Metabolic engineering of biosynthesis and sequestration of artemisinin

Bo Wang

Thesis committee

Promotor
Prof. Dr Harro J. Bouwmeester
Professor of Laboratory Plant Physiology
Wageningen University

Co-promotor
Prof. Dr.ir. AR (Sander) van der Krol
Associate Professor of Laboratory Plant Physiology
Wageningen University

Other members
Prof. Dr AHJ Bisseling, Wageningen University
Prof. Dr A Tissier, Leibniz Institute of Plant Biochemistry, German
Prof. Dr SC de Vries, Wageningen University
Prof. Dr HJ Bosch, Utrecht University / Plant Research International Wageningen

This research was conducted under the auspices of
the Graduate School of Experimental Plant Sciences

Metabolic engineering of biosynthesis and sequestration of artemisinin

Bo Wang

CONTENTS

Chapter   *1*

General Introduction

1 Plant secondary metabolites and terpenes

1.1 plant secondary metabolites

In 1803, the first bio-active plant natural product, morphine, was isolated in its crystalline pure form from the medicinal plant opium poppy by Friedrich Serturner (1783-1841). Although medicinal plants had been used for many centuries by many different cultures, this marked the beginning of modern scientific phytochemistry (Goerig and Schulte am Esch, 1991). In the next two centuries, more and more bio-active plant products were isolated and their structure elucidated. At present over 200,000 different plant derived metabolites have been reported (Hartmann, 2007). This provided the chemical basis for the research on plant secondary metabolism and lead to the development of a natural products industry, which over the last 50 years has begun to exploit the rich diversity of plant secondary metabolites. On the human side, research addresses the understanding of the bioactivity of secondary metabolites, trying to identify their molecular action, improve their target specificity and target delivery. On the plant side, the research mainly addresses the question why and how plant cells synthesize such a rich diversity of secondary metabolites. Moreover, scientists study the possibility to use biotechnology, including genetic engineering, to introduce or enhance the production of secondary metabolites in crops, to improve, for example, nutritional value, environmental stress tolerance or disease resistance. An excellent example of such genetic engineering is the purple tomato, which has strongly enhanced levels of the health-promoting anthocyanin (Butelli et al., 2008). This was achieved through genetic modification by introducing the *Antirrhinum majus* transcription factors *Delila* and *Rosea1* into tomato. As many plant secondary metabolites are toxic to the plant itself, research into the use of biotechnology to increase or introduce secondary metabolite formation in plants should also involve the question how plants can escape or tolerate these potentially toxic compounds.

Secondary metabolites may be classified into different families based on their biosynthetic pathways. The phenolics, which are products of the phenylpropanoid biosynthesis pathway; the terpenoids and steroids, which are products of the terpenoid or isoprenoid biosynthesis pathways; and the alkaloids, which are products of the alkaloid biosynthesis pathway (Bourgaud et al., 2001). Some classes of secondary metabolites are specific to certain plant species. For instance, glucosinolates occur as secondary metabolites in almost all plant species of the order *Brassicales* (e.g. families *Crucifer-*

ae, *Capparidaceae*, and *Caricaceae*) and in the genus *Drypetes* (family *Euphorbiaceae*) (Abdel-Farid et al., 2009; Rodman et al., 1996). While each of these biosynthesis pathways alone can already produce a myriad of complex molecules, the biochemical diversity of plant-derived compounds is further extended by the production of combined products from different pathways. An important example of such hybrid product formation is indole-alkaloid biosynthesis in *Catharanthus roseus*. Indole-alkaloids are composed of a monoterpenoid and an alkaloid moiety. Examples are vincristine and vinblastine, which are important anti-cancer drugs, making them interesting targets for metabolic engineering. Indole-alkaloid biosynthesis involves many enzymatic steps, catalyzed by enzymes that belong to a wide range of protein families (El-Sayed and Verpoorte, 2007). Metabolic engineering approaches for this type of compounds form a good illustration of the importance of sequestration and transport. When the first enzyme of the *C. roseus* monoterpenoid indole alkaloid biosynthetic pathway, geraniol synthase, was heterologously expressed in *tobacco* this resulted in the formation of multiple geraniol derivatives, including further oxidized and double bond reduced geraniol derivatives and their glycosides (Dong et al., 2013). When subsequently a large part of the terpenoid branch of the pathway was reconstructed through transient expression in *N. benthamiana* only a very low flux to the end product was observed (Miettinen et al., 2014). The authors showed that competing side reactions occurred in *N. benthamiana* draining away carbon from the main pathway. Suitable compartmentation and improved transport and sequestration could therefore potentially solve these issues.

1.2 Terpenoids

In my thesis, I will focus on the terpenoids, one of the most important families of plant derived natural secondary metabolites. The terpenoids have a large diversity of functions in plants, for example as plant hormones (gibberellins, ABA, strigolactones and cytokinins), to attract pollinators or in defense against insects (Abel et al., 2009; Chen et al., 2003; Paré and Tumlinson, 1999). They are extensively used as flavors, fragrances, and drugs for humans (Ikram et al., 2015; Rohloff, 2004; Russo, 2011; Singh and Sharma, 2015; Tholl and Lee, 2011). Terpenes are formed by the condensation of 5-carbon isoprene units, the ubiquitous precursors isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). Depending on the number of isoprene building blocks, terpenes may be classified into mono-, sesqui-, di-, sester-, tri-, tetra- and polyterpenes (Table 1). The precursors of terpenes are synthesized from two separate pathways: the

