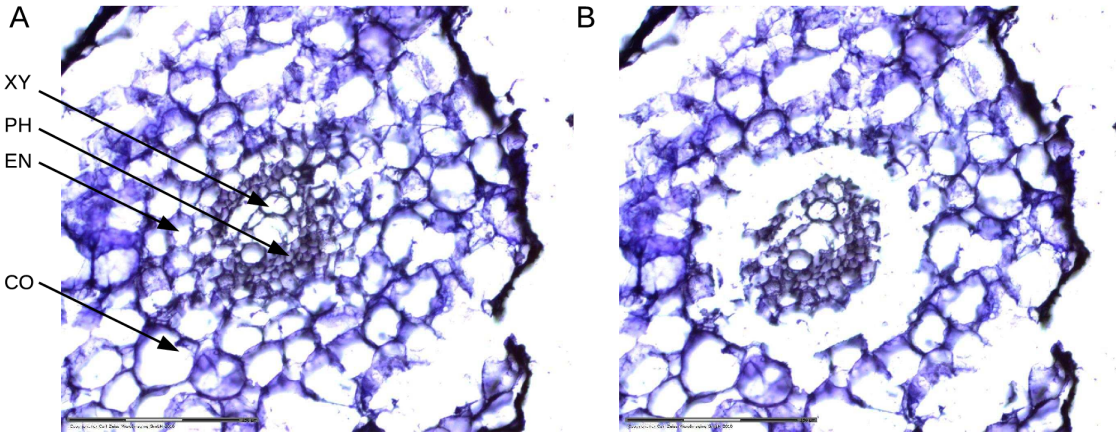


Supplementary figure 2. Test for contamination of cDNA samples with genomic DNA. Test for contamination of cDNA samples with genomic DNA. Primers for *hss* were placed on two exons separated by an intron to distinguish PCR products of genomic DNA and cDNA by size. A, PCR products of *hss* (1), *gapdh* (2) and *sodao5* (3) on a 3 % (m/v) agarose gel from cDNA of endodermis cells (a), the surrounding root cells (b) and a no template control (NTC). The 70 bp PCR product for *hss* indicates that the sample is gDNA free. B, control PCR reactions of the same primers (1, *hss*; 2, *gapdh*) but with genomic DNA as template to test the strategy. As expected when using gDNA as template the PCR product for hss is significantly larger (200 bp) because of remaining introns. Arrow heads indicating different PCR products. Yellow, PCR products of *gapdh* and *sodao5* ; white, products of *hss* without intron; pink, *hss* with intron; M, marker (GeneRuler 100bp Plus DNA Ladder, Thermo Scientific).

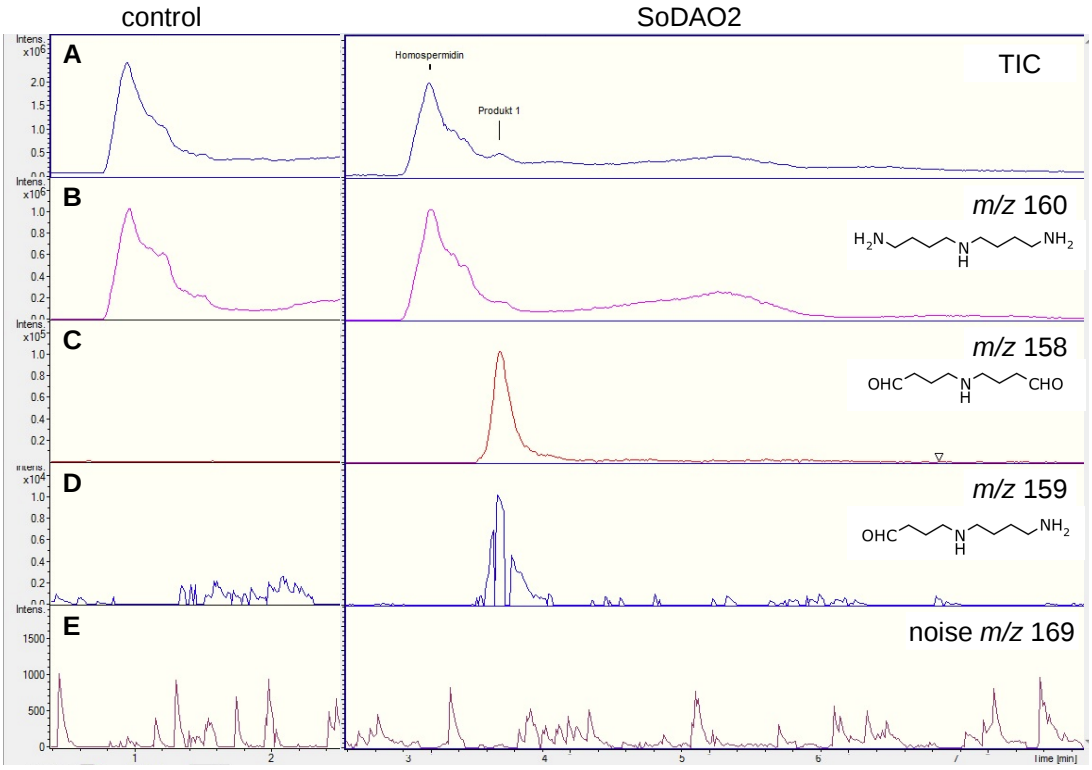
.



Supplementary figure 3. Specificity tests of primers for real-time PCR analyses. For all tested diamine oxidases each primer pair was tested for specificity against all other diamine oxidases. As template plasmid DNA of the corresponding DAO sequence was used. A, Test of primer pairs against *sodao1*. B, Test against *sodao2*. C, Test against *sodao3* and *sodao4*. D, Test against *sodao5* and no template control (NTC). 1, *sodao1*; 2, *sodao2*; 3, *sodao3*; 4, *sodao4*; 5, *sodao5*; White arrow head indicates expected size of PCR product (100bp). M, marker (GeneRuler 100bp Plus DNA Ladder, Thermo Scientific).



Supplementary figure 4. Laser-captured microdissection of *S. officinale* roots. Cryo-cross-sections of *S. officinale* roots before (A) and after (B) microdissection of the endodermis cells. After staining with cresyl violet, endodermis cells were harvested from several cross-sections and collected in the cap of a collection tube. The control "surrounding cells" was harvested accordingly, but from different cross-sections to avoid possible cross-contaminations during the collection process. CO, cortex parenchyma; EN, endodermis; PH, phloem; XY, xylem.



Supplementary figure 5. LC-MS ion $[M+H]^+$ chromatograms of an incubation of SoDAO2 with homospermidine. A, Total ion chromatogram (TIC) of a control incubation without enzyme and with enzyme. B, Extracted ion chromatogram of the corresponding mass *m/z* 160 of homospermidine. C, Extracted ion chromatogram of the corresponding mass *m/z* 159 of 4,4'-iminobisbutanal(dialdehyde). D, Extracted ion chromatogram of the corresponding mass *m/z* 159 of 4-(4-aminobutylamino)butanal (monoaldehyde). E, extracted ion chromatogram of the mass *m/z* 169 (example for noise).

5. CHAPTER 3

In Planta Evidence for the Involvement of Homospermidine Synthase in Pyrrolizidine Alkaloid Biosynthesis – Establishment of a RNAi-mediated knockdown assay in the non-model plant *Symphytum officinale* (Comfrey).

Lars H. Kruse, Julia Jensen-Kroll, Annika Engelhardt, and Dietrich Ober

Botanisches Institut, CAU Kiel, Am Botanischen Garten 1-9, 24118 Kiel

Detailed author contributions are listed at the end of the thesis.

5.1. ABSTRACT

Pyrrolizidine alkaloids (PAs) are a group of toxic plant metabolites with 400 different known structures. So far, only the first pathway specific enzyme, the homospermidine synthase (HSS), is known. HSS produces homospermidine, which is exclusively incorporated into PAs. HSS was recruited several times independently in different plant lineages during evolution by duplication of the gene encoding the primary metabolism enzyme deoxyhypusine synthase (DHS). For a better understanding of the mechanisms involved in the evolution of PA biosynthesis in the angiosperms, it is of great importance to identify more enzymes of the pathway. Previous studies identified several candidate enzymes for further steps of the biosynthesis but resilient evidence is still missing (e.g. diamine oxidases). Here, we describe the establishment of RNAi knockdown hairy root mutants of HSS as a proof-of-concept experiment in *Symphytum officinale*. A knockdown of HSS by 60-80 % resulted in a significant reduction of the major PA components 7-acetylintermedine and 3-acetylechiupinine by 60 %. To our knowledge, this is the first *in vivo* evidence for the involvement of HSS in PA biosynthesis and a strong encouragement for future studies, to test the involvement of further candidate genes by this technique. In addition, this strategy might be also interesting for pharmaceutical applications. *Symphytum* is a medicinal plant, which extracts are traditionally used for the treatment of painful, inflammatory muscle and joint issues. The generation of PA depleted hairy roots could be a cost-efficient way to get rid of these toxic by-products that limit the medicinal applicability of *S. offcinale* extracts.

5.2. INTRODUCTION

Pathways of secondary metabolism are, in contrast to pathways of primary metabolism, not found generally in all plants and are often highly specific to a plant family, a genus, or even to a single species (Pichersky and Gang, 2000; Hartmann, 2007). The capacity to produce secondary metabolites evolved to a great extent under the selection pressure of co-evolving organisms, e.g., floral scent to attract pollinators, toxins to fend off herbivores, or allopathic substances to reduce growth of neighboring plants (Mol et al., 1998; Dudareva and Pichersky, 2000; Pichersky and Gang, 2000; Hartmann, 2007). Involved enzymes, which are responsible for the synthesis of these secondary metabolites are highly specialized and specific for the reaction they catalyze (Pichersky and Gang, 2000). A good example for a quite diverse group of secondary metabolites are the pyrrolizidine alkaloids (PAs). PAs are found in a number of angiosperm lineages, e.g., Asteraceae, Boraginaceae, Convolvulaceae, Fabaceae, and several other plant families including the monocot families Poaceae and Orchidaceae (Hartmann and Witte, 1995; Koulman et al., 2008). PAs are toxic plant secondary metabolites produced as part of the chemical defense against herbivores. Their backbone structure consists of a necine base, which is esterified with one or more necic acids (Hartmann and Ober, 2000). Until today, only the first pathway specific enzyme, homospermidine synthase (HSS) was identified. HSS transfers the aminobutyl moiety of spermidine to putrescine resulting in the formation of homospermidine (Böttcher et al., 1993; Hartmann and Ober, 2000). The PA precursor homospermidine is exclusively incorporated in the necine base backbone of PAs (Böttcher et al., 1993). HSS evolved by gene duplication from a gene encoding the deoxyhypusine synthase (DHS) of primary metabolism, which is involved in the post-translational activation of the eukaryotic initiation factor 5A (Ober and Hartmann, 1999a; Ober and Hartmann, 1999b). HSS evolved several times independently in different plant lineages that are able to produce PAs (Anke et al., 2004; Reimann et al., 2004; Kaltenegger et al., 2013; Irmer et al., 2015). Studies hypothesized that after the formation of homospermidine, *inter alia* diamine oxidases (DAO), alcohol dehydrogenases (ADH) and acetohydroxyacid synthases (AHAS) are involved in the biosynthesis of PAs (Robins, 1982; Kelly and Robins, 1988; Böttcher et al., 1993; Böttcher et al., 1994; Graser and Hartmann, 1997; Weber et al., 1999; Frölich et al., 2007). Studies aiming to prove an involvement of those enzymes in the

PA pathway were unsuccessful to the present day. A popular method that was successfully applied for the identification of a gene function in a specific pathway are gene knockouts by T-DNA insertion. Unfortunately, this method is restricted to well-established model organisms like *Arabidopsis thaliana* (Sønderby et al., 2010; Olson-Manning et al., 2013; Lächler et al., 2015). One method to conduct loss-of-function experiments by gene knockdown in non-model plants is RNA silencing (RNA inteference, RNAi). Since the first report of Napoli et al. (1990) about RNA silencing of a chalcone synthase in the flavonoid biosynthesis in *Petunia* spec., RNA interference became a powerful tool to study gene function (Hamilton and Baulcombe, 1999; Baulcombe, 2004; Sen and Blau, 2006). It was successfully applied to an array of genes in species of model and non-model plants (Kumagai and Kouchi, 2003; Bayindir et al., 2008; DeBoer et al., 2009; Runguphan et al., 2009; DeBoer et al., 2011; Song and Wang, 2011). To facilitate RNAi studies in a plant of interest, a method for stable transformation has to be available. Until now there are only a few universally applicable techniques for stable transformation of non-model plants. The most important method is the infection with *Agrobacterium* (Klee et al., 1987; Pitzschke and Hirt, 2010; Păcurar et al., 2011; Hwang et al., 2015). *Agrobacterium tumefaciens* and *Agrobacterium rhizogenes* are used for the transformation of many different plant species (Chandra, 2012; Ron et al., 2014). Infection of a plant with *A. rhizogenes* induces the so called “hairy root disease” leading to the growth of artificial roots with indefinite growth, strong branching pattern, and an often-enhanced production of secondary metabolites, from the site of infection (Giri and Narasu, 2000; Oksman-Caldentey and Inzé, 2004; Banerjee et al., 2012; Georgiev et al., 2012; Sharma et al., 2013; Ludwig-Müller et al., 2014). Another advantage of the *A. rhizogenes*-mediated plant transformation is the relatively broad host specificity of those bacteria (Porter and Flores, 1991; Veena and Taylor, 2007).
This is the first report of a successful RNAi mediated knockdown study in *S. officinale* after successful transformation with *A. rhizogenes* and the first *in vivo* evidence for the involvement of HSS in the biosynthesis of PAs. Quantitative real-time PCR and LS-MS based PA quantification was assessed in order to identify the consequences of an HSS knockdown in hairy root mutants on PA level in comparison to a control group of hairy root individuals.

5.3. RESULTS

5.3.1 Establishment of RNAi knock down mutants

In order to examine *S. officinale* hairy roots as a test system for candidate sequences of PA biosynthesis, we chose HSS as a target for a proof-of-concept experiment. Reduced HSS transcript levels should result in decreased homospermidine levels and therefore should have an effect on PA levels in the hairy root mutants.

To test if the generated RNAi mutants show a down-regulation of the target gene *hss*, we analyzed different mutants that resulted from independent transformation events on transcript level. Seven well growing mutant lines were cultivated for 14 days and total RNA was extracted. HSS transcripts in these lines were quantified using quantitative real-time PCR (qRT-PCR) with *gapdh* and *ef1α* as reference genes and compared to the mean *hss* transcript levels of seven hairy root wild-type lines. These wild type hairy roots were generated using an vector that contains the *gus* gene instead of a specific RNAi construct. Four HSS knockdown mutants showed a significant down-regulation of HSS transcript level to 23 %-37 % of the transcript level of wild type hairy roots (HR05, HR07, HR09, and HR16; Fig. 1). These four mutant lines were chosen to serve as biological replicates for an RNAi knockdown of HSS in the following analysis. Mutants that showed no significant down- or up-regulation of HSS transcript levels were excluded from the following analyses (HR01, HR06, and HR14; Fig. 1). In order to add more biological significance to our analyses, we compared the mean HSS transcript level of RNAi mutants against the mean HSS transcript level of wild type hairy roots. We found that HSS expression was significantly ($p=0.003$) reduced in the four chosen RNAi mutant lines in comparison to the wild type (Fig. 3A).

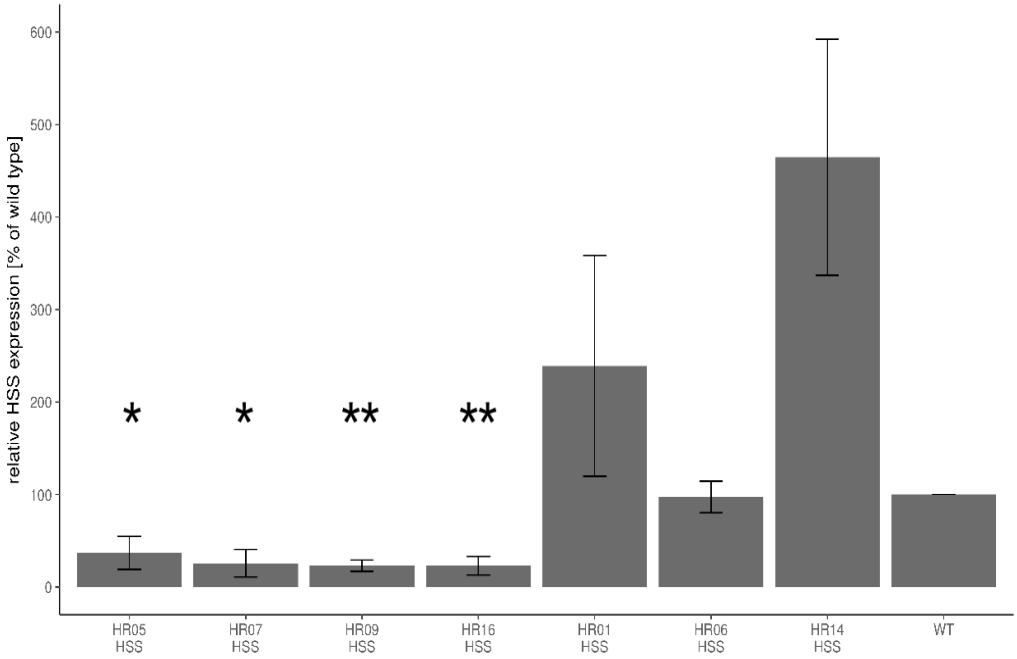


Figure 1. Establishment of HSS RNAi knockdown lines. Seven different hairy root mutants harboring RNAi constructs for knockdown of HSS transcript levels were analyzed in comparison to the mean HSS transcript level of wild type hairy roots (WT). Relative *hss*-transcript levels of each mutant were calculated using *ef1α* and *gapdh* as reference genes. Relative expression is given in percent of wild type. Error bars represent the standard error of the mean (SEM) of three independently grown hairy root cultures of the same line (technical replicates). The mean of 7 independent mutants (biological replicates) that were also grown in triplicate each is given for WT and set to 100 %. Asterisk indicate a significant ($p<0.05$), and two asterisk a high significant ($p<0.01$) down-regulation in comparison to the wild type hairy roots (WT).

5.3.2 HSS is functionally relevant for PA biosynthesis

In order to determine if knockdown of HSS transcript level results in decreased PA levels, we quantified PAs of RNAi mutants and wild type hairy roots with LC-ESI-MS and subsequent broadband MSn fragmentation. Analyzing three masses $[M+H]^+$, we quantified the major components of the PA composition of *S. officinale*: *m/z* 358.18*, m/z* 398.21, *m/z* 440.22 (Fig. 2). By GC-MS analysis, we were able to confirm that the three most abundant PAs that were quantified by LC-MS, are most likely 7-actetylintermedine (*m/z* 358.18), the stereoisomeres echiupinine and myoscorpine (*m/z* 398.21), and 3-acetylechiupinine (*m/z* 440.22). The same stoichiometry of these PAs was observed in both, GC-MS and LC-MS analyses (data not shown). The identified PAs were earlier described as the characteristic PA components of *S. officinale* hairy roots (Frölich et al., 2007).

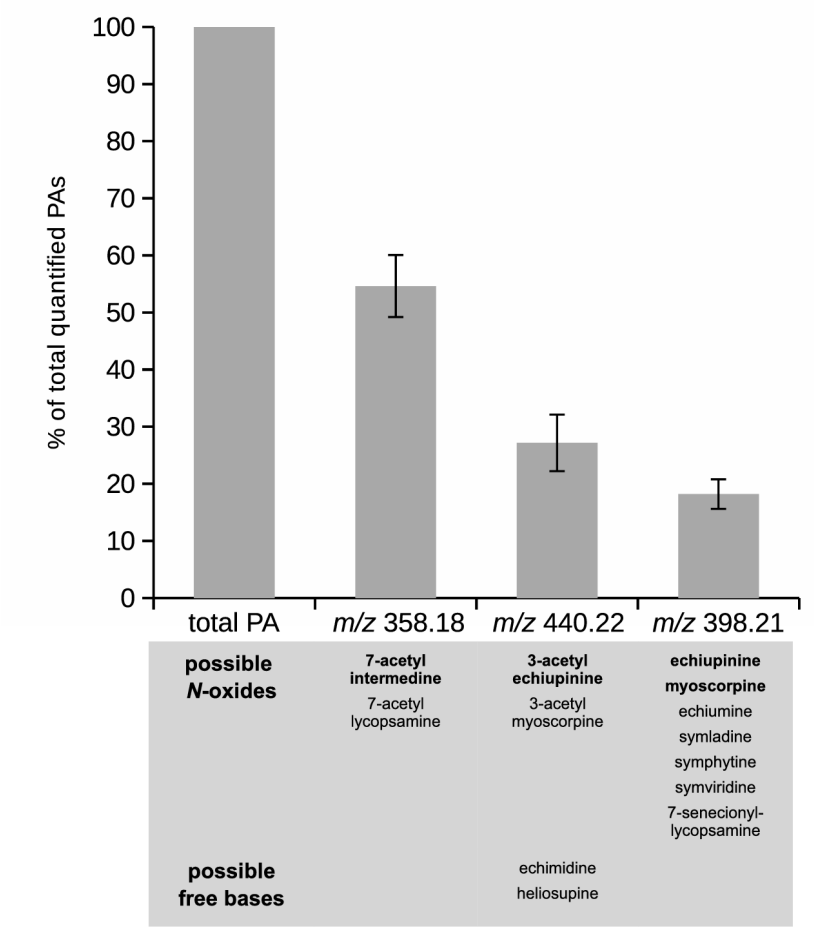


Figure 2. Abundance of the three predominant masses of PAs in wild type hairy roots. Because the chosen LC-MS mode did not allow exact identification of the quantified PAs, each of the three masses can represent different possible PA *N*-oxides or free bases. Therefore we identified the three major components of the PA bouquet with GC-MS. Bold letters indicate the PAs that were detected by GC-MS to be the most abundant in hairy roots. To compare the abundance of the three quantified masses $[M+H]^+$ (*m/z* 358.18, *m/z* 440.22, *m/z* 398.21, we set the sum of the relative intensity of the three quantified masses to 100 % and subsequently determined the proportion of each mass. Error bars give the standard deviation of seven independent wild type hairy root lines. We only quantified the most abundant PAs that account for approximately 95 % of the total PAs (see also Supp. Fig. 1).

A down-regulation of HSS transcript levels should result in reduction of homospermidine levels and subsequently in a reduction of PA content. Therefore, we compared the relative intensities of the three most characteristic masses of PAs in *Symphytum officinale* hairy roots. The relative intensity of the mass *m/z* 358.18 of the PA *N*-oxide 7-acetylintermedine showed a significant reduction in the knockdown mutants compared to the wild type (hedge's g=1.34, p=0.004, Fig. 3B). Also the less abundant PA-*N*-oxide, 3-acetylechiupinine, with the corresponding mass *m/z* 440.22, was significantly decreased in HSS knockdown mutants (hedge's g=2.43, p= 0.015,

Fig. 3C). The relative intensities of both PAs (m/z 358.18 and m/z 440.22) were reduced by approximately 60 % in HSS knockdown lines compared to the wild type hairy roots. This is in the same range of magnitude as the observed down-regulation of HSS transcript levels in the tested knockdown lines by 60-80 % (Fig. 1). When testing the correlation between HSS transcript level and 7-acetylintermedine in knockdown and wild type hairy roots, we found a significant ($p=0.006$) strong positive correlation (Pearson's product-moment coefficient = 0.76, Fig. 4A). Also for the less abundant PA (3-acetylechiupinine), we found a significant ($p=0.014$) strong positive correlation (Pearson's product-moment coefficient = 0.76, Fig. 4B). In addition, we found a significant ($p=0.049$) strong correlation between the third quantified PA specific mass (*m/z* 398.21, echiupinine and myoscorpine) and HSS expression (Pearson's product-moment coefficient = 0.60, Fig. 4C). In contrast to the PAs 7-acetylintermedine (*m/z* 358.18) and 3-acetylechiupinine (m/z 440.22), we found only slight differences in the relative intensities of the mass (*m/z* 398.21, echiupinine and myoscorpine) between HSS knockdown mutants and wild type hairy roots (hedge's $g=0.59$, $p=0.292$, Fig. 3D).

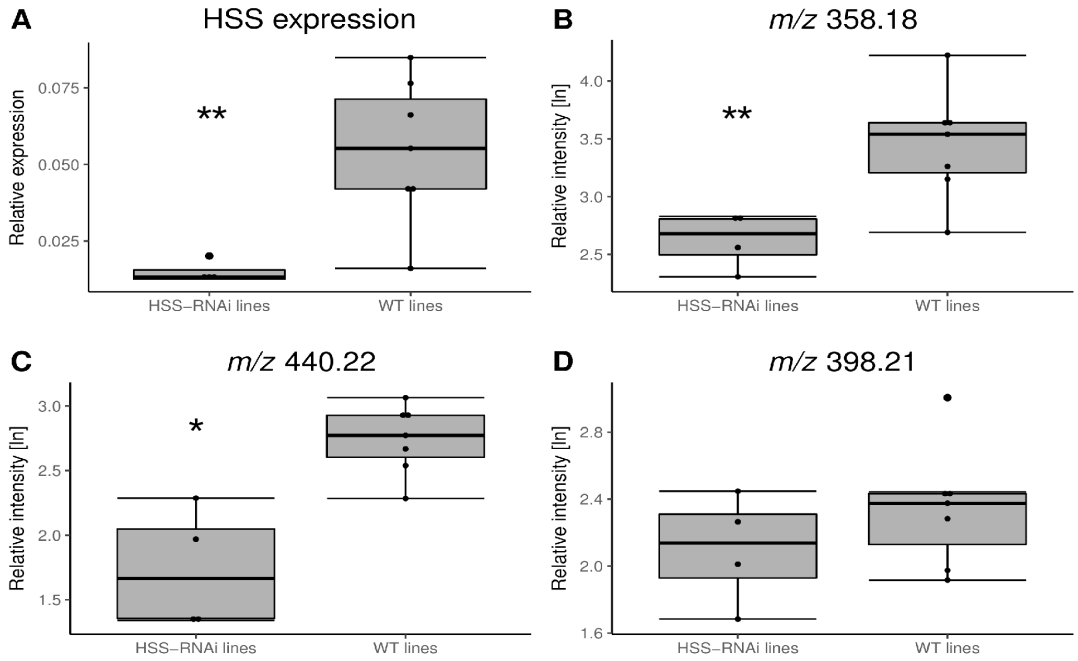


Figure 3. Comparison of HSS transcript level and PA content between HSS knockdown mutants and wild type hairy roots. A, HSS transcript level in percent of the reference genes *ef1α* and *gapdh*. B, relative intensity of the mass $[M+H]^+$ *m/z* 358.18 corresponding to 7-acetylintermedine. C, relative intensity of the mass $[M+H]^+$ *m/z* 440.22 corresponding to 3-acetylechiupinine. D, relative intensities of the mass $[M+H]^+$ *m/z* 398.21 corresponding to myoscorpine and echiupinine. Relative intensities are calculated in relation to the internal standard monocrotaline and extracted biomass. Asterisk indicate significant differences of $p<0.05$ (*) and $p<0.01$ (**).

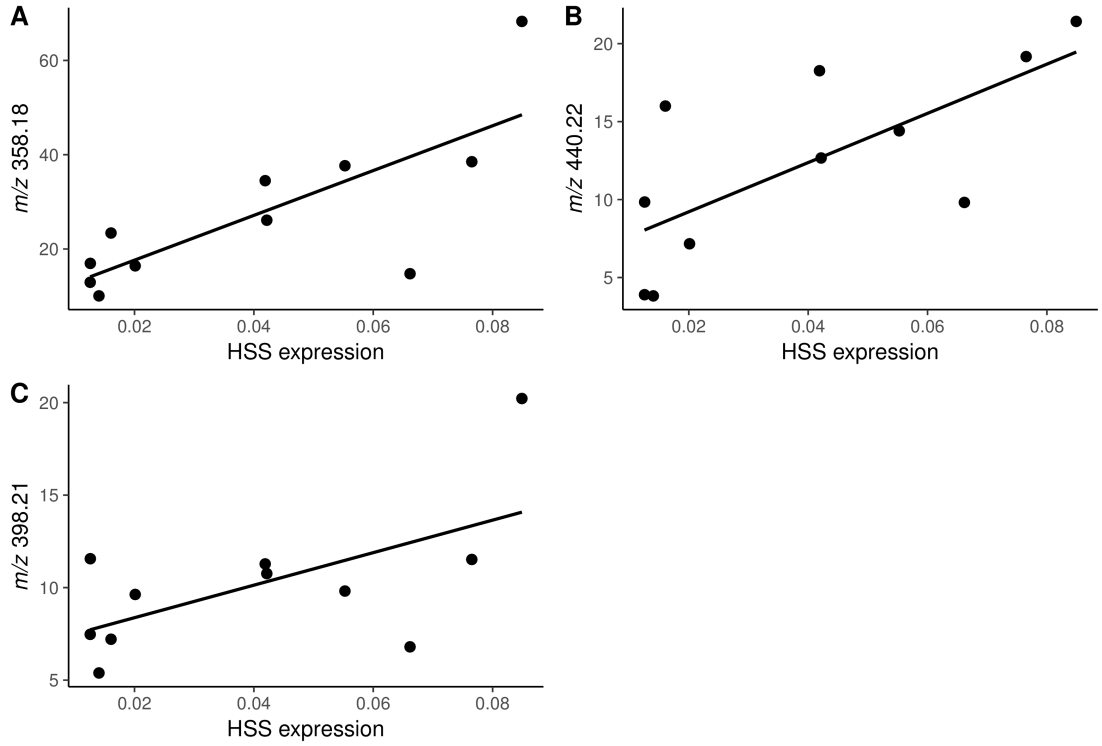


Figure 4. Correlation between HSS transcript level and relative intensity of specific masses of typical pyrrolizidine alkaloids. A, HSS transcript level versus relative intensity of the mass $[M+H]^+$ *m/z* 358.18 corresponding to 7-acetylintermedine ($p<0.01$, Pearson's product-moment coefficient = 0.76). B, HSS transcript level versus relative intesity of mass $[M+H]^+$ *m/z* 440.22 corresponding to 3-acetylechiupinine ($p<0.05$, Pearson's product-moment coefficient = 0.76). C, HSS transcript level versus relative intensities of the mass $[M+H]^+$ *m/z* 398.21 corresponding to myoscorpine and echiupinine ($p<0.05$, Pearson's product-moment coefficient = 0.60). Relative intensities are calculated in relation to the internal standard monocrotaline and extracted biomass.