mevalonic acid (MVA) pathway, localized in the cytosol, endoplasmic re-
ticulum and peroxisomes and the 2-C-methyl-D-erythritol 4-phosphate
(MEP) pathway, which is localized in the plastids (Figure 1) (Vranová et
al., 2013). Terpene biosynthesis in the mitochondria may be supported by
import of IPP from the cytosolic MVA pathway (Figure 1) (Rodríguez-Con-
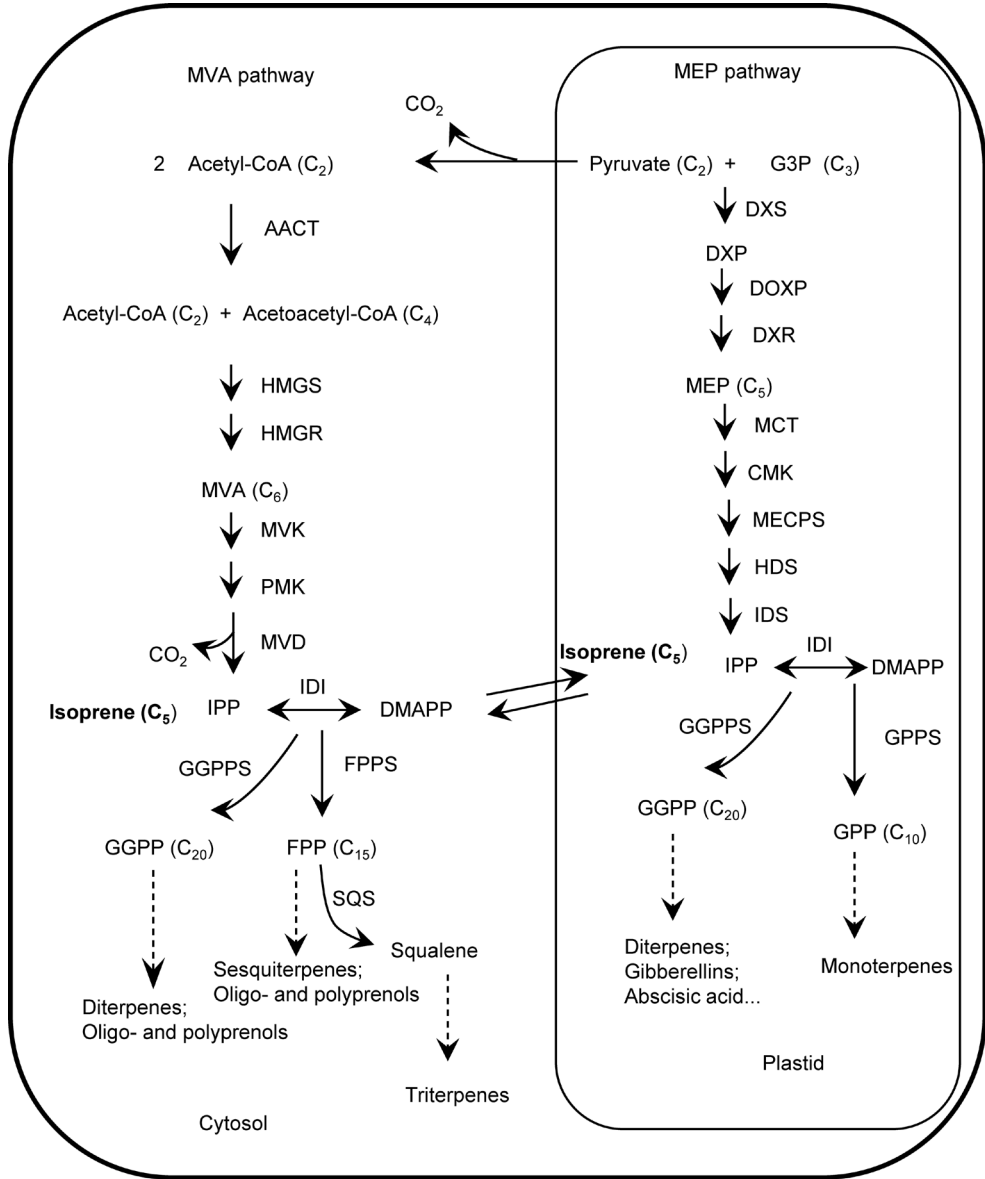cepción and Boronat, 2002; Tholl and Lee, 2011; Vranová et al., 2013).

Figure 1. The biosynthesis pathway of MVA and MEP pathway. Acetoacetyl-CoA thiolase (AACT), HMG-CoA synthase (HMGS), HMG-CoA reductase (HMGR), mevalonate kinase (MVK), phosphomevalonate kinase (PMK), and mevalonate diphosphate decarboxylase (MVD), isopentenyl diphosphate isomerase (IDI), farnesyl diphosphate synthase (FPPS), geranylgeranyl diphosphate synthase (GGPPS), squalene synthase (SQS), deoxy-d-xylulose-5-phosphate synthase (DXS), 1-deoxy-D-xylulose-5-phosphate reductoisomerase (DXR), 2-C-methyl-d-erythritol 4-phosphate cytidylyltransferase (MCT), 4-(cytidine 5'-diphospho)-2-C-methyl-d-erythritol kinase (CMK), 2-C-methyl-d-erythritol 2,4-cyclodiphosphate synthase (MECPS), (E)-4-hydroxy-3-methylbut-2-enyl diphosphate synthase (HDS) and isopentenyl diphosphate synthase (IDS), Geranyl diphosphate synthase (GPPS), terpene synthases (TPS).