5.4. DISCUSSION

Over the last 25 years, several *in vitro* studies and tracer feeding experiments were able to show that the HSS is the first pathway specific enzyme of PA biosynthesis and produces homospermidine, which is exclusively incorporated into PA biosynthesis (Böttcher et al., 1993; Böttcher et al., 1994; Ober and Hartmann, 1999a). But resilient *in planta* evidence for the finding that the HSS is necessary for the PA biosynthesis was missing to the present day. Because HSS produces the first pathway specific intermediate homospermidine, it seems to be an ideal target for a gene-knockdown in order to reduce PA content in the plant. In this study we described the first *in planta* evidence for the involvement of HSS in PA biosynthesis by successful RNAi-mediated down-regulation of HSS that resulted in significant reduction of the predominant PAs of *S. officinale* hairy roots.

5.4.1 PA content of hairy roots of *S. officinale*

We observed only slight differences in the PA composition between hairy roots in this study and the earlier published study of Frölich et al. (2007). The three masses that were quantified in our analysis represent 95 % of the PAs that were found in *Symphytum* hairy roots. The *N*-oxide 7-acetylintermedine (*m/z* 358.18) accounted for 47 to 62 % (Fig. 2) of the total PA content quantified in our analysis, which is in accordance with an earlier study by Frölich et al. (2007) who described 7-acetylintermedine-*N*-oxide as the most abundant in hairy roots, contributing approximately 56 % to the total PA content. The *N*-oxide 3-acetylechiupinine (*m/z* 440.22) contributed 19 to 34 % to the total PA content (Fig. 2), which in the range of magnitude of the study by Frölich et al. (2007) who found that 3-acetylechiupinine-*N*-oxide accounts for approximately 10 %. PA-*N*-oxides with the corresponding mass *m/z* 398.21 (myoscorpine and echiupinine), had the lowest abundance and contributed for 15 to 21 % to the total quantified PA content (Fig. 2). These were also described by Frölich et al. (2007) to contribute approximately 20 % of total PA content. That hairy roots differ in the amount of secondary metabolites produced even when generated from the same plant individual is a known phenomenon and is presumably due to the random insertion of T-DNA into the plants genome (Hu and Du, 2006), the altered phytohormone balance (auxin production),

and production of opines that can influence the biosynthetic capacity (Georgiev et al., 2007; Chandra and Chandra, 2011; Chandra, 2012).

5.4.2 RNAi mediated down-regulation of HSS leads to reduced PA levels

The relative intensities of the two predominant masses (*m/z* 358.18 and *m/z* 440.22), contributing together the vast majority of total PAs in hairy roots, are significantly reduced in HSS knockdown mutants. These results clearly demonstrate the importance of HSS for the PA biosynthesis. Reduced transcript levels that lead to reduced protein levels and therefore to a reduction of the first pathway-specific precursor homospermidine, successfully inhibited the production of alkaloids. That the relative intensity of the mass of *m/z* 398.21 is only slightly reduced in HSS knockdown mutants in comparison to wild-type hairy roots could have various reasons. The PA-*N*-oxides that are represented by this mass, namely myoscorpine-*N*-oxide and echiupinine-*N*-oxide, are the precursor for 3-acetyl myoscorpine/ echiupinine-*N*-oxide (*m/z* 440, Frölich et al., 2007). It is possible that when the amount of those precursor PAs is reduced, the acetylation, leading to the 3-acteyl derivatives is inhibited, which could result in a less strong reduction of the PA content of myoscorpine and echiupinine. Myoscorpine-*N*-oxide and echiupinine-*N*-oxide are the PAs with the lowest abundance in our analysis. It might also be that the sample size was not sufficient to detect a significant difference between the knockdown mutants and wild-type hairy roots because of the small effect size (hedge's g = 0.59).
A reduction of PA level by roughly 60 % is in the range of RNAi knockdown experiments in plants synthesizing other alkaloids. DeBoer et al. (2011) found for example a similar decrease in nicotine levels of tobacco hairy roots suppressing ornithine decarboxylase (ODC). Another group found decreased amounts of the tanshinones dihydrotanshinone I by 53 % and cryptotanshinone by 38 % in RNAi knockdown hairy roots of *Salvia miltiorrhiza* supressing copalyldiphosphate synthase (CPS) (Cheng et al., 2014). A few studies also showed that RNAi knockdown of secondary metabolism genes can lead to almost undetectable levels of the respective metabolite (Bayindir et al., 2008; Runguphan et al., 2009).

5.4.3 Perspectives for future studies

The described experimental setup offers different opportunities for future studies: (1) The elucidation of the PA biosynthesis by testing candidate genes for specific steps

of the pathway by gene-knockdown and (2), the opportunity to generate PA-reduced, or even PA-free specimen for the production of herbal medicines.

In vivo knockdown studies have the advantage that a gene function is studied in an almost unchanged environment and the effect of a reduced gene product is tested directly in the plant of interest. Many known and unknown factors influencing the gene *in vivo,* like co-factors and transcription factors, are included in the experiment. In addition to knockdown studies, many secondary pathways have been elucidated by biochemical characterization of recombinantly expressed enzymes, for example enzymes from the tropane alkaloid biosynthesis, lycopodium alkaloid biosynthesis, terpenoid biosynthesis, and flavonoid biosynthesis (Bohlmann et al., 1998; Kaltenbach et al., 1999; Heim et al., 2007; Katoh et al., 2007; Bunsupa et al., 2012; Sun et al., 2012; Naconsie et al., 2014). A major challenge of those *in vitro* studies is the huge diversity of enzymes involved in secondary metabolism from which many are undescribed. Therefore, for each new identified enzyme a new method has to be developed, which is time consuming and bears a great risk to be unsuccessful. RNAi based knockdown studies in hairy roots of *S. offcinale* could be a fast and cost-efficient way to elucidate further steps of PA biosynthesis.

Symphytum officinale is a medicinal plant commonly used for the ectopical treatment of painful, inflamed or swelled muscles and joints, and the major pharmaceutical active compounds are believed to be allantoin and rosmarinic acid (Staiger, 2012). From a pharmacological point of view, *S. officinale* contains PAs as by-product that are toxic and cancerogenous, and restrict the usage of unprocessed herbal medicines from *S. officinale* to ectopical applications (Staiger, 2012). For other forms of applications, laborious purification processes have to be conducted in order to remove PAs efficiently from the prepared extracts (Ludwig-Müller et al., 2014). RNAi mediated knockdown of HSS or other PA pathway-specific target genes might be a way to reduce the content of PAs while keeping the content of allantoin and rosmarinic acid high. It has to be confirmed if the generated HSS knockdown mutants contain sufficiently high amounts of allantoin and rosmarinic acid.

5.4.4 Summary

These results are the first *in planta* evidence for the involvement of HSS in PA biosynthesis and a strong encouragement for further RNAi mediated knockdown studies in hairy roots of *S. officinale*. We believe that this *in planta* method will be a

useful tool to test further candidate genes of PA biosynthesis by knockdown experiments and subsequent phenotyping by PA quantification. Currently, hairy root mutants harboring contructs for RNA silencing of five different diamine oxidases (DAOs) are under investigation. DAOs are hypothesized to facilitate the second step of PA biosynthesis, the oxidation of homospermidine and subsequent formation of the bicyclic ring structure of the necine base (Robins, 1982; Kelly and Robins, 1988; Böttcher et al., 1993; Böttcher et al., 1994). A down-regulation of a PA pathway specific DAO should result in a similar pattern of reduced PA content, as observed in the present study for HSS knockdown mutants, and an accumulation of homospermidine. Candidate genes for further steps of PA biosynthesis might also be taken into consideration for future experiments.

5.5. EXPERIMENTAL

5.5.1 Plant material

Symphytum officinale (Rühlemann's Kräuter & Duftpflanzen, Horstedt, Germany) was grown in pots with a mixture of TKS2 (Floragard) and lava granulate in a ratio of 3:1 in the green-house from April to September in the Botanical Gardens Kiel (54°20'45.2"N 10°07'05.3"E).

5.5.2 Cloning of RNAi constructs

For cloning of RNAi constructs we used the Gateway® cloning system (Invitrogen). For the target gene, a 230 bp long part of the 3' UTR and 3' coding region was amplified with specific primers (Tab. 1) that contained attB1/2-sites for homologous recombination with attP1/2-sites in the pDONR-221 vector (Invitrogen). Gateway® BP Clonase II enzyme mix was used for the integration of PCR products into pDONR221 as recommended by the manufacturer. The resulting plasmids were transformed into chemically competent *Escherichia coli* TOP10 cells (Invitrogen) for vector propagation. For integration into the binary pH7GWIWG2(II) (Karimi et al., 2002) destination vector, the positive pDONR221 entry constructs and the pH7GWIWG2(II) vector were mixed in a molar ratio of 2:1 and homologous recombination was carried out with Gateway® LR Clonase II Enzyme Mix (Karimi et al., 2007). To generate control hairy root mutants that are selectable on antibiotics, in this study called wild type hairy roots, we used a pH7GWIWG2(II) vector that contained the *gus* gene in both multiple cloning sites. The control plasmid containing *gus* was generated using the control plasmid (pENTR™-gus) that was shipped with the Gateway® LR Clonase™ Enzyme Mix using the same cloning strategy as used for the other plasmid. Transformation with this plasmid was considered to not cause any RNAi effect, because the bacterial *gus* gene is not present in *Symphytum officinale*. Plasmids were propagated in *E. coli* TOP 10 cells and correct insertion was controlled by sequencing at MWG Eurofins (Ebersberg, Germany).

5.5.3 Transformation of *Agrobacterium rhizogenes*

Binary vectors were transformed into chemically competent *Agrobacterium rhizogenes* strain ATCC 15834 cells by following procedure: chemically competent cells were thawed on ice, 2 µg of plasmid DNA was added and the mixture was incubated for additional 30 minutes on ice (freeze/thaw method, Wise et al., 2006). Then, the mixture was frozen in liquid nitrogen and thawed in a water bath at 37 °C for approximately 45 s until completely liquified. To allow formation of resistance against spectinomycin, 1 ml of YEB media (5 g/l beef extract, 1 g/l yeast extract, 5 g/l peptone, 5 g/l sucrose, 0.5 g/l $MgSO_4$, pH 7) was added, followed by incubation at 28 °C for 5 hours on a rotary shaker. Cells were streaked out on solidified YEB media (1.5 % agar) containing 100 µg/ml spectinomycin and cultured for 48 hours at 28 °C. Positive clones were grown for additional 16 hours in liquid YEB. Stock cultures were stored in YEB media containing 25 % (v/v) glycerol at -80 °C until use.

5.5.4 Generation of transgenic hairy root mutants

For the generation of hairy roots, young leaves (max. 20 cm length) were harvested and washed with unsterile water supplemented with Tween 20 (0,01 %). Sterilization of the surface was achieved by washing the leaves for 30 seconds in 70 % ethanol and for 5 minutes in calcium hypochlorite (6 % (w/v) solution of a 70 % (w/w) powder). After this treatment the leaves were thoroughly washed three times in sterile water. The sterile leaves were cut into pieces with 2 to 3 cm edge length and co-cultivated with the agrobacteria. Therefore, a freshly grown overnight culture of *A. rhizogenes* strain ATCC 15834, harboring the respective construct, was centrifuged for 10 min at 3500 g, the supernatant was discarded and the cell pellet was re-suspended to a final OD_{600} of 0.2 in MS media with 20 % of the original amount of NH_4NO_3 (Murashige and Skoog, 1962; Sievert et al., 2015). The explants were incubated in the *Agrobacterium* containing media for 3 hours at 26 °C with gentle agitation before transferring the explants upside down to solid MS20 media (1% (w/v) agar). After three days at 23 °C under constant light/dark regime (16 h of 200 µmol m^2 s^{-1}), the explants were transferred to solid MS20 media containing 250 µg/ml of a mixture of ticarcillin and clavulanic acid (mixed in a ratio of 15:1, Duchefa). The infected explants were cultivated in the climate chamber until hairy roots emerged and were transferred to fresh media when growth of agrobacteria or fungi contaminants was observed. Emerging hairy roots of

approximately 1 cm length were removed from the leaf and transferred to fresh MS20 plates containing ticarcillin, clavulanic acid and hygromycin B (25 µg/ml) for the selection of successful transformed hairy root mutants. Hairy roots mutants that survived the selection were tested by PCR on the integration of the respective constructs into the genome by amplification of the *hpt* resistance cassette and the *rolA* gene. The absence of remaining agrobacteria was tested by trying to amplify the *virD* gene. When amplifcation was unsuccessful, the hairy roots were considered agrobacteria-free. Primers used for PCR analyses are listed in Tab. 2.

5.5.5 Hairy root growth and experimental setup

For quantification of PAs and HSS transcripts, hairy root mutants were grown three times independently under the same conditions in a climate chamber to give technical replicates. From a well grown hairy root line, equal amounts of root material were transferred into 70 ml liquid MS20 media (supplemented with 285 µg/ml ticarcillin/clavulanic acid mixture and 8.5 µg/ml nystatin to prevent growth of agrobacteria or contaminating microorganisms) and grown in the dark at 21 °C for 14 days on an orbital shaker at 130 rpm. At the end of the experiment, the hairy roots were dipped dry with tissue paper, weighed and one half of the grown roots were freeze-dried for PA extraction. The other half was flash-frozen in liquid nitrogen and stored until transcript quantification at -80 °C.