Table 1. Classification of the terpenes

Terpenes	Isoprene units	Carbon atoms
Monoterpene	2	10
Sesquiterpene	3	15
Diterpene	4	20
Sesterterpenes	5	25
Triterpenes	6	30
Tetraterpenes	8	40
Polyterpenes	> 40	> 40

The MEP pathway starts with the condensation of pyruvate and d-glyceraldehyde 3-phosphate (G3P) to 1-deoxy-xylulose-5-phosphate (DXP) by DXP synthase (DXS) (Rodríguez-Concepción, 2006). Then, the formation of the C5 unit MEP is catalyzed by DXP reducto-isomerase (DXR) (Rodríguez-Concepción and Boronat, 2002). The downstream steps from MEP to IPP are catalyzed by 2-C-methyl-d-erythritol 4-phosphate cytidylyltransferase (MCT), 4-(cytidine 5'-diphospho)-2-C-methyl-d-erythritol kinase (CMK), 2-C-methyl-d-erythritol 2,4-cyclodiphosphate synthase (MECPS), (E)-4-hydroxy-3-methylbut-2-enyl diphosphate synthase (HDS) and isopentenyl diphosphate synthase (IDS). Cytosolic and plastidal IPP can be exchanged, and in both compartments can be isomerized into DMAPP by isopentenyl diphosphate isomerase (IDI), providing an intrinsic link between the plastidal MEP pathway and the cytosolic MVA pathway (Vickers et al., 2014; Vranová et al., 2013). Geranyl diphosphate synthase (GPPS) catalyzes the condensation of one molecule each of IPP and DMAPP to produce GPP (C10), the universal precursor of the monoterpenes (Hemmerlin et al., 2012; Tholl and Lee, 2011; Vickers et al., 2014; Vranová et al., 2013). In addition, DMAPP and three molecules of IPP is catalyzed by Geranylgeranyl diphosphate synthase (GGPPS) into GGPP (C20) in plastids, cytosol, endoplasmic reticulum (ER) and mitochondria (Okada et al., 2000).

The first step in the MVA pathway is the condensation of two molecules of acetyl-CoA in the cytosol to acetoacetyl-CoA by the acetoacetyl-CoA thiolase (AACT) (Lombard and Moreira, 2011). HMG-CoA synthase (HMGS) and HMG-CoA reductase (HMGR) subsequently catalyze the formation of MVA (C6) (Gardner and Hampton, 1999) (Figure 1). Mevalonate is further converted to IPP by mevalonate kinase (MVK), phosphomevalonate kinase (PMK), and mevalonate diphosphate decarboxylase (MVD), respectively. IPP is transformed into DMAPP by IDI, a divalent metal ion-requiring enzyme in the cytosol (Dorsey and Porter, 1968; Vranová et al., 2013). Condensation of one DMAPP and two IPP molecules catalyzed by farnesyl diphosphate synthase (FPPS) leads to the formation of FPP (C15), the natural precursor of sesquiterpenes (Cunillera et al., 1996).

Following the formation of the prenyl diphosphate precursors GPP and FPP, the monoterpene and sesquiterpene skeleton is generated by the action of terpene synthases (TPS). The primary products of these terpene synthases are often volatile. Subsequent enzymatic steps catalyzed by cytochrome P450s, reductases, dehydrogenases, acyl transferases, glycosyl transferases, etc then create the enormous structural diversity in the mono- and sesquiterpenoids which may be less volatile and with a wide range of hydrophilic or hydrophobic. Diterpenes are formed from GGPP. For the biosynthesis of triterpenes, FPP is converted into 2,3-oxidosqualene by the squalene synthase (SQS) and squalene monooxygenase or epoxidase (SQE). Then 2,3-oxidosqualene is cyclized by oxidosqualene cyclase (OSC) to form triterpenes (Abe et al., 1993). Also in the di- and triterpenes an enormous structural diversity is subsequently created through the action of additional enzymes as described above for the mono- and sesquiterpenes.

Function of induced terpene biosynthesis

Terpenes are often produced and stored in specialized glands or trichomes below the cuticular wax layer of the cells, so that also potential terpenes do not evaporate easily. However, when a plant is damaged by herbivorous insects, the leaves release stored volatiles and newly produced volatiles from the induction of terpene synthases (Paré and Tumlinson, 1999). This blend of volatile terpenes functions in communication between and among species, and as messengers between members of the same species (Paré and Tumlinson, 1999; Witzany, 2006). They may be toxic or repellent to the attacking insect, or they attract parasitic and predatory enemies of the insects (Law and Regnier, 1971; Paré and Tumlinson, 1999). Also flowers release

volatiles. These are important for plant reproduction as they attract pollinators (Chen et al., 2003; Dudareva et al., 2013; Dudareva et al., 2006; Huang et al., 2012). (*E*)-β-caryophyllene, for example, of which the biosynthesis is catalyzed by caryophyllene synthase (CST), *AtTPS21*, that is specifically expressed in the flowers, is the major volatile organic compound emitted from *Arabidopsis thaliana* flowers (Chen et al., 2003; Huang et al., 2012). In addition to its role in attracting pollinators it was shown to be an effective defense compound against a bacterial pathogen (Abel et al., 2009).

1.3 Metabolic engineering of secondary metabolism

With the identification of the biosynthetic pathway genes of secondary metabolites, metabolic engineering to increase their production comes within reach. However, initial attempts showed that production of engineered products was low (Farhi et al., 2011; Lange et al., 2011; Wallaart et al., 2001). Therefore strategies should be investigated to improve this production, for example by increasing the precursor pool, eliminating competing pathways, targeting the enzymes to a different subcellular compartment with higher precursor pool and intron-enhanced gene transcription.

Increasing the precursor pool

Increasing the precursor pool could provide more substrate for the biosynthesis genes. So far, most efforts have focused on ectopic expression of genes encoding the early, rate-limiting, steps of the pathway. HMGR is a key rate-limiting step in the MVA pathway (Goldstein and Brown, 1990). Overexpression of a truncated HMGR resulted in increased production of endogenous sterols (Chappell et al., 1995) as well as the heterologous sesquiterpenoid, patchoulol, in tobacco (Wu et al., 2006; Zhan et al., 2014). Overexpression of DXS, the key rate-determining enzyme in the MEP pathway (Rodríguez-Concepción, 2006), in *Arabidopsis*, increased accumulation of many plastidic isoprenoids including chlorophyll, α-tocopherol, carotenoids, and abscisic acid (Estévez et al., 2001; Wright et al., 2014).