5.5.6 Quantification of PAs

The freeze-dried material was pulverized with mortar and pestle. 10 mg material was transferred to 2 ml reaction tubes and further crushed with metal balls in a Mixer Mill MM 400 (Retsch, Haan, Germany). PAs were extracted twice for 30 s at room temperature in 600 µl of 80 % (v/v) methanol containing 0.5 µg/ml of the PA monocrotaline (Sigma-Aldrich, Munich, Germany) as internal standard. After centrifugation for 10 min with 12.000 g at 4 °C, the supernatants were combined and transferred into a new tube. Before LC-MS measurements, 40 µl 0.1 % formic acid were added to 160 µl of the sample to allow precipitation of potential remaining substances that are insoluble under LC conditions prior to injection.

To assess LC-MS measurement stability, quality control (QC) samples were used following Demetrowitsch et al. (2015). From each sample 20 µl of the extract were

combined and treated the same way as the original samples, e.g., same freeze-thaw cycles, and measured repeatedly over whole LC-MS run. All samples were divided into three batches, each containing one replicate of cultivation and each sample was measured three times. The measurements were performed on three consecutive days on a LC-coupled QTOF-MS system. The components of the extracts were separated on a reversed-phase LC column (Nucleodur C18 Gravity column, 100 mm x 2 mm, 1.8 µm, Macherey-Nagel) with an Infinity 1260 UHPLC system (Agilent Technologies). A flow rate of 250 µl/min and a gradient from 0 % solvent B from minute 1 to 90 % at minute 9 (solvent A: water, 0.1 % formic acid; solvent B: acetonitrile, 0.1 % formic acid) was used. The column was washed with 90 % of solvent B for 30 s after which the system was equilibrated again for 5 min and 30 s with starting conditions.

A microTOF-QII mass spectrometer (Bruker Daltonik, Bremen, Germany) with electrospray ionization (ESI) source was used for mass-to-charge measurements. Positive ionization mode with broadband collision-induced dissociation method (bbCID) was chosen. The parameters were: dry gas at 210 °C, flow rate of 6 l/min, a nebulizer pressure of 100 kPa, and an ionization energy of 10 eV for MS and 20 eV for MS/MS.

Relative quantification of PAs was performed with the Compass PathwayScreener 1.0 software (Bruker Daltonik, Bremen). The parameters were: extracted ion chromatogram width of ±5 mDa, mSigma tolerance set to 1000, and the area and intensity threshold set to 100 counts. The minimum peak area was set to 1 % by a sensitivity level of 99 % and a smoothing (Gauss) width of 0.2. All peaks were extracted and manually searched for peaks of the respective masses, which were selected. In the analysis we concentrated on the masses $[M+H]^+$ of the PAs with the highest abundance in *S. officinale* as reported earlier by Frölich et al. (2007), that were *m/z* 358.18 (7-acetylintermedine), *m/z* 398.21 (echiupinine and myoscorpine), and *m/z* 440.22 (3-acetylechiupinine).

5.5.7 Identification of the quantified PAs by GC-MS

For explicit identification of the quantified PAs we extracted PAs from a wild type hairy root line for GC-MS analysis. The PAs were extracted as described by Kempf et al. (2008). GC-MS data were obtained with a Thermo Trace Ultra DSQ system equipped with an Optima-MN1 column (30 m x 0.25 mm i.d, 0.25 µm film thickness, Macherey & Nagel). EI-mass spectra were recorded at 70 eV (ion source

temperature 200 °C). GC conditions were: injector 280 °C, temperature program 80 °C for 5 min, 80 to 300 °C at 15 °C/min, following 5 min at 300 °C, carrier gas helium 1 ml/min, MS transfer line 250 °C. Individual PAs were identified by their Kovats indices (R_i values), molecular ions and MS fragmentation patterns in comparison to our own comprehensive MS data base of PAs. GC-MS data (*m/z* 50-650) for 7-acetylintermedine were: R_i 2216 (on MN1), *m/z* 341 [M^+]; MS spectrum, *m/z* (relative intensity): 180 (100), 120(75), 93(72), 136(29), 181(28); for echiupinine: R_i 2480 , *m/z* 381 [M^+]; for myoscorpine: R_i 2489, *m/z* 381 [M^+], MS spectrum, *m/z*: 220(100), 136(84), 120(67), 83(53), 93(52); for 3-acetylechiupinine: R_i 2554, *m/z* 423 [M^+], MS spectrum, *m/z*: 220(100), 136(67), 120(64), 93(58), 83(47).

5.5.8 RNA extraction and quantitative real time PCR (qRT-PCR)

Hairy root samples were pulverized in liquid nitrogen with mortar and pestle and total RNA was extracted with Trizol® (Life Technologies) according to the manufacturer's protocol including optional steps as reported before by Kruse et al. (2017). RNA was dissolved in RNAse-free water and RNA integrity was tested by agarose gel electrophoresis and by 260/280 nm and 260/230 nm ratio measurements using a NanoDrop® ND2000 UV/VIS spectrometer. Reverse transcription was performed with RevertAid® Premium Reverse Transcriptase following Kruse et al. (2017). For each sample one control reaction was set up without reverse transcriptase (no-RT control) to control for genomic DNA contaminants. qRT-PCR was performed in a Rotor-Gene® Q System (Qiagen) using GoTaq® qPCR Master Mix (Promega) following a protocol described earlier (Kruse et al., 2017). The $2^{\Delta ct}$ method was used to calculate transcripts levels (Schmittgen and Livak, 2008). Glyceraldehyde-3-phosphate-dehydrogenase (*gapdh*) and elongation factor 1α (*ef1α*) served as reference genes to normalize expression levels. Used primers are listed in Tab. 1.

Table 1. Primer sequences. Sequences of primers used for generation of the RNAi constructs, tests for insertion of T-DNA into the mutant's genome and for transcript quantification by qRT-PCR.

gene	primer name	Annealing		sequence (3' to 5')	reference
hss	LkSoHSSrnai-f2	58 °C	RNAi construct	GGG GAC AAG TTT GTA CAA AAA AGC AGG TTT TAT CTT AAG TGA TGA GAA GGG ATG	This study
hss	LKSoHSSrnai-r	58 °C	RNAi construct	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC TAA CTC ATT TTA TTC AAT TTG AGA TAA CA	
gapdh	LK-RT-GAPDH-F1	60 °C	qPCR	AAG GCA GTC GGT AAA GTG CTT	1.
gapdh	LK-RT-GAPDH-R2	60 °C	qPCR	TCC TTC TCT AGC CTC ACA GTG A	
ef1a	SoEf1a-F2	60 °C	qPCR	CCA CCA CCC CAA AAT ATT CCA AG	
ef1a	SoEf1a-R2	60 °C	qPCR	CAG TCC AGG TTG GTG GAC CT	
hss	Sohss-F4	60 °C	qPCR	AGT GCT ATG GAC AAT GAA TCA GTG A	2.
hss	Sohss-R6	60 °C	qPCR	CAG CAA AAT CAG CGC CTC CA	
rolA	DErolAfor	50 °C	Test for insertion of T-DNA	ACG GTG AGT GTG GTT GTA GG	3.
rolA	DErolArev	50 °C	Test for insertion of T-DNA	GCC ACG TGC GTA TTA ATC CC	
hpt	IH53hpt_f	57 °C	Test for insertion of T-DNA	ATG TCG CAA GGA CGT AAG CCC A	4.
hpt	IH54hpt_re	57 °C	Test for insertion of T-DNA	GGA GTC TTT CAG CAT GGA GCA A	
virD	DEvirDfor	53 "C	Test for insertion of T-DNA	ATG TCG CAA GGA CGT AAG CCC A	5.
virD	DEvirDrev	53 "C	Test for insertion of T-DNA	GGA GTC TTT CAG CAT GGA GCA A	

1. Kruse et al., in preparation; 2. Kruse et al., 2017; 3. Alpizar et al., 2008; 4. Hajdas, unpublished, 5. Alpizar et al., 2006

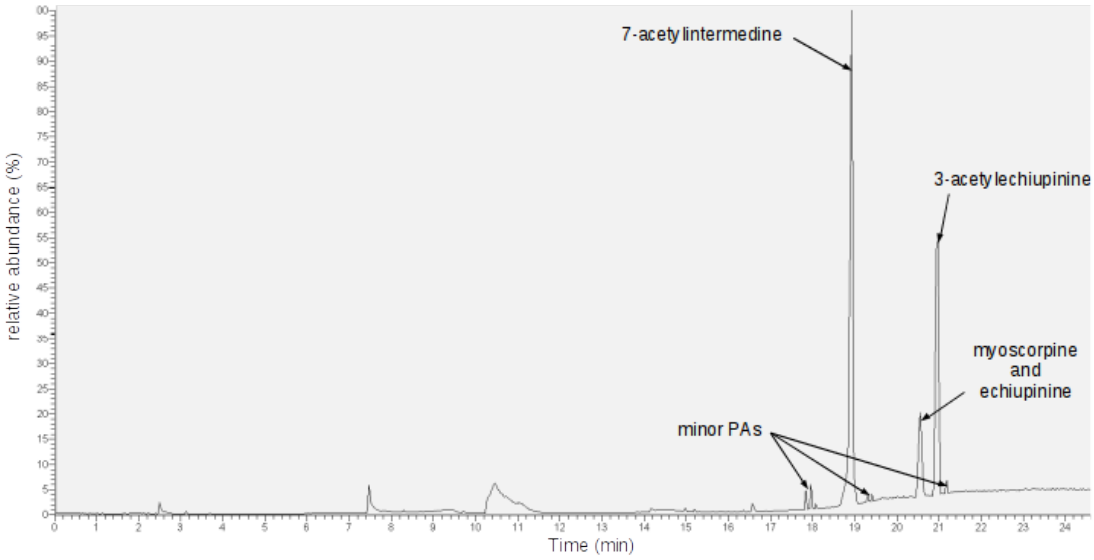
5.5.9 Statistical analyses

Statistical analyses were performed in R version 3.3.3 (R Core Team, 2017). Differences in transcript levels and PA content were firstly tested for normal distribution applying Shapiro-Wilk test, and, if required, data were ln transformed prior to re-fitting models. The Welch two sample t-test was used to compare relative expression between HSS-RNAi and WT lines. Effect sizes, using Hedge's g (Hedges, 1981), were calculated to identify the magnitude of the detected effects. For the interpretation of Hedge's g effect size, it was distinguished between small (<0.2), medium ($0.2 \leq g \leq 0.8$), and large effects (> 0.8). For correlation testing between different PAs and HSS transcript level, Pearson's product correlation was used. R^2 values were interpreted with 0.1 - 0.3 as weak, 0.3 - 0.6 as moderate, and ≥0.6 as strong correlation. Statistical analyses were performed using the packages stats, graphics, and base.

5.6. Acknowledgments

We thank Margret Doose and Brigitte Schemmerling for support in the laboratory. We thank Dr. Tobias Demetrowitsch and Dr. Anne-Maria Wesseling for help with the analysis of LC-MS data. We thank Thomas Stegemann for help with the GC-MS measurements. We are grateful to Dr. Jessica Garzke for sharing her R skills with us and providing helpful comments on the manuscript.

5.7. Supplementary material



Supplementary Figure 1. GC-MS base peak chromatogram (50-300 *m/z*) of a strong cation exchange-solid phase extraction from a wild type hairy root. Peaks of enriched PAs were identified as 7-acetylintermedine (~55 % of total PA content), 3-acetylechiupinine (~31%), a mixture of myoscorpine and echiupinine (~11 %), and other minor PAs (~2%). The stoichiometry of the identified PAs is well in accordance to the stoichiometry of the PAs quantified with LC-MS (Fig. 2).

6. CONCLUSIONS AND GENERAL PERSPECTIVES

The results that are presented in this work add novel knowledge to the field of pyrrolizidine alkaloid biosynthesis in general and particular in comfrey (*Symphytum officinale,* Boraginaceae).

In the past, it was discovered that the PA biosynthesis, and at least the first specific enzyme HSS, is differently regulated and localized in the investigated species. Even between species in which only one common evolutionary origin for PA biosynthesis was found, the site where PAs are produced in the plant varied with regard to involved cells, plant tissues and developmental phase. The research in this thesis reveals that even in one species the regulation of the biosynthesis of PAs is complex and varies between different plant organs, viz., *Symphytum officinale* activates a second site in young leaves, in addition to roots, when inflorescences develop (**Chapter 1**). This enables the plant to adjusts its usually constitutive production of PAs to situations in which additional protection is needed, for example, the growth of young, smooth textured, and nutrient rich inflorescences that could attract herbivores. The observation that the highest concentrations of defense molecules (PAs) are found in reproductive tissues (inflorescences of *Symphytum*) is well in accordance with the optimal defense theory (OD) (McKey, 1979) and also documented for other species (Hartmann et al., 1989; van Dam et al., 1994; Zangerl and Rutledge, 1996). From an ecological perspective, the production of toxins is costly for the plant and subsequently have evolved mechanisms to reduce the costs of secondary metabolism to optimize evolutionary fitness (Gershenzon, 1994; Purrington, 2000). In addition to constitutive defense, another mechanism that is observed in nature is the induced defense (Ohnmeiss and Baldwin, 2000). To produce certain defense

molecules, for example nicotine, only when herbivores feed on the plant might be beneficial for the plant, because the cost (nitrogen in the case of alkaloids,) for the synthesis of such molecules can be reduced when the plant is not attacked (Baldwin, 1988; Ohnmeiss and Baldwin, 2000; Heil and Baldwin, 2002). The greatest disadvantage of induced defense is believed to be the risk to produce effective concentrations in time (which can take hours to days), before the damage of the plant is too high (Heil and Baldwin, 2002; Wittstock and Gershenzon, 2002). It is discussed that the evolution of induced or constitutive defense relies on the frequency a plant is subject to herbivorous damage (McKey, 1979; reviewed in Wittstock and Gershenzon, 2002). A plant that is under constant attack by herbivores therefore relies more on constitutive defense mechanisms, whereas a plant that is rarely attacked uses induced mechanisms to fend off herbivores (reviewed in Wittstock and Gershenzon, 2002). The ecological costs, of one or the other strategy, determine which strategy evolved in the respective species in the ecological niche (cost-benefit) (reviewed in Zangerl and Rutledge, 1996; and Wittstock and Gershenzon, 2002). It seems that *Symphytum* has an advantage to have constant high PA content in the inflorescences, presumably because of persistent grazing pressure. That such a mechanism was selected during evolution implicates that the disadvantage, to allow specialized herbivores to adapt to the high PA concentrations was less strong than the advantage to successfully fend off generalist herbivores. Because plants rely not only on one defense mechanisms, a complete understanding of the evolution that led to the observed defense strategy is likely more complicated (Freeman and Beattie, 2008).