Eliminating competing pathways

Endogenous enzymes in plants cells may modify the pathway intermediates of end products that are produced by ectopic expression of a pathway in a heterologous host, such conversions may decrease the production of the

target compound (Staniek et al., 2013). Competing modification enzymes may be reductases, dehydrogenases and oxidases, as well as conjugating enzymes such as glycosyl transferases and glutathione S-transferases (Dudareva et al., 2004; Renault et al., 2014; Shi and Xie, 2014; Staniek et al., 2013; Strasser et al., 2014). These modifications and conjugations formation has been reported for many metabolically engineered compounds, such as geraniol (formation of glycosides), costunolide (formation of glutathione conjugates) and artemisininic acid (formation of glycosides) (Dong et al., 2013; Liu et al., 2011b; Ting et al., 2013). A successful example of blocking the competing enzymes was carried out in *E. coli*. A dehydrogenases, *YjgB*, was knocked out to increase heterologous geraniol production by preventing geraniol dehydrogenation. This resulted in increased purity and production of geraniol by more than 20% (Zhou et al., 2014).

Targeting expression

Plant cells contain multiple membrane-delimited compartments (Small et al., 1998). Enzymes and products of secondary metabolism may be in different compartments of the cell. For example, flavonoid metabolism occurs on the cytoplasmic side of the endoplasmic reticulum but the products of the flavonoid biosynthesis pathway are transported to the vacuole by transporters (Zhao and Dixon, 2010). In addition, enzymes in one metabolic biosynthesis pathway may be localized in different compartments. Moreover, alkaloids are produced within multiple cell types and precursors in an alkaloid biosynthesis pathway may be translocated intracellularly - between cytosol, vacuole, ER, vesicles, and thylakoid membrane - as well as intercellularly (Deus-Neumann and Zenk, 1986; McCaskill et al., 1988; Qu et al., 2015; Roytrakul and Verpoorte, 2007; St-Pierre et al., 1999).

Enzymes may show targeting to different compartments in different plant species. In terpene biosynthesis, for example, geraniol synthases have been reported with both plastidic and cytosolic localization (Dong et al., 2013). Using metabolic engineering, a default subcellular compartmentation may be artificially altered or circumvented. For example, artificially plastid targeted geraniol synthase produced 30% more than upon cytosolic targeting (Dong et al., 2013). Furthermore, in an exciting metabolic engineering approach, (Kumar et al., 2012) remodeled the cytosolic MVA pathway into the tobacco chloroplast, which increased the synthesis of isoprenoid metabolites including mevalonate, carotenoids, sterols, and squalene.

Introns-enhanced gene transcription

Introns have a dramatic effect on gene expression in plants because they may contain enhancer elements or alternative promoters, or may increase mRNA accumulation by intron-mediated enhancement (IME) (Rose, 2008). For example, a heterologous intron increased expression in *Arabidopsis* relative to the intronless control, even for tissue specific genes, which indicates that different genes respond similarly to the same stimulatory intron (Emami et al., 2013). This approach has not fully been exploited yet to increase the production of secondary metabolites in metabolic engineering approaches but the results suggest that it could be worthwhile to study the possible effect of introns in enhancing of secondary metabolites biosynthesis upon heterologous expression of pahtway genes in plant cells.

1.4 Transport in secondary metabolism

The formation of side products in metabolic engineering, as described above, suggests that transport and sequestration may be important issues to solve if we want to be able to exploit metabolic engineering to the full. Many secondary metabolites are produced in certain cell types only, such as trichomes in tobacco or laticifers in chicory and sequestered potential toxic secondary metabolites outside the cells in subcuticular space or in laticifers. Biotechnological approaches should therefore at least consider this, as sequestration outside of the cell could potentially improve the yield of the engineered natural product. Several membrane transport events may be involved in secondary metabolite transport and/or sequestration, such as glycosylation formation and then transport to the vacuole (1), transport into the vacuole directly (2), transport the secondary molecules into ER for the further modification (3), secretion of metabolites outside the cell by vesicles or non-vesicular trafficking (4), export from the cytosol to the apoplast through the plasma membrane - a process often carried out by ABC transporters (5), and export of plastid-produced secondary metabolites to ER or cytosol (6) (Figure 2). In addition, we considered a potential role of LTPs which reside in the cell wall (7).

1.4.1 ATP-binding cassette (ABC) transporters

ABC proteins are large, ubiquitous, and diverse membrane transporters that carry their substrates over biological membranes by an ATP-dependent reaction. Plant ABC transporters can be localized in most membranes, such as the vacuolar membrane, the plasma membrane, and the membranes of plastids,
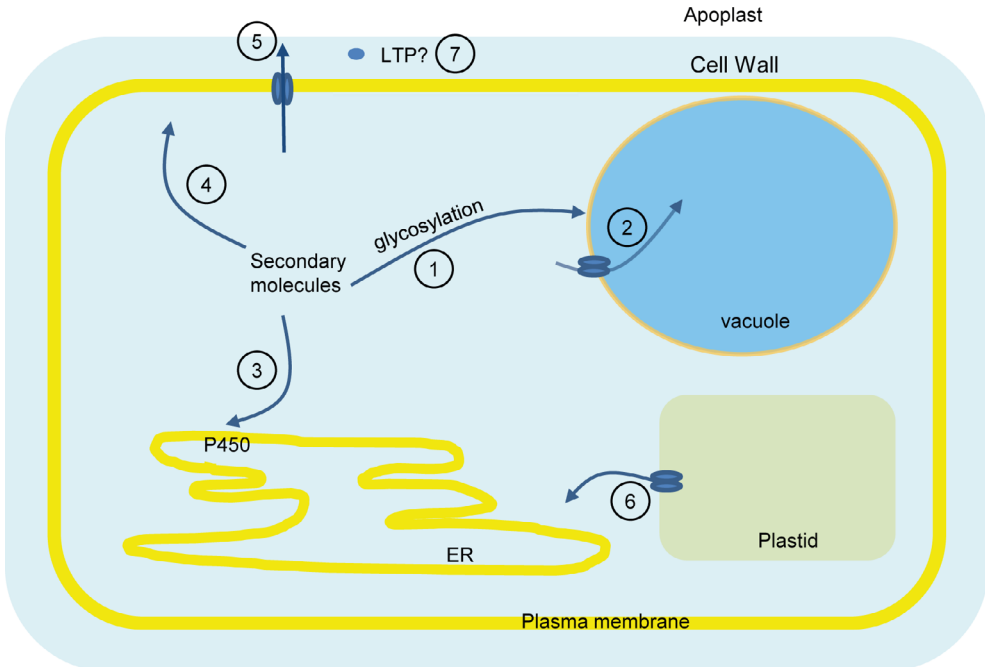
Figure 2. Potential transport processes for secondary metabolites in plants. 1, glycosides formation and transport to the vacuole; 2, transport of cytosol-produced secondary metabolites into the vacuole; 3, transport the secondary molecules into ER for the further modification; 4, secretion of metabolites outside the cell by vesicles or non-vesicular trafficking; 5, export of secondary metabolites from the cytosol to the apoplast; 6, export of plastid-produced secondary metabolites to ER or cytosol.

peroxisomes, mitochondria and endoplasmic reticulum. In plants, ABC transporters contribute to multiple functions, ranging from detoxification and a role in resistance against pathogens and abiotic stress, to a role in plant development.

The core structure of ABC-transporters consists of two different types of domains, a nucleotide-binding domain (NBD, or ATP-binding cassettes) and a transmembrane domain (TMD). The NBD has a highly conserved sequence. Two NBD monomers form a closed dimer which can bind two ATP molecules at the dimer interface in the cytoplasm. This will result in the closed dimer to open up and ATP hydrolysis and release of ADP and Pi (Higgins and Linton, 2004). This sustained cyclic opening and closing of the NBD dimer during ATP binding and release of ADP and Pi provides energy for transporting a substrate across the membrane (Doshi et al., 2010). The TMD is embedded in the membrane bilayer to connect the cytosolic and apoplastic side. The substrate binding site of the TMD recognizes the substrate. Then two TMD modules form a solvent-accessible space to transport the substrate across the membrane.

The ABCG subfamily contains both half-size proteins, called White–Brown Complex (WBC) transporters, and full-size proteins coined Pleiotropic Drug Resistance (PDR) transporters. WBC transporters are involved in transport of eye pigment precursors in *Drosophila melanogaster* (Mackenzie et al., 2000). NtWBC1 was the first reported transporter of this subfamily in plants. It was shown to play an important role in plant reproduction probably because it is involved in exudate production in the stigma and anther dehiscence (Otsu et al., 2004). *Arabidopsis* ABCG11/WBC11 interacts with ABCG12/CER5 and together these transporters form a key component of the export pathway for cutin and wax secretion through the plasma membrane of Arabidopsis epidermis cells to the cell wall and cuticle (Bird et al., 2007; Panikashvili et al., 2007).

PDRs play a role in the export of metabolites before their concentration reaches cellular toxic levels. Goossens *et al* reported that *PDR* genes can be used to stimulate the secretion of secondary metabolites in plant cell suspension cultures (Goossens et al., 2003) and when *Saccharomyces cerevisiae PDR5* was expressed in tobacco BY2 (*Nicotiana tabacum* L. cv Bright-Yellow 2) cells, this resulted in improved secretion of alkaloids. NpPDR1 transports the antifungal diterpene, sclareol, onto the leaf surface of *Nicotiana plumbaginifolia,* and NpPDR1 knock out lines displayed increased susceptibility to a group of fungal and oomycete pathogens (Bultreys et al., 2009; Stukkens et al., 2005). Treatment of suspension-cultured tobacco BY-2 cells with the bacterial elicitor, flagellin, induced *NtPDR1* expression (Sasabe et al., 2002). The plasma membrane localized NtPDR1 transports diterpenes involved in the plant defense response in *Nicotiana tabacum* (Crouzet et al., 2013). Another PDR transporter, from *C. roseus*, *CrTPT2*, is expressed predominantly in the epidermis of young leaves and the level of catharanthine–vindoline drug dimers was increased when it was silenced (Yu and De Luca, 2013). The suppression of *CrTPT2* expression blocks the efflux of catharanthine, resulting in a decline in the leaf-surface catharanthine levels while increasing catharanthine levels within the leaf. Despite this scattered evidence for the role of ABC transporters in secondary metabolite transport, still little is known about the (possible) involvement of PDR transporters in secondary metabolite formation, and especially in the transport of mono- and sesquiterpenes.

1.4.2 Lipid transfer proteins (LTPs)

LTPs are abundant, small, basic proteins that are able to bind and exchange hydrophobic compounds between membranes (Yeats and Rose, 2008). This transfer process is distinct from the secretory vesicles of the ER (Holthuis