Despite scientists are aware of PAs since the early 1950s and the discovery of homospermidine synthase in 1993, only little progress was made in spotting further enzymes of PA biosynthesis. This thesis presents a strategy to identify candidate genes of PA biosynthesis by next-generation-sequencing based approaches (**Chapter 2**) and subsequent testing by RNAi-mediated gene knockdown (**Chapter 3**). The fact that *Symphytum* produces PAs in specific tissues was used for transcriptome sequencing of PA-producing and non-PA-producing tissues. Candidate sequences for genes of the PA biosynthesis were identified under the assumption that the genes involved in PA biosynthesis are only expressed in biosynthetic active tissues. In particular, genes coding for diamine oxidases (DAOs) were examined in more detail and will serve as a valuable source for identifying the proposed second step of PA biosynthesis, the oxidation of homospermidine. This thesis lays the basis for a number of future research questions, for example, if PA biosynthesis in *Symphytum* is catalyzed by the same or different enzymes in the two PA producing tissues. In **Chapter 3**, an alternative experiment, shows to be a promising strategy to

test the identified candidate genes. By a proof-of-concept experiment it was possible to show that the HSS is essential for PA biosynthesis for the first time *in planta*. Knockdown of HSS in hairy roots of *Symphytum* revealed that reduced levels of HSS resulted in a significant decrease of the predominant PA components. This work is a strong encouragement for such experiments with other candidate enzymes and will promote future attempts in the identification of PA biosynthesis genes significantly. In addition, *Symphytum* is an important medicinal plant and used for a number of pharmaceutical applications. The generation of PA-depleted or PA-free hairy root specimen might be a way to advance the development of pharmaceuticals, containing *Symphytum* extracts, for new therapeutic approaches.

Furthermore, this thesis provides valuable new data for many aspects of future research in the field of PA biosynthesis but might also help to get a better understanding of pathway evolution in general.

7. OUTLOOK

Some general aspects of the perspectives for future research were discussed in general in the previous section. In this section, I will give more precise recommendations for future studies.

7.1. PA biosynthesis in comfrey (*Symphytum officinale*)

In **Chapter 1** it was shown that *Symphytum* activates a second site of PA biosynthesis to boost PA content in inflorescences (Kruse et al., 2017). The analysis showed that the PA content in inflorescences increased during flower development, and significantly decreased in the last analyzed stage. For a complete understanding of the ecological function of PA biosynthesis in *Symphytum,* two questions arise from this observation. First, it would be important to investigate in which parts of the flowers the PAs accumulate, and second, what happens to the PAs when the PA content drops in inflorescences. It has to be clarified if the PAs were metabolized or translocated, and, if translocated, where they were transported to. To know where the flower accumulates PAs could give a hint on the ecological function. In **Chapter 1**, it was hypothesized that higher PA concentrations in petals could fend off nectar robbers. To verify this hypothesis, it should be investigated if petals really accumulate high PA concentrations. According to the optimal defense theory (McKey, 1979), reproductive tissues are of special importance for the plant and need increased

protection. Several studies have shown that especially young tissues show higher concentrations of anti-herbivorous toxins (Zangerl and Rutledge, 1996; Ohnmeiss and Baldwin, 2000; Pavia et al., 2002; McCall and Fordyce, 2010). Following this theory, it is also conceivable that other parts of the flower have a higher value for reproduction than petals, e.g., the ovary, pistil, or stamens. It might be, that high PA concentrations are not important anymore when the inflorescences are withered, because tough and dry tissue is of low interest for herbivores. Therefore, the concentrations of PAs in the different parts of the flower should be subject of further investigations. It could also be beneficial to know which herbivores feed on *Symphytum* inflorescences to get a better understanding of the ecological interactions of *Symphytum* with herbivores during flowering. Special attention should be drawn also on the developing seeds to clarify if and how PAs are removed from the flowers and seeds to protect animals that might be involved in seed dispersal, for example, seed carrying ants.

7.2. Heterologous expression of DAOs and biochemical characterization

Chapter 2 highlights the identification of enzymes putatively catalyzing the second step of PA biosynthesis, presumably diamine oxidases (DAOs). Five different DAO enzymes were identified, and those were subject to intensive transcript level analysis of PA producing tissues. Two DAO coding genes, namely SoDAO1 and SoDAO5, were classified as the most promising candidates. Subsequently, it was tried to heterologously express those enzymes for biochemical characterization. After optimization of *E. coli* expression protocols it was possible to express soluble protein of all identified sequences, but unfortunately only SoDAO2 and 3 from cluster II (Fig. 4., **Chapter 2**) were expressed in active form reproducibly in *E. coli*. This tree cluster contains also sequences from the peroxisomal *N*-methylputrescine oxidase (MPO) from *Nicotiana tabacum* and a peroxisomal localized DAO from *Arabidopsis thaliana.* Both were expressed in active form in *E. coli* as it was observed for SoDAO2 and 3. It might be that other DAOs, clustering with other secretory localized DAOs, need different expression conditions. It was shown that a secretory DAO from *Pisum sativum* was successfully expressed in active form in *Pichia pastoris* (Koyanagi et al., 2000). For the expression of the hitherto inactive DAOs (SoDAO1 and SoDAO5), a switch to the eukaryotic expression system (*Pichia pastoris*) was

already initiated. Vectors for the expression of SoDAO1 and SoDAO5 are prepared and should be used in following experiments, offering different strategies for expression in miscellaneous *Pichia* strains. It was shown, that the expression in yeast and subsequent measurements of enzyme activity without intensive purification in yeast extracts could be sufficient to measure enzyme activity, as it was done to analyze the activity of cytochrome P450 monooxygenases in microsomes (Glawischnig et al., 1999; Hamann and Møller, 2007). Even though, some authors outlined how to reconstitute a complete pathway in yeast (Pyne et al., 2016). Nonetheless, the expression in *E. coli* should be repeated to find out if further variations of the expression and purification conditions could result in active proteins. Preliminary, not reproduceable expression experiments in *E. coli* with SoDAO1 showed trace activities with homospermdine, which should encourage further investigations. It might be beneficial to test an array of buffer conditions that fit to the predicted subcellular localization.

7.3. Localization of HSS and DAOs by fusion with fluorescent proteins

An interesting aspect is that homospermidine accumulates after synthesis when it is not incorporated into PAs (Böttcher et al., 1993; Böttcher et al., 1994; Frölich et al., 2007), therefore, has to be protected from enzymatic decay or excretion. The candidates for a PA-specific DAO are predicted to have a signal peptide for the secretory protein pathway (Suppl. Tab. 1, **Chapter 2**). Other DAOs are either predicted to possess also a secretory signal peptide (SoDAO4), to be localized in the chloroplast (SoDAO3), or the peroxisome (SoDAO2) (Suppl. Tab. 1, **Chapter 2**). The HSS has no detectable signal peptide or sorting signal. It was localized in the cytosol of HSS expressing cells by immunolocalization (Moll et al., 2002; Anke et al., 2004; Niemüller et al., 2012; Kruse et al., 2017). This raises the question where the second step of PA biosynthesis is localized and sets up the need for a homospermidine transport into one of the possible organelles. It is established, that the secretory pathway can guide proteins, for example, to the vacuole (Rojo and Denecke, 2008). Ehmke et al. (1987; 1988) hypothesized a PA specific transporter for the import of PA *N*-oxides into the vacuole which makes it unlikely that PAs are synthesized in vacuoles, but could explain the accumulation of homospermidine when downstream reactions are inhibited or HSS is expressed in a non-PA-producing plant (Böttcher et

al., 1993; Frölich et al., 2007; Abdelhady et al., 2009). Co-localization of HSS with the respective DAOs could be a way to entangle the different possibilities of the subcellular localization of PA biosynthesis. As an easy and widely used method, the generation of fusion constructs with fluorescent proteins (green fluorescent protein, GFP; yellow fluorescent protein, YFP; red fluorescent protein, RFP) (Shaner et al., 2005; Giepmans et al., 2006) could be a convenient way to validate or rule out existing models for the PA pathway localization. As suitable techniques for first experiments, I recommend the ballistic transformation of epidermal cells of onion peels or the PEG-mediated transformation of tobacco protoplasts with appropriate fusion proteins. Subsequent confocal laser scanning microscopy could reveal new insights into the localization of the different proteins. The site of protein localization could also give hints to adjust the heterologous expression protocols to fit the respective properties of the organelle.

7.4. Generation of hairy root DAO knockdown mutants

Chapter 3 describes a successful proof-of-concept experiment proving the involvement of HSS in PA biosynthesis. The suppression of HSS by RNAi-mediated gene silencing resulted in a decrease of PA content in the analyzed mutants. This is a strong encouragement for further experiments using RNAi to downregulate candidate genes of PA biosynthesis. As of the two candidate DAOs (SoDAO1 and 5), only SoDAO5 is expressed in roots, this enzyme should be investigated in more depth. Plant binary vectors containing constructs that result in the formation of double stranded hairpin RNA of SoDAO5 were cloned in the course of this thesis and should be tested in following experiments. Two different constructs were designed that contain in one case a 110 bp long part of the coding region and in another case a 180 bp long part of the 3' untranslated region of SoDAO5. Both constructs should be used for the generation of *Symphytum* hairy root mutants and following growth in highly standardized conditions in liquid media. Subsequently, the homospermidine and PA content should be analyzed in comparison to parallel grown wild type hairy roots in order to determine an effect on PA biosynthesis. In the case that no suitable heterologous expression system can be established for the identified diamine oxidases, this strategy could be a valuable alternative to prove their function as the postulated homospermidine oxidase.

CONTRIBUTIONS

The study described in **Chapter 1** was designed by me and my supervisor Dietrich Ober. I performed the transcript and protein level analysis, performed the experiments (except the radioactive tracer feeding) and the sampling, analyzed the data, and interpreted the results under the supervision of Dietrich Ober. Thomas Stegemann performed the HPLC and GC-MS measurements. Christian Sievert undertook the feeding of the leaves with radioactive tracers that were analyzed by Thomas Stegemann, and provided comments on the manuscript. I prepared all figures and wrote the manuscript together with Dietrich Ober. I have to acknowledge the support of a professional language editing service for this chapter also.

The presented study in **Chapter 2** was designed by me under the supervision of Dietrich Ober. I did all experiments and analyses unlike otherwise stated. Annika Jonathas did the laser assisted-micro-dissection and provided primer specificity tests under my supervision. Under my guidance, Yu-Chen Wu helped to establish a heterologous expression protocol for diamine oxidases in *E. coli*. I analyzed all data, interpreted the results, prepared the figures, and wrote the manuscript. Dietrich Ober provided comments on the manuscript.

The study of **Chapter 3** was designed by me under the supervision of Dietrich Ober. I established transgenic hairy root mutants of *Symphytum officinale*, controlled for insertion of the respective constructs, selected positive mutant lines, and verified LC-MS measurements by GC-MS. Julia Jensen-Kroll and me performed the growth experiment of selected lines. Julia Jensen-Kroll performed the LC-MS measurements, partially supervised by me with respect to optimization of root culturing, harvest and sample preparation. Annika Jonathas performed the qPCR runs under my supervision. I analyzed all data, performed the statistical analyses, interpreted the results, prepared all figures, and wrote the manuscript. Dietrich Ober discussed the results with me and reviewed the manuscript.

DANKSAGUNG & ACKNOWLEDGEMENTS

As the mandatory part of this thesis is completed now, it is the time to thank all the people that helped to let this thesis come true:

Als Erstes möchte ich meinem Doktorvater Prof. Dietrich Ober danken. Dafür, dass Du mir die Möglichkeit gegeben hast, meine Doktorarbeit bei Dir anzufertigen, dass Du mich die ganze Zeit voll unterstützt hast, stets ein offenes Ohr und das Vertrauen in meine Arbeit hattest. Ich danke dir Dietrich!

Des Weiteren möchte ich mich bei Dr. Christian Sievert, PD Dr. Christoph Plieth, Dr. Ullrich Girreser, Prof. Dr. Axel Scheidig, Prof. Dr. Karin Schwarz, Prof. Dr. Jutta Ludwig-Müller für Ihre Hilfe bei verschiedenen Aspekten dieser Thesis danken.

A special thanks goes to all current and former members of the Ober lab for the great time I had doing my PhD. Thank you guys! Especially I want to thank Anne, Moritz, Arun, Elisabeth and Thomas for our great “mensa” times, the long and inspiring discussions and, of course, for our wonderful “Kaffee” group. Ladies and gentlemen, it has been an honour for me!

Besonders möchte ich auch Margret und Brigitte danken! Dir Brigitte dafür, dass Du den Laden hier zusammen hältst, mir bei unzähligen Gelegenheiten geholfen hast und die super Philadelphia-Torte! Dir Margret, möchte ich für die viele Hilfe im Labor, beim Ansetzen von Medien, der Pflege der Pflanzen und unsere gemeinsame Zeit im Büro danken. Außerdem möchte ich dir (und natürlich auch Fritz) danken, dass du dich besonders in der letzten, anstrengenden Zeit des Zusammenschreibens so lieb um Happy mitgekümmert hast.

Natürlich gebührt mein Dank auch meinen Freunden, ohne die mein Studium und die anschließende Promotion bedeutend langweiliger geworden wären. Danken möchte

ich Arne, Daniel, Hannes, Jan und Jan, Irene, Julia, Karo und Matthias für die geniale “Kieler Zeit”! Besonders möchte ich Arne dafür danken, dass Du so ein guter Freund, der weltbeste Mitbewohner und hilfsbereiter Mensch bist. Danke dafür, dass Du die Schreibphase mit mir ausgehalten hast, für das eine oder andere Aufbaubier auf dem Balkon, dass gemeinsame Kicken und Deine große Loyalität!

Meiner Familie möchte ich auch von ganzem Herzen danken! Euch, Mama und Papa, dass Ihr immer für mich da wart und mich in jeder erdenklichen Art und Weise unterstützt habt. Ohne euch hätte ich das alles niemals geschafft. Nils, auch Dir möchte ich danken. Danke, dass Du mir beigestanden hast! Bei meiner Oma Helene möchte ich mich dafür bedanken, dass Du immer unerschüttlich an mich geglaubt hast und die ständige Unterstützung, im Kleinen wie im Großen! Du bist die beste Oma der Welt!

Jessica, dir möchte ich ebenfalls danken. Danke für deine Unterstützung bei Allem, ohne Dich wäre ich nicht da, wo ich bin. Ich hätte nicht halb so viel erlebt und noch weniger Spaß daran gehabt. Danke für Deinen wissenschaftlichen Rat und Deine persönliche Unterstützung. Und danke natürlich dafür, dass du mich jeden Tag so wunderbar geweckt hast, egal ob du in Kanada, Helgoland oder Büsum warst! And of course, I have to thank Happy, the best Post-Dog in town. Thank you for making me happy everyday, for reminding me that a PhD thesis is not everything, and making sure that I got fresh air everyday! Thank you!

And last but not least, I want to thank all the people that might have contributed to any part of this thesis, personally or scientifically, but haven’t been mentioned above. You are not forgotten! Thank you for helping me to become the person I am now and making this thesis possible.

REFERENCES

Abd El-Mawla AMA (2010) Effect of certain elicitors on production of pyrrolizidine alkaloids in hairy root cultures of *Echium rauwolfii*. Pharm - Int J Pharm Sci **65**: 224–226

Abdelhady MIS, Beuerle T, Ober D (2009) Homospermidine in transgenic tobacco results in considerably reduced spermidine levels but is not converted to pyrrolizidine alkaloid precursors. Plant Mol Biol **71**: 145–155

Alcázar R, Altabella T, Marco F, Bortolotti C, Reymond M, Koncz C, Carrasco P, Tiburcio AF (2010) Polyamines: molecules with regulatory functions in plant abiotic stress tolerance. Planta **231**: 1237–1249

Amann M, Wanner G, Zenk MH (1986) Intracellular compartmentation of two enzymes of berberine biosynthesis in plant cell cultures. Planta **167**: 310–320

Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Anke S, Gondé D, Kaltenegger E, Hänsch R, Theuring C, Ober D (2008) Pyrrolizidine alkaloid biosynthesis in *Phalaenopsis* orchids: developmental expression of alkaloid-specific homospermidine synthase in root tips and young flower buds. Plant Physiol **148**: 751–760

Anke S, Niemüller D, Moll S, Hänsch R, Ober D (2004) Polyphyletic origin of pyrrolizidine alkaloids within the Asteraceae. Evidence from differential tissue expression of homospermidine synthase. Plant Physiol **136**: 4037–4047

Baldwin IT (1988) Damage-induced alkaloids in tobacco: Pot-bound plants are not inducible. J Chem Ecol **14**: 1113–1120

Bale NM, Crout DHG (1975) Determination of the relative rates of incorporation of arginine and ornithine into retronecine during pyrrolizidine alkaloid biosynthesis. Phytochemistry **14**: 2617–2622

Banerjee S, Singh S, Rahman LU (2012) Biotransformation studies using hairy root cultures — A review. Biotechnol Adv **30**: 461–468

Baulcombe D (2004) RNA silencing in plants. Nature **431**: 356–363

Bayindir Ü, Alfermann AW, Fuss E (2008) Hinokinin biosynthesis in *Linum corymbulosum* Reichenb. Plant J **55**: 810–820

Becerra JX (2003) Synchronous coadaptation in an ancient case of herbivory. Proc Natl Acad Sci **100**: 12804–12807

Becerra JX (2007) The impact of herbivore–plant coevolution on plant community structure. Proc Natl Acad Sci **104**: 7483–7488

Benderoth M, Textor S, Windsor AJ, Mitchell-Olds T, Gershenzon J, Kroymann J (2006) Positive selection driving diversification in plant secondary metabolism. Proc Natl Acad Sci **103**: 9118–9123

Berenbaum M, Feeny P (1981) Toxicity of angular furanocoumarins to swallowtail butterflies: Escalation in a coevolutionary arms race? Science **212**: 927–929

Berenbaum MR, Zangerl AR (2006) Parsnip webworms and host plants at home and abroad: Trophic complexity in a geographic mosaic. Ecology **87**: 3070–3081

Bevan M, The EU Arabidopsis Genome Project, Bancroft I, Bent E, Love K, Goodman H, Dean C, Bergkamp R, Dirkse W, Van Staveren M, et al (1998) Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*. Nature **391**: 485–488

Bird DA, Franceschi VR, Facchini PJ (2003) A tale of three cell types: Alkaloid biosynthesis Is localized to sieve elements in opium poppy. Plant Cell **15**: 2626–2635

Bohlmann J, Meyer-Gauen G, Croteau R (1998) Plant terpenoid synthases: Molecular biology and phylogenetic analysis. Proc Natl Acad Sci **95**: 4126–4133

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinforma Oxf Engl **30**: 2114–2120

Böttcher F, Adolph R-D, Hartmann T (1993) Homospermidine synthase, the first pathway-specific enzyme in pyrrolizidine alkaloid biosynthesis. Phytochemistry **32**: 679–689

Böttcher F, Ober D, Hartmann T (1994) Biosynthesis of pyrrolizidine alkaloids: putrescine and spermidine are essential substrates of enzymatic homospermidine formation. Can J Chem **72**: 80–85

Bottomley W, Gheissman TA (1964) Pyrrolizidine alkaloids. The biosynthesis of retronecine. Phytochemistry **3**: 357–360

Bunsupa S, Yamazaki M, Saito K (2012) Quinolizidine alkaloid biosynthesis: recent advances and future prospects. Front Plant Sci **3**: 239

Burlat V, Oudin A, Courtois M, Rideau M, St-Pierre B (2004) Co-expression of three MEP pathway genes and geraniol 10-hydroxylase in internal phloem parenchyma of *Catharanthus roseus* implicates multicellular translocation of

intermediates during the biosynthesis of monoterpene indole alkaloids and isoprenoid-derived primary metabolites. Plant J **38**: 131–141

Cai D, Klinman JP (1994) Copper amine oxidase: Heterologous expression, purification, and characterization of an active enzyme in *Saccharomyces cerevisiae*. Biochemistry (Mosc) **33**: 7647–7653

Cai D, Williams NK, Klinman JP (1997) Effect of metal on 2,4,5-trihydroxyphenylalanine (topa) quinone biogenesis in the *Hansenula polymorpha* copper amine oxidase. J Biol Chem **272**: 19277–19281

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. BMC Bioinformatics **10**: 421

Castro S, Silveira P, Navarro L (2008) Consequences of nectar robbing for the fitness of a threatened plant species. Plant Ecol **199**: 201–208

Catalfamo JL, Martin WB, Birecka H (1982) Accumulation of alkaloids and their necines in *Heliotropium curassavicum*, *H. spathulatum* and *H. indicum*. Phytochemistry **21**: 2669–2675

Chandra S (2012) Natural plant genetic engineer *Agrobacterium rhizogenes*: role of T-DNA in plant secondary metabolism. Biotechnol Lett **34**: 407–415

Chandra S, Chandra R (2011) Engineering secondary metabolite production in hairy roots. Phytochem Rev **10**: 371

Chang A, Hartmann T (1998) Solubilization and characterization of a senecionine *N*-oxygenase from *Crotalaria scassellatii* seedlings. Phytochemistry **49**: 1859–1866

Chen T, Mei N, Fu PP (2010) Genotoxicity of pyrrolizidine alkaloids. J Appl Toxicol **30**: 183–196

Cheng Q, Su P, Hu Y, He Y, Gao W, Huang L (2014) RNA interference-mediated repression of SmCPS (copalyldiphosphate synthase) expression in hairy roots of *Salvia miltiorrhiza* causes a decrease of tanshinones and sheds light on the functional role of SmCPS. Biotechnol Lett **36**: 363–369

Chizzola R (1993) Repartition of the pyrrolizidine alkaloids acetylintermedine and acetyllycopsamine within the *Symphytum officinale* plant. Planta Med **59**: A644–A645

Colegate SM, Edgar JA, Knill AM, Lee ST (2005) Solid-phase extraction and HPLC-MS profiling of pyrrolizidine alkaloids and their *N*-oxides: a case study of *Echium plantagineum*. Phytochem Anal **16**: 108–119

Conant GC, Wolfe KH (2008) Turning a hobby into a job: How duplicated genes find new functions. Nat Rev Genet **9**: 938–950

Conesa A, Götz S (2008) Blast2GO: A Comprehensive Suite for Functional Analysis in Plant Genomics. Int J Plant Genomics **2008**: 1–13

Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics **21**: 3674–3676

Cordell GA (2013) Fifty years of alkaloid biosynthesis in Phytochemistry. Phytochemistry **91**: 29–51

Couet CE, Crews C, Hanley AB (1996) Analysis, separation, and bioassay of pyrrolizidine alkaloids from comfrey (*Symphytum officinale*). Nat Toxins **4**: 163–167

Cox ED, Cook JM (1995) The Pictet-Spengler condensation: a new direction for an old reaction. Chem Rev **95**: 1797–1842

van Dam NM, Verpoorte R, van der Meijden E (1994) Extreme differences in pyrrolizidine alkaloid levels between leaves of *Cynoglossum officinale*. Phytochemistry **37**: 1013–1016

van Dam NM, Vuister LWM, Bergshoeff C, de Vos H, van der Meijden E (1995a) The "Raison D'être" of pyrrolizidine alkaloids in *Cynoglossum officinale*: deterrent effects against generalist herbivores. J Chem Ecol **21**: 507–523

van Dam NM, Witte L, Theuring C, Hartmann T (1995b) Distribution, biosynthesis and turnover of pyrrolizidine alkaloids in *Cynoglossum officinale*. Phytochemistry **39**: 287–292

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

DeBoer KD, Dalton HL, Edward FJ, Hamill JD (2011) RNAi-mediated down-regulation of ornithine decarboxylase (ODC) leads to reduced nicotine and increased anatabine levels in transgenic *Nicotiana tabacum* L. Phytochemistry **72**: 344–355

DeBoer KD, Lye JC, Aitken CD, Su AK-K, Hamill JD (2009) The A622 gene in *Nicotiana glauca* (tree tobacco): evidence for a functional role in pyridine alkaloid synthesis. Plant Mol Biol **69**: 299

Demetrowitsch TJ, Petersen B, Keppler JK, Koch A, Schreiber S, Laudes M, Schwarz K (2015) Validation of a two-step quality control approach for a large-scale human urine metabolomic study conducted in seven experimental batches with LC/QTOF-MS. Bioanalysis **7**: 103–112

Diatchenko L, Lukyanov S, Lau Y-FC, Siebert PD (1999) Suppression subtractive hybridization: A versatile method for identifying differentially expressed genes. **303**: 349–380

Dixon RA, Strack D (2003) Phytochemistry meets genome analysis, and beyond..... Phytochemistry **62**: 815–816

Dudareva N, Pichersky E (2000) Biochemical and molecular genetic aspects of floral scents. Plant Physiol **122**: 627–634

Ehmke A, Borstel K von, Hartmann T (1988) Alkaloid *N*-oxides as transport and vacuolar storage compounds of pyrrolizidine alkaloids in *Senecio vulgaris* L. Planta **176**: 83–90

Ehmke A, Borstel KV, Hartmann T (1987) Specific uptake of the *N*-oxides of pyrrolizidine alkaloids by cells, protoplasts and vacuoles from *Senecio* cell cultures. *In* B Marin, ed, Plant Vacuoles. Springer US, New York, pp 301–304

Ehrlich PR, Raven PH (1964) Butterflies and plants: A study in coevolution. Evolution **18**: 586

Elmore BO, Bollinger JA, Dooley DM (2002) Human kidney diamine oxidase: heterologous expression, purification, and characterization. J Biol Inorg Chem **7**: 565–579

El-Shazly A, Wink M (2014) Diversity of pyrrolizidine alkaloids in the Boraginaceae structures, distribution, and biological properties. Diversity **6**: 188–282

Emanuelsson O, Nielsen H, Brunak S, von Heijne G (2000) Predicting subcellular localization of proteins based on their *N*-terminal amino acid sequence. J Mol Biol **300**: 1005–1016

Evans PT, Malmberg RL (1989) Do polyamines have roles in plant development? Annu Rev Plant Physiol Plant Mol Biol **40**: 235–269

Facchini PJ (2001) Alkaloid biosynthesis in plants: Biochemistry, cell biology, molecular regulation, and metabolic engineering applications. Annu Rev Plant Physiol Plant Mol Biol **52**: 29–66

Facchini PJ, Bird DA (1998) Developmental regulation of benzylisoquinoline alkaloid biosynthesis in opium poppy plants and tissue cultures. Vitro Cell Dev Biol - Plant **34**: 69–79

Facchini PJ, St-Pierre B (2005) Synthesis and trafficking of alkaloid biosynthetic enzymes. Curr Opin Plant Biol **8**: 657–666

Fraenkel GS (1959) The raison d'Être of secondary plant substances. Science **129**: 1466–1470

Freeman BC, Beattie GA (2008) An overview of plant defenses against pathogens and herbivores. Plant Health Instr.

Frölich C, Hartmann T, Ober D (2006) Tissue distribution and biosynthesis of 1,2-saturated pyrrolizidine alkaloids in *Phalaenopsis* hybrids (Orchidaceae). Phytochemistry **67**: 1493–1502

Frölich C, Ober D, Hartmann T (2007) Tissue distribution, core biosynthesis and diversification of pyrrolizidine alkaloids of the lycopsamine type in three Boraginaceae species. Phytochemistry **68**: 1026–1037

Fu PP, Xia Q, Lin G, Chou MW (2004) Pyrrolizidine alkaloids - Genotoxicity, metabolism enzymes, metabolic activation, and mechanisms. Drug Metab Rev **36**: 1–55

Geldner N (2013) The endodermis. Annu Rev Plant Biol **64**: 531–558

Georgiev MI, Agostini E, Ludwig-Müller J, Xu J (2012) Genetically transformed roots: from plant disease to biotechnological resource. Trends Biotechnol **30**: 528–537

Georgiev MI, Pavlov AI, Bley T (2007) Hairy root type plant in vitro systems as sources of bioactive substances. Appl Microbiol Biotechnol **74**: 1175

Gershenzon J (1994) Metabolic costs of terpenoid accumulation in higher plants. J Chem Ecol **20**: 1281–1328

Gershenzon J, McCaskill D, Rajaonarivony JIM, Mihaliak C, Karp F, Croteau R (1992) Isolation of secretory cells from plant glandular trichomes and their use in biosynthetic studies of monoterpenes and other gland products. Anal Biochem **200**: 130–138

Giepmans BNG, Adams SR, Ellisman MH, Tsien RY (2006) The fluorescent toolbox for assessing protein location and function. Science **312**: 217–224

Gill SS, Tuteja N (2010) Polyamines and abiotic stress tolerance in plants. Plant Signal Behav **5**: 26–33

Giri A, Narasu ML (2000) Transgenic hairy roots: recent trends and applications. Biotechnol Adv **18**: 1–22

Glawischnig E, Grün S, Frey M, Gierl A (1999) Cytochrome P450 monooxygenases of DIBOA biosynthesis: Specificity and conservation among grasses. Phytochemistry **50**: 925–930

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol **29**: 644–652

Graser G, Hartmann T (1997) Biosynthetic incorporation of the aminobutyl group of spermidine into pyrrolizidine alkaloids. Phytochemistry **45**: 1591–1595

Grotewold E (2005) Plant metabolic diversity: a regulatory perspective. Trends Plant Sci **10**: 57–62