and Menon, 2014; Lev, 2010). The size of LTPs ranges from 6.5 ~ 10.5 kD (Liu et al., 2015) and an eight-cysteine residue conserved motif forms the backbone of the LTPs (Liu et al., 2015; Yeats and Rose, 2008). The LTP protein forms a large internal hydrophobic cavity, containing lipid-binding sites (Liu et al., 2015; Wang et al., 2012). Four disulfide bonds between the cysteine residues stabilize this hydrophobic cavity structure (Liu et al., 2015; Samuel et al., 2002). In mammals and yeast, LTPs play an important role in non-vesicular trafficking of cholesterol, phospholipids, and sphingolipids between biological membranes (Lev, 2010; Prinz, 2007). LTPs regulate diverse lipid-mediated cellular processes in ER-to-Golgi and Golgi-to-plasma membrane trafficking (De Matteis et al., 2007; Lev, 2010; Prinz, 2007). The first report about LTPs playing an essential role in non-vesicular trafficking was in yeast, in which *Sec14* loss of function produced defects in membrane trafficking from the Golgi complex to the plasma membrane (Bankaitis et al., 1990; Bankaitis et al., 1989). *Sec14* was shown to be involved in controlling the production of diacylglycerol and phosphatidylinositol-4-P that play a key role in the formation of plasma-membrane-destined transport carriers at the Golgi complex (Bankaitis et al., 2012; Mousley et al., 2007). In mammalian cells, the lipid transfer protein CERT (CERamide Transfer protein) specifically binds ceramide from the ER membrane and delivers it to the Golgi membrane (Hanada et al., 2003; Kawano et al., 2006; Simanshu et al., 2013). In addition to the activity at the surface of the Golgi complex, the trafficking from the outer to the inner Golgi membrane by LTPs was reported in mammals and yeast(Gatta et al., 2015; Georgiev et al., 2011; Mesmin et al., 2011). Deleting all OSBP (sterol specific LTPs, OxySterol Binding Protein) homologs in yeast reduced the sterol trafficking from the plasma membrane to the ER (Georgiev et al., 2011). Apart from OSBPs, STeroidogenic Acute Regulatory Transfer (START) proteins are also sterol specific LTPs, and they can transfer sterols from the plasma membrane to the ER in both mammals and yeast (Gatta et al., 2015; Mesmin et al., 2011).

In plants, not much is known about the roles of LTPs, with a few exceptions. Almost all plant LTPs carry an N-terminal signal peptide directing the protein to the apoplast (Edstam et al., 2011). Arabidopsis DIR1 is an example of such a putative apoplastic LTP. It seems to interact with lipophilic molecules to promote long distance pathogen defense signaling during systemic acquired resistance and a plasma membrane receptor seems to be required to activate the resistance response (Lascombe et al., 2006; Maldonado et al., 2002). Blein et al. (Blein et al., 2002) showed that wheat LTP1 binds with elicitin receptors in the plasma membrane to induce plant cell responses (agonist behavior). However, whether this LTP1 forms a com-

plex with lipids to interact with the elicitin receptor is still unknown. MtN5 belongs to a group of LTPs having a role in plant-microbe interaction (Pii et al., 2010). Overexpression of *MtN5* in *Medicago truncatula* led to an increased number of nodules when induced by *Sinorhizobium meliloti* (Pii et al., 2010). Silencing of *OsC6*, one of the LTPs in rice, resulted in defective development of lipidic orbicules (i.e. Ubisch bodies) and pollen exine during anther development and reduced pollen fertility (Zhang et al., 2010). Although OsC6 displays lipid binding activity *in vitro*, there is no clue of which specific lipid molecules it could bind to for regulating anther development.

In addition to the function to regulate cell signalling, plant LTPs are also active in lipid secretion. Tobacco *NtLTP1*, a glandular-trichome specific LTP, is required for diterpene secretion from glandular trichomes and transgenic tobacco plants overexpressing *NtLTP1* displayed increased resistance against aphids (Choi et al., 2012). Interestingly, LTP genes show very high transcriptional activity in glandular trichomes, which produce terpenes, in many plant species (peppermint (Lange et al., 2000), alfalfa (Aziz et al., 2005), *Artemisia annua* (Bertea et al., 2005), hops (Wang et al., 2008), *Salvia fruticosa* (Chatzopoulou et al., 2010), tomato (Schilmiller et al., 2010) and tobacco (Harada et al., 2010). This suggests that LTPs may play a role in terpene transport and sequestration.

1.4.3 Glutathione S-transferases (GSTs)

GSTs are abundant soluble proteins encoded by a highly divergent, ancient gene family, which can be divided on the basis of sequence identity into the phi, tau, theta, zeta and lambda classes (Edwards and Dixon, 2005; Sheehan et al., 2001). Plant GSTs conjugate electrophilic xenobiotics with glutathione (GSH) for vacuolar sequestration and catalyze the reduction of toxic organic hydroperoxides (Edwards et al., 2000). The detoxification process of electrophilic compounds is mediated by three groups of enzymes (Ishikawa, 1992; Marrs, 1996; Sandermann Jr, 1992). Substrate specific enzymes such as cytochrome P450 monooxygenases introduce functional groups to substrates in the transformation phase; enzymes such as UDP-glucosyltransferases and GSTs use this functional group as a site of conjugation resulting in a less toxic and more water-soluble conjugate in the conjugation stage; transporter enzymes, often ATP-dependent membrane pumps such as the ABC transporters described above, recognize and transfer conjugates across membranes for excretion or sequestration in the compartmentation stage.

GSTs might also have non-catalytic roles and serve as carrier proteins for transport of phytochemicals between cellular compartments (Edwards et al.,

2000). Since cytosolic accumulation of anthocyanin is toxic to cells and results in feed-back inhibition (Licciardello et al., 2014), GSH conjugation is required for anthocyanin transport and sequestration into the vacuole (Licciardello et al., 2014; Mueller et al., 2000). *Bronze-2(Bz2)* is a GST involved in vacuolar transfer of flavonoids in maize (Marrs et al., 1995). Bz2 conjugates GSH to cyanidin 3-glucoside and transports the resulting product into the vacuole via a tonoplast Mg-ATP–requiring glutathione pump (GS-X pump) (Alfenito et al., 1998). In *Arabidopsis*, the GST, Transparent testa19 (TT19), functions as a carrier to transport cyanidin and/or anthocyanins across the tonoplast (Sun et al., 2012). Multidrug Resistance-associated Protein (MRP)-type ABC transporters can transport glutathione conjugates into the vacuole in maize (Goodman et al., 2004; Klein et al., 2006). ABCC1 from grape berry, which shares a high amino acid sequence identity with Zm-MRP3 (Francisco et al., 2013; Goodman et al., 2004), acts as an anthocyanin transporter that depends on GSH without the formation of an anthocyanin-GSH conjugate (Francisco et al., 2013). Except MRP/ABCC transporters, multidrug and toxic extrusion (MATE) transporter, *Transparent testa12* (*TT12*), was shown to be involved in flavonoid sequestration in the vacuole of cells in the seed coat endothelium (Debeaujon et al., 2001). In plants, GSTs bind indole-3-acetic acid (IAA) as carrier proteins while only a small amount of IAA-GSH conjugates was detected (LeClere et al., 2002; Pěnčík et al., 2013), which allow temporary storage or IAA uptake from membranes and trafficking to receptors (Marrs, 1996).