Grue-Sørensen G, Spenser ID (1982) The biosynthesis of retronecine. Can J Chem **60**: 643–662

Haas B (2016) Transcriptome contig Nx and ExN50 stats. https://github.com/trinityrnaseq/trinityrnaseq/wiki/Transcriptome Contig Nx and ExN50 stats

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, et al (2013) *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc **8**: 1494–1512

Hamann T, Møller BL (2007) Improved cloning and expression of cytochrome P450s and cytochrome P450 reductase in yeast. Protein Expr Purif **56**: 121–127

Hamilton AJ, Baulcombe DC (1999) A species of small antisense RNA in posttranscriptional gene silencing in plants. Science **286**: 950–952

Hartmann T (2007) From waste products to ecochemicals: Fifty years research of plant secondary metabolism. Phytochemistry **68**: 2831–2846

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

Hartmann T, Ober D (2000) Biosynthesis and metabolism of pyrrolizidine alkaloids in plants and specialized insect herbivores. *In* FJ Leeper, JC Vederas, eds, Biosynthesis. Springer, Berlin, Heidelberg, pp 207–243

Hartmann T, Ober D (2008) Defense by Pyrrolizidine Alkaloids: Developed by Plants and Recruited by Insects. *In* A Schaller, ed, Induc. Plant Resist. Herbiv. Springer, Netherlands, pp 213–231

Hartmann T, Witte L (1995) Chapter four - Chemistry, biology and chemoecology of the pyrrolizidine alkaloids. *In* SW Pelletier, ed, Alkaloids Chem. Biol. Perspect. Pergamon Press, Oxford, pp 155–233

Hartmann T, Zimmer M (1986) Organ-specific distribution and accumulation of pyrrolizidine alkaloids during the life history of two annual *Senecio* species. J Plant Physiol **122**: 67–80

Hashimoto T, Yamada Y (2003) New genes in alkaloid metabolism and transport. Curr Opin Biotechnol **14**: 163–168

He X, Zhang J (2005) Rapid subfunctionalization accompanied by prolonged and substantial neofunctionalization in duplicate gene evolution. Genetics **169**: 1157–1164

Heckel DG (2014) Insect Detoxification and Sequestration Strategies. *In* C Voelckel, G Jander, eds, Annu. Plant Rev. John Wiley & Sons, Ltd, Chichester, UK, pp 77–114

Hedges LV (1981) Distribution Theory for Glass's Estimator of Effect size and Related Estimators. J Educ Stat **6**: 107–128

Heil M, Baldwin IT (2002) Fitness costs of induced resistance: emerging experimental support for a slippery concept. Trends Plant Sci **7**: 61–67

Heim WG, Sykes KA, Hildreth SB, Sun J, Lu R-H, Jelesko JG (2007) Cloning and characterization of a *Nicotiana tabacum* methylputrescine oxidase transcript. Phytochemistry **68**: 454–463

Horton P, Park K-J, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K (2007) WoLF PSORT: protein localization predictor. Nucleic Acids Res **35**: W585–W587

Hu Z-B, Du M (2006) Hairy root and its application in plant genetic engineering. J Integr Plant Biol **48**: 121–127

Hughes CA, Letcher R, Warren FL (1964) 956. The Senecio alkaloids. Part XVI. The biosynthesis of the “necine” bases from carbon-14 precursors. J Chem Soc Resumed 4974–4978

Hwang H-H, Gelvin SB, Lai E-M (2015) Editorial: “*Agrobacterium* biology and its application to transgenic plant production.” Front Plant Sci **6**: 265

Innan H, Kondrashov F (2010) The evolution of gene duplications: classifying and distinguishing between models. Nat Rev Genet **11**: 97–108

Irmer S, Podzun N, Langel D, Heidemann F, Kaltenegger E, Schemmerling B, Geilfus C-M, Zörb C, Ober D (2015) New aspect of plant–rhizobia interaction: Alkaloid biosynthesis in *Crotalaria* depends on nodulation. Proc Natl Acad Sci **112**: 4164–4169

Irwin RE, Maloof JE (2002) Variation in nectar robbing over time, space, and species. Oecologia **133**: 525–533

Jank B, Rath J (2017) The risk of pyrrolizidine alkaloids in human food and animal feed. Trends Plant Sci **22**: 191–193

Kaltenbach M, Schröder G, Schmelzer E, Lutz V, Schröder J (1999) Flavonoid hydroxylase from *Catharanthus roseus*: cDNA, heterologous expression, enzyme properties and cell-type specific expression in plants. Plant J **19**: 183–193

Kaltenegger E, Eich E, Ober D (2013) Evolution of homospermidine synthase in the Convolvulaceae: A story of gene duplication, gene loss, and periods of various selection pressures. Plant Cell Online **25**: 1213–1227

Kaltenegger E, Ober D (2015) Paralogue interference affects the dynamics after gene duplication. Trends Plant Sci **20**: 814–821

Kanegae T, Kajiya H, Amano Y, Hashimoto T, Yamada Y (1994) Species-dependent expression of the hyoscyamine 6 beta-hydroxylase gene in the pericycle. Plant Physiol **105**: 483–490

Karimi M, Depicker A, Hilson P (2007) Recombinational cloning with plant Gateway vectors. Plant Physiol **145**: 1144–1154

Karimi M, Inzé D, Depicker A (2002) GATEWAY™ vectors for *Agrobacterium*-mediated plant transformation. Trends Plant Sci **7**: 193–195

Katoh A, Shoji T, Hashimoto T (2007) Molecular cloning of *N*-methylputrescine oxidase from tobacco. Plant Cell Physiol **48**: 550–554

Kaur N, Reumann S, Hu J (2009) Peroxisome Biogenesis and Function. Arab Book **7**: e0123

Kelly HA, Robins DJ (1988) Evidence for an immonium ion intermediate in pyrrolizidine alkaloid biosynthesis. J Chem Soc Chem Commun **5**: 329–330

Kempf M, Beuerle T, Bühringer M, Denner M, Trost D, von der Ohe K, Bhavanam VBR, Schreier P (2008) Pyrrolizidine alkaloids in honey: Risk analysis by gas chromatography-mass spectrometry. Mol Nutr Food Res **52**: 1193–1200

Khan HA, Robins DJ (1981) Pyrrolizidine alkaloids: evidence for *N*-(4-aminobutyl)-1, 4-diaminobutane (homospermidine) as an intermediate in retronecine biosynthesis. J Chem Soc Chem Commun **11**: 554–556

Khan HA, Robins DJ (1985) Pyrrolizidine alkaloid biosynthesis. Synthesis of ^{14}C-labelled homospermidines and their incorporation into retronecine. J Chem Soc [Perkin 1] 819–824

Klee H, Horsch R, Rogers S (1987) *Agrobacterium*-mediated plant transformation and its further applications to plant biology. Annu Rev Plant Physiol **38**: 467–486

Koch L (2014) Plant genomics: Insights into duplicate gene fate in plants. Nat Rev Genet **15**: 442–442

Kotelnikova KV, Tretjakova DJ, Nuraliev MS (2011) Shoot structure of *Symphytum officinale* L. (Boraginaceae) in relation to the nature of its axially shifted lateral branches. WULFENIA **18**: 63–79

Koulman A, Seeliger C, Edwards PJB, Fraser K, Simpson W, Johnson L, Cao M, Rasmussen S, Lane GA (2008) *E*/*Z*-Thesinine-*O*-4′-α-rhamnoside, pyrrolizidine conjugates produced by grasses (Poaceae). Phytochemistry **69**: 1927–1932

Koyanagi T, Matsumura K, Kuroda S, Tanizawa K (2000) Molecular Cloning and Heterologous Expression of Pea Seedling Copper Amine Oxidase. Biosci Biotechnol Biochem **64**: 717–722

Kruse LH, Stegemann T, Sievert C, Ober D (2017) Second site of pyrrolizidine alkaloid biosynthesis in Comfrey boosts plant defense in floral stage. Plant Physiol pp.00265.2017

Kumagai H, Kouchi H (2003) Gene silencing by expression of hairpin RNA in *Lotus japonicus* roots and root nodules. Mol Plant Microbe Interact **16**: 663–668

Kumar V, Dooley DM, Freeman HC, Guss JM, Harvey I, McGuirl MA, Wilce MC, Zubak VM (1996) Crystal structure of a eukaryotic (pea seedling) copper-containing amine oxidase at 2.2 å resolution. Structure **4**: 943–955

Lächler K, Imhof J, Reichelt M, Gershenzon J, Binder S (2015) The cytosolic branched-chain aminotransferases of *Arabidopsis thaliana* influence methionine supply, salvage and glucosinolate metabolism. Plant Mol Biol **88**: 119–131

Lange BM, Wildung MR, Stauber EJ, Sanchez C, Pouchnik D, Croteau R (2000) Probing essential oil biosynthesis and secretion by functional evaluation of expressed sequence tags from mint glandular trichomes. Proc Natl Acad Sci **97**: 2934–2939

Langel D, Ober D (2011) Evolutionary recruitment of a flavin-dependent monooxygenase for stabilization of sequestered pyrrolizidine alkaloids in arctiids. Phytochemistry **72**: 1576–1584

Langel D, Ober D, Pelser PB (2011) The evolution of pyrrolizidine alkaloid biosynthesis and diversity in the Senecioneae. Phytochem Rev **10**: 3–74

Larghi EL, Amongero M, Bracca ABJ, Kaufman TS (2005) The intermolecular Pictet-Spengler condensation with chiral carbonyl derivatives in the stereoselective syntheses of optically-active isoquinoline and indole alkaloids. ARKIVOC. doi: 10.3998/ark.5550190.0006.c09

Leegood RC (2008) Roles of the bundle sheath cells in leaves of C3 plants. J Exp Bot **59**: 1663–1673

Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics **12**: 323

Li N, Xia Q, Ruan J, P. Fu P, Lin G (2011) Hepatotoxicity and tumorigenicity induced by metabolic activation of pyrrolizidine alkaloids in herbs. Curr Drug Metab **12**: 823–834

Lindigkeit R, Biller A, Buch M, Schiebel H-M, Boppré M, Hartmann T (1997) The two faces of pyrrolizidine alkaloids: the role of the tertiary amine and its *N*-oxide in chemical defense of insects with acquired plant alkaloids. Eur J Biochem **245**: 626–636

Ludwig-Müller J, Jahn L, Lippert A, Püschel J, Walter A (2014) Improvement of hairy root cultures and plants by changing biosynthetic pathways leading to pharmaceutical metabolites: Strategies and applications. Biotechnol Adv **32**: 1168–1179

Luk LYP, Bunn S, Liscombe DK, Facchini PJ, Tanner ME (2007) Mechanistic studies on norcoclaurine synthase of benzylisoquinoline alkaloid biosynthesis: An enzymatic Pictet−Spengler reaction. Biochemistry (Mosc) **46**: 10153–10161

Ma X, Wang Z, Zhang X (2010) Evolution of dopamine-related systems: Biosynthesis, degradation and receptors. J Mol Evol **71**: 374–384

Macel M (2011) Attract and deter: a dual role for pyrrolizidine alkaloids in plant–insect interactions. Phytochem Rev **10**: 75–82

Macmanes MD (2014) On the optimal trimming of high-throughput mRNA sequence data. Front Genet **5**: 13

Mattocks AR (1968) Toxicity of pyrrolizidine alkaloids. Nature **217**: 723–728

Mayer F, Lüthy J (1993) Heliotrope poisoning in Tadjikistan. The Lancet **342**: 246–247

McCall AC, Fordyce JA (2010) Can optimal defence theory be used to predict the distribution of plant chemical defences? J Ecol **98**: 985–992

McCarthy DJ, Chen Y, Smyth GK (2012) Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res **40**: 4288–4297

McGrath AP, Caradoc-Davies T, Collyer CA, Guss JM (2010) Correlation of active site metal content in human diamine oxidase with trihydroxyphenylalanine quinone cofactor biogenesis. Biochemistry (Mosc) **49**: 8316–8324

McKey D (1979) The distribution of secondary compounds within plants. *In* GA Rosenthal, DH Janzen, eds, Herbiv. Their Interact. Second. Plant Metab. Academic Press, New York, pp 55–133

Mithöfer A, Boland W (2012) Plant defense against herbivores: Chemical aspects. Annu Rev Plant Biol **63**: 431–450

Mol J, Grotewold E, Koes R (1998) How genes paint flowers and seeds. Trends Plant Sci **3**: 212–217

Moll S, Anke S, Kahmann U, Hänsch 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 Physiol **130**: 47–57

Morita M, Shitan N, Sawada K, Montagu MCEV, Inzé D, Rischer H, Goossens A, Oksman-Caldentey K-M, Moriyama Y, Yazaki K (2009) Vacuolar transport of nicotine is mediated by a multidrug and toxic compound extrusion (MATE) transporter in *Nicotiana tabacum*. Proc Natl Acad Sci **106**: 2447–2452

Muetterlein R, Arnold C-G (1993) Investigations concerning the content and the pattern of pyrrolizidine alkaloids in *Symphytum officinale* L. (Comfrey). Pharm Ztg **138**: 119–125

Murashige T, Skoog F (1962) A revised medium for rapid growth and bio assays with tobacco tissue cultures. Physiol Plant **15**: 473–497

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

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

Naconsie M, Kato K, Shoji T, Hashimoto T (2014) Molecular evolution of *N*-methylputrescine oxidase in tobacco. Plant Cell Physiol **55**: 436–444

Nakajima K, Oshita Y, Kaya M, Yamada Y, Hashimoto T (1999) Structures and expression patterns of two tropinone reductase genes from *Hyoscyamus niger*. Biosci Biotechnol Biochem **63**: 1756–1764

Napoli C, Lemieux C, Jorgensen R (1990) Introduction of a chimeric chalcone synthase gene into *Petunia* results in reversible co-suppression of homologous genes in trans. Plant Cell **2**: 279–289

Neumann D (1985) Storage of alkaloids. *In* K Mothes, H Schütte, M Luckner, eds, Biochem. Alkaloids. VCH, Weilheim, Germany, pp 49–55

Niemüller D, Reimann A, Ober D (2012) Distinct cell-specific expression of homospermidine synthase involved in pyrrolizidine alkaloid biosynthesis in three species of the Boraginales. Plant Physiol **159**: 920–929

Nowacki E, Byerrum RU (1962) A study on the biosynthesis of the *Crotalaria* alkaloids. Life Sci **1**: 157–161

Nurhayati N, Gondé D, Ober D (2009) Evolution of pyrrolizidine alkaloids in *Phalaenopsis* orchids and other monocotyledons: Identification of deoxyhypusine synthase, homospermidine synthase and related pseudogenes. Phytochemistry **70**: 508–516

Nurhayati N, Ober D (2005) Recruitment of alkaloid-specific homospermidine synthase (HSS) from ubiquitous deoxyhypusine synthase: Does *Crotalaria* possess a functional HSS that still has DHS activity? Phytochemistry **66**: 1346–1357