2 Artemisinin and its biosynthesis pathway

Malaria is the most devastating infectious disease in the tropics (Talisuna et al., 2004). Current estimates indicate that up to three million people die of malaria each year in Africa (Breman et al., 2004; Ntoumi F, 2004). The medicinal plant *Artemisia annua (*commonly called Sweet wormwood, Qinghao*)* has been used for treatment of malaria symptoms for over 2000 years in China. The medical use of *A. annua* as herbal tea to treat malaria was recorded in "Compendium of Materia Medica" by Li Shenzhen in 1596. The main bio-active compound in *A. annua* is artemisinin, a sesquiterpene lactone containing an endoperoxide bridge. Artemisinin was extracted and for the first time purified from leaves of *A. annua* by Project 523 between 1969 and 1972 under the supervision of Youyou Tu, and was named *Qinghaosu* (Chinese: 青蒿素) (Faurant, 2011). Artemisinin is widely used for the treatment of malaria, and has saved millions of lives. Therefore, this year the Nobel Prize in Physiology or Medicine was awarded to Youyou Tu. The World

health Organization (WHO) has recommended artemisinin-based combined therapy (ACT) for the treatment of malaria as the first choice from 2001. However, the production of artemisinin in *A. annua* varies from 0.01% to 0.8% DW (Liu et al., 2011a) and even the high artemisinin yielding hybrid selected by Mediplant in Switzerland only produces up to 1.95% DW of artemisinin (Simonnet et al., 2008). Because of these low contents artemisinin supply by isolation from *A. annua* does not satisfy the global demand (Durante et al., 2011) and several initiatives to solve this using biotechnology in both plants as well as micro-organisms were developed as I will explain below.

2.1 Biosynthesis of artemisinin

Artemisinin is produced from the precursor FPP, which is generated in the MVA pathway (Figure 3). The first committed step of artemisinin biosynthesis is the cyclization of FPP to amorpha-4,11-diene (AD) by amorphadiene synthase (ADS) (Bouwmeester et al., 1999). AD is subsequently oxidized to artemisinic alcohol (AAOH), artemisinic aldehyde (AAA) and artemisinic acid (AA) by a cytochrome P450 enzyme, CYP71AV1, which is anchored to the endoplasmic reticulum (ER) (Covello et al., 2007; Teoh et al., 2006). In *A. annua*, AA is non-enzymatically converted to arteannuin B (AB) (Figure 3) (Brown and Sy, 2007). However, the exocyclic double bond of AAA can also be reduced by the soluble cytosolic protein DBR2, resulting in dihydroartemisinic aldehyde (DHAAA) (Zhang et al., 2008). DHAAA is oxidized either by CYP71AV1 or an alcohol dehydrogenase (ALDH1) to dihydroartemisinic acid (DHAA), which is the direct precursor of artemisinin (Teoh et al., 2009). It is assumed that the oxidation from DHAA to artemisinin occurs non-enzymatically (Figure 3) (Brown, 2010).

Chemotypes of A. *annua*

A. annua has two different chemotypes according to the relative level of the end products in the branched biosynthesis pathway: arteannuin B and artemisinin (Figure 3). The high artemisinin production (HAP) chemotypes also have a high DHAA level and a relatively low AA level, while the low artemisinin production (LAP) chemotypes have a relatively low level of DHAA compared with AA (Wallaart et al., 2000). On the other hand, these two chemotypes also show a differential response to elicitation by methyl jasmonate (MeJA) (Wu et al., 2011). As a result of exogenous application of MeJA, HAP plants display increased accumulation of DHAA and artemisinin while only a trace amount of AA can be detected. For the HAP chemotype there is a cor-
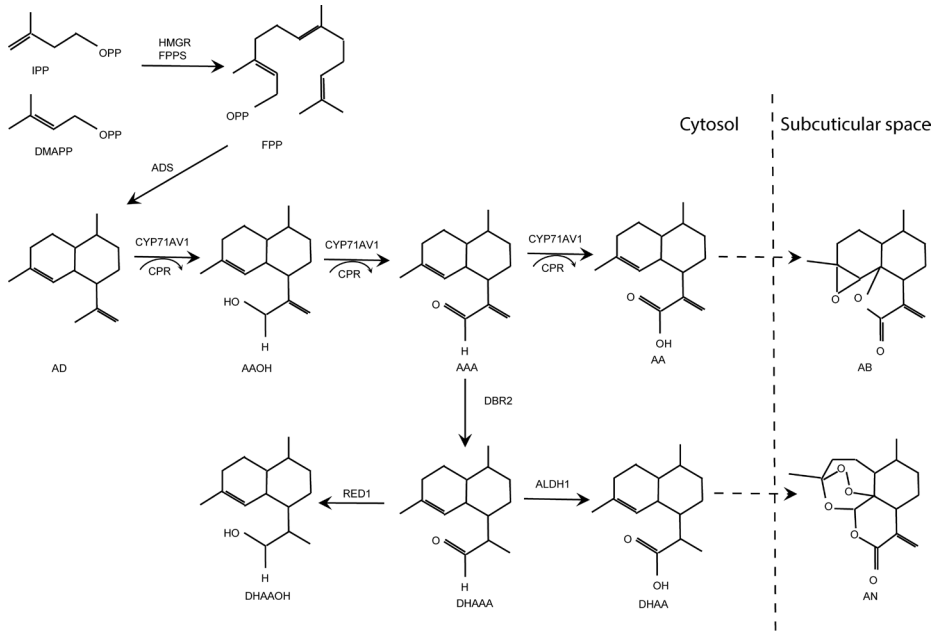
Figure 3. Biosynthesis pathway of artemisinin. IPP, isopentenyl diphosphate; DMAPP, di-methylallyl diphosphate; FPP, farnesyl diphosphate; AD, amorpha-4,11-diene; AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; AB, arteannu-in B; DHAAOH, dihydroartemisinic alcohol; DHAAA, dihydroartemisinic aldehyde; DHAA, dihydroartemisinic acid; FPPS, farnesyl diphosphate synthase; HMGR, 3-hy-droxy-3-methylglutarylCoA reductase; ADS, amorphadiene synthase; CYP71AV1, amor-phadiene oxidase; CPR, cytochrome P450 reductase; DBR2, artemisinic aldehyde double-bond reductase; RED1, dihydroartemisinic aldehyde reductase 1; ALDH1, alde-hyde dehydrogenase 1. Broken arrows indicate the involvement of more than one step.