Ober D (2005) Seeing double: gene duplication and diversification in plant secondary metabolism. Trends Plant Sci **10**: 444–449

Ober D, Gibas L, Witte L, Hartmann T (2003) Evidence for general occurrence of homospermidine in plants and its supposed origin as by-product of deoxyhypusine synthase. Phytochemistry **62**: 339–344

Ober D, Hartmann T (1999a) Homospermidine synthase, the first pathway-specific enzyme of pyrrolizidine alkaloid biosynthesis, evolved from deoxyhypusine synthase. Proc Natl Acad Sci **96**: 14777–14782

Ober D, Hartmann T (1999b) Deoxyhypusine Synthase from Tobacco. cDNA isolation, characterization, and bacterial expression of an enzyme with extended substrate specificity. J Biol Chem **274**: 32040–32047

Ohnmeiss TE, Baldwin IT (2000) Optimal defense theory predicts the ontogeny of an induced nicotine defense. Ecology **81**: 1765–1783

Ohno DS (1970) Duplication for the sake of producing more of the same. Evol Gene Duplic. Springer Berlin Heidelberg, pp 59–65

Ohta T (1988) Evolution by gene duplication and compensatory advantageous mutations. Genetics **120**: 841–847

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

Olson-Manning CF, Lee C-R, Rausher MD, Mitchell-Olds T (2013) Evolution of flux control in the glucosinolate pathway in *Arabidopsis thaliana*. Mol Biol Evol **30**: 14–23

Păcurar DI, Thordal-Christensen H, Păcurar ML, Pamfil D, Botez C, Bellini C (2011) *Agrobacterium tumefaciens*: From crown gall tumors to genetic transformation. Physiol Mol Plant Pathol **76**: 76–81

Panchy N, Lehti-Shiu M, Shiu S-H (2016) Evolution of gene duplication in plants. Plant Physiol **171**: 2294–2316

Papp B, Pál C, Hurst LD (2003) Dosage sensitivity and the evolution of gene families in yeast. Nature **424**: 194–197

Park MH, Nishimura K, Zanelli CF, Valentini SR (2010) Functional significance of eIF5A and its hypusine modification in eukaryotes. Amino Acids **38**: 491–500

Parret AH, Besir H, Meijers R (2016) Critical reflections on synthetic gene design for recombinant protein expression. Curr Opin Struct Biol **38**: 155–162

Pavia H, Toth GB, Åberg P (2002) Optimal defense theory: Elasticity analysis as a tool to predict intraplant variation in defenses. Ecology **83**: 891–897

Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods **8**: 785–786

Pichersky E, Gang DR (2000) Genetics and biochemistry of secondary metabolites in plants: an evolutionary perspective. Trends Plant Sci **5**: 439–445

Pichersky E, Lewinsohn E (2011) Convergent evolution in plant specialized metabolism. Annu Rev Plant Biol **62**: 549–566

Pichersky E, Noel JP, Dudareva N (2006) Biosynthesis of plant volatiles: Nature's diversity and ingenuity. Science **311**: 808–811

Pictet A, Spengler T (1911) Über die Bildung von Isochinolin-Derivaten durch Einwirkung von Methylal auf Phenyläthylamin, Phenylalanin und Tyrosin. Berichte Dtsch Chem Ges **44**: 2030–2036

Pitzschke A, Hirt H (2010) New insights into an old story: *Agrobacterium*-induced tumour formation in plants by plant transformation. EMBO J **29**: 1021–1032

Porter JR, Flores H (1991) Host range and implications of plant infection by *Agrobacterium rhizogenes*. Crit Rev Plant Sci **10**: 387–421

Purrington CB (2000) Costs of resistance. Curr Opin Plant Biol **3**: 305–308

Pyne ME, Narcross L, Fossati E, Bourgeois L, Burton E, Gold ND, Martin VJJ (2016) Chapter Nine - Reconstituting Plant Secondary Metabolism in Saccharomyces cerevisiae for Production of High-Value Benzylisoquinoline Alkaloids. *In* SE O'Connor, ed, Methods Enzymol. Academic Press, pp 195–224

R Core Team (2017) R: A Language and Environment for Statistical Computing, https://www.R-project.org/.

Ramsay H, Rieseberg LH, Ritland K (2009) The correlation of evolutionary rate with pathway position in plant terpenoid biosynthesis. Mol Biol Evol **26**: 1045–1053

Reimann A, Nurhayati N, Backenköhler A, Ober D (2004) Repeated evolution of the pyrrolizidine alkaloid–mediated defense system in separate angiosperm lineages. Plant Cell Online **16**: 2772–2784

Rensing SA (2014) Gene duplication as a driver of plant morphogenetic evolution. Curr Opin Plant Biol **17**: 43–48

Robins DJ (1989) Biosynthesis of pyrrolizidine alkaloids. Chem Soc Rev **18**: 375–408

Robins DJ (1982) A biogenetically patterned synthesis of the pyrrolizidine alkaloid trachelanthamidine. J Chem Soc Chem Commun 1289–1290

Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics **26**: 139–140

Robinson MD, Oshlack A (2010) A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol **11**: R25

Robinson MD, Smyth GK (2007) Moderated statistical tests for assessing differences in tag abundance. Bioinformatics **23**: 2881–2887

Robinson MD, Smyth GK (2008) Small-sample estimation of negative binomial dispersion, with applications to SAGE data. Biostatistics **9**: 321–332

Roeder E (2000) Medicinal plants in China containing pyrrolizidine alkaloids. Pharm **55**: 711–726

Roeder E (1995) Medicinal plants in Europe containing pyrrolizidine alkaloids. Pharmazie **50**: 83–98

Rojo E, Denecke J (2008) What Is moving in the secretory pathway of plants? Plant Physiol **147**: 1493–1503

Romeo JT, Saunders JA, Barbosa P, eds (1996) Phytochemical diversity and redundancy in ecological interactions. Doi: 10.1007/978-1-4899-1754-6

Ron M, Kajala K, Pauluzzi G, Wang D, Reynoso MA, Zumstein K, Garcha J, Winte S, Masson H, Inagaki S, et al (2014) Hairy root transformation using *Agrobacterium rhizogenes* as a tool for exploring cell type-specific gene

expression and function using tomato as a model. Plant Physiol **166**: 455–469

Rosini E, Nossa S, Valentino M, D'Arrigo P, Marinesco S, Pollegioni L (2012) Expression of rat diamine oxidase in *Escherichia coli*. J Mol Catal B Enzym **82**: 115–120

Runguphan W, Maresh JJ, O'Connor SE (2009) Silencing of tryptamine biosynthesis for production of nonnatural alkaloids in plant culture. Proc Natl Acad Sci **106**: 13673–13678

Sander H, Hartmann T (1989) Site of synthesis, metabolism and translocation of senecionine *N*-oxide in cultured roots of *Senecio erucifolius*. Plant Cell Tissue Organ Cult **18**: 19–31

Schilling B, Lerch K (1995) Cloning, sequencing and heterologous expression of the monoamine oxidase gene from *Aspergillus niger*. Mol Gen Genet MGG **247**: 430–438

Schmittgen TD, Livak KJ (2008) Analyzing real-time PCR data by the comparative CT method. Nat Protoc **3**: 1101–1108

Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. Nat Methods **9**: 671–675

Sen GL, Blau HM (2006) A brief history of RNAi: the silence of the genes. FASEB J **20**: 1293–1299

Shaner NC, Steinbach PA, Tsien RY (2005) A guide to choosing fluorescent proteins. Nat Methods **2**: 905–909

Sharma P, Padh H, Shrivastava N (2013) Hairy root cultures: A suitable biological system for studying secondary metabolic pathways in plants. Eng Life Sci **13**: 62–75

Shine J, Dalgarno L (1974) The 3′-Terminal sequence of *Escherichia coli* 16S ribosomal RNA: Complementarity to nonsense triplets and ribosome binding sites. Proc Natl Acad Sci U S A **71**: 1342–1346

Sievert C, Beuerle T, Hollmann J, Ober D (2015) Single cell subtractive transcriptomics for identification of cell-specifically expressed candidate genes of pyrrolizidine alkaloid biosynthesis. Phytochemistry **117**: 17–24

Somerville C, Somerville S (1999) Plant functional genomics. Science **285**: 380–383

Sønderby IE, Geu-Flores F, Halkier BA (2010) Biosynthesis of glucosinolates – gene discovery and beyond. Trends Plant Sci **15**: 283–290

Song J, Wang Z (2011) RNAi-mediated suppression of the phenylalanine ammonia-lyase gene in *Salvia miltiorrhiza* causes abnormal phenotypes and a reduction in rosmarinic acid biosynthesis. J Plant Res **124**: 183–192

Staiger C (2012) Comfrey: A Clinical Overview. Phytother Res **26**: 1441–1448

Stegelmeier BL, Edgar JA, Colegate SM, Gardner DR, Schoch TK, Coulombe RA, Molyneux RJ (1999) Pyrrolizidine alkaloid plants, metabolism and toxicity. J Nat Toxins **8**: 95–116

Steudle E, Peterson CA (1998) How does water get through roots? J Exp Bot **49**: 775–788

Stoltzfus A (1999) On the Possibility of Constructive Neutral Evolution. J Mol Evol **49**: 169–181

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

Sun J, Morita H, Chen G, Noguchi H, Abe I (2012) Molecular cloning and characterization of copper amine oxidase from *Huperzia serrata*. Bioorg Med Chem Lett **22**: 5784–5790

Suzuki K, Yamada Y, Hashimoto T (1999a) Expression of *Atropa belladonna* putrescine *N*-methyltransferase gene in root pericycle. Plant Cell Physiol **40**: 289–297

Suzuki K, Yun D-J, Chen X-Y, Yamada Y, Hashimoto T (1999b) An *Atropa belladonna* hyoscyamine 6β-hydroxylase gene is differentially expressed in the root pericycle and anthers. Plant Mol Biol **40**: 141–152

Tiburcio AF, Altabella T, Bitrián M, Alcázar R (2014) The roles of polyamines during the lifespan of plants: from development to stress. Planta **240**: 1–18

Tropf S, Lanz T, Rensing SA, Schröder J, Schröder G (1994) Evidence that stilbene synthases have developed from chalcone synthases several times in the course of evolution. J Mol Evol **38**: 610–618

Utelli A-B, Roy BA (2001) Causes and consequences of floral damage in *Aconitum lycoctonum* at high and low elevations in Switzerland. Oecologia **127**: 266–273

Veena V, Taylor CG (2007) *Agrobacterium rhizogenes*: recent developments and promising applications. Vitro Cell Dev Biol - Plant **43**: 383–403

Vimolmangkang S, Deng X, Owiti A, Meelaph T, Ogutu C, Han Y (2016) Evolutionary origin of the NCSI gene subfamily encoding norcoclaurine synthase is associated with the biosynthesis of benzylisoquinoline alkaloids in plants. Sci Rep **6**: 26323

Wang J, Wang Y, Gao C, Jiang L, Guo D (2017) PPero, a computational model for plant PTS1 type peroxisomal protein prediction. PLOS ONE **12**: e0168912

Weber S, Eisenreich W, Bacher A, Hartmann T (1999) Pyrrolizidine alkaloids of the lycopsamine type: biosynthesis of trachelanthic acid. Phytochemistry **50**: 1005–1014

Weid M, Ziegler J, Kutchan TM (2004) The roles of latex and the vascular bundle in

morphine biosynthesis in the opium poppy, *Papaver somniferum*. Proc Natl Acad Sci U S A **101**: 13957–13962

Weng J-K, Philippe RN, Noel JP (2012) The rise of chemodiversity in plants. Science **336**: 1667–1670

Wiedenfeld H, Edgar J (2011) Toxicity of pyrrolizidine alkaloids to humans and ruminants. Phytochem Rev **10**: 137–151

Wise A, Liu Z, Binns A (2006) Three Methods for the Introduction of Foreign DNA into *Agrobacterium*. *In* K Wang, ed, Agrobacterium Protoc. Humana Press, pp 43–54

Wittstock U, Gershenzon J (2002) Constitutive plant toxins and their role in defense against herbivores and pathogens. Curr Opin Plant Biol **5**: 300–307

Yang Y, Smith SA (2013) Optimizing de novo assembly of short-read RNA-seq data for phylogenomics. BMC Genomics **14**: 328

Zangerl AR, Rutledge CE (1996) The probability of attack and patterns of constitutive and induced defense: A test of optimal defense theory. Am Nat **147**: 599–608

Zarei A, Trobacher CP, Cooke AR, Meyers AJ, Hall JC, Shelp BJ (2015) Apple fruit copper amine oxidase isoforms: Peroxisomal MdAO1 prefers diamines as substrates, whereas extracellular MdAO2 exclusively utilizes monoamines. Plant Cell Physiol **56**: 137–147

Zenk MH, Juenger M (2007) Evolution and current status of the phytochemistry of nitrogenous compounds. Phytochemistry **68**: 2757–2772

Zhang J (2003) Evolution by gene duplication: an update. Trends Ecol Evol **18**: 292–298

Zhou X, Lindsay H, Robinson MD (2014) Robustly detecting differential expression in RNA sequencing data using observation weights. Nucleic Acids Res **42**: e91

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

LEBENSLAUF

Persönliche Daten

Geburtstag:	07.04.1987
Geburtsort:	Preetz
Staatsangehörigkeit:	Deutsch

Promotion

seit 01/2012	**Promotion unter der Betreuung von** **Prof. Dr. Dietrich Ober am Botanischen Institut** Christian-Albrechts-Universität zu Kiel

Studium

10/2009 - 10/2011	**Master of Science in Biologie** Christian-Albrechts-Universität zu Kiel
10/2006 - 10/2009	**Bachelor of Science in Biologie** Christian-Albrechts-Universität zu Kiel

Schule

06/2006	**Abitur bzw. Allgemeine Hochschulreife** Fachgymnasium Preetz

ERKLÄRUNG

Die vorliegende Doktorarbeit ist - abgesehen von der Beratung durch meinen Betreuer Prof. Dr. Dietrich Ober - nach Inhalt und Form meine eigene Arbeit.

Die Arbeit wurde weder ganz noch zum Teil an der CAU Kiel oder an einer anderen Hochschule im Rahmen eines Prüfungsverfahrens vorgelegt.

Die vorliegende Arbeit entstand unter Einhaltung der Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft.

Teilergebnisse aus dieser Arbeit wurden in folgenden Beiträgen vorab veröffentlicht:

Kruse LH, Stegemann T, Sievert C, Ober D (2017) Second site of pyrrolizidine alkaloid biosynthesis in Comfrey boosts plant defense in floral stage. Plant Physiol pp.00265.2017

Lars Hendrik Kruse