responding activation of most artemisinin pathway genes by MeJA except for the cytochrome P450 reductase (CPR) (Wu et al., 2011). In contrast, AA and artemisinin decreased dramatically under MeJA elicitation for the LAP plants in favor of AA and arteannuin B production (Wu et al., 2011), which matches with the fact that the expression of the biosynthesis genes *ADS*, *CYP71AV1* and *DBR2* (but not *ALDH1*) were downregulated at various time points after JA treatment. The accumulation of AA in LAP plants and DHAA in HAP plants suggests that the conversion of these precursors represents a rate-limiting step in the biosynthesis of both arteannuin B and artemisinin, respectively.

2.2 Metabolic engineering of artemisinin

2.2.1 Metabolic engineering in *A. annua*

Because of its highly valued product artemisinin, many attempts have been made to increase artemisinin production in *A. annua*. Efforts have been aimed at selecting HAP chemotypes (Wallaart et al., 2000; Wu et al., 2011), higher numbers of glandular trichomes (Lommen et al., 2006; Maes et al., 2011) and stimulation of the pathway genes by (Me)JA (Maes et al., 2011; Wu et al., 2011). Also, a number of genes involved in artemisinin biosynthesis have been overexpressed in *A. annua* to enhance artemisinin production (Tang et al., 2014). Indeed, artemisinin production was increased by overexpression of the genes encoding ADS and CYP71AV1 (Alam and Abdin, 2011; Ma et al., 2009; Shen et al., 2012). Alternatively, it was studied whether production may be increased by blocking competing reactions. By blocking the FPP consuming enzyme squalene synthase (SQS), for example, artemisinin production was increased 3.14-fold compared with the wildtype (Zhang et al., 2009). Since DHAAA reductase 1 (RED1) competes with ALDH1 for DHAAA (Figure 3), it was suggested that knocking-down RED1 may lead to increased production of artemisinin in *A. annua* (Rydén et al., 2010).

The possibility to increase the expression of all artemisinin biosynthetic genes at once was explored by identifying transcription factors involved in the coordinated expression of the pathway genes. The transcription factor AaWRKY1 was reported to regulate artemisinin production (De Muynck et al., 2009) and overexpression of *AaWRKY1* in *A. annua* resulted in an artemisinin content in transgenic lines of 24.5 mg/g (DW), which is 4.4-fold that of the control lines (De Muynck et al., 2009; Tang et al., 2014). The jasmonate-responsive AP2/ERF-type transcription factors *AaERF1* and *AaERF2* also positively elevate the transcript levels of both *ADS* and *CYP71AV1,* and both artemisinin and AA content were reduced in transgenic lines of *A. annua* in which *AaERF1* and *AaERF2* expression were suppressed through RNAi (Yu et al., 2012). Up-regulation or down-regulation of the expression of *AaORA*, a trichome-specific AP2/ERF transcription factor, in *A. annua* resulted in a significant increase or decrease, respectively, in artemisinin and DHAA (Lu et al., 2013). AaORA did not only positively regulate the expression of *ADS*, *CYP71AV1* and *DBR2* but also stimulated the expression of *AaERF1* in *A. annua.*

Because it had been shown that JA can increase the expression of the artemisinin biosynthesis pathway genes attempts were made to improve artemisinin production in *A. annua* by constitutive expression of allene

oxide cyclase (AOC), the key enzyme in the biosynthetic pathway of jas-monates. Expression of the *A. annua* AOC resulted in an almost 2-fold increase (87.9 %) in artemisinin in transgenic lines (Tang et al., 2014).

2.2.2 Metabolic engineering in heterologous plants

Although all artemisinin biosynthetic pathway genes have been identified and cloned, there is no or only very low levels of artemisinin detected when these biosynthetic genes are overexpressed in heterologous plant hosts, either using transient expression (Ting et al., 2013; Zhang et al., 2011) or in sta-bly transformed tobacco plants (Farhi et al., 2011). Heterologous expression of *ADS*, *CYP71AV1*, *DBR2* and *ALDH1* in tobacco resulted in the accumula-tion of amorphadiene and AAOH and DHAAOH, but not AA and DHAA (Zhang et al., 2011). Transient expression of a multi-gene construct containing *ADS*, *FPPS*, t-*HMGR* and *CYP71AV1* - with ribosomal skipping sequences and mitochondrial targeting of ADS - in *N. benthamiana* resulted in the formation of AA which was efficiently conjugated by endogenous glycosyl transferases (Van Herpen et al., 2010). The transient expression of *CYP71AV1* (that was isolated from a LAP chemotype) together with the other artemisinin biosyn-thetic genes in *N. benthamiana* resulted in high production of glycoside-con-jugated artemisinin precursors (Ting et al., 2013). Transgenic tobacco, expressing five plant- and yeast-derived genes involved in the MVA and arte-misinin pathways in a single vector produced 0.48 - 6.8 mg/g dry weight arte-misinin, which is comparable to the production in *A. annua* (Farhi et al., 2011).

2.2.3 Metabolic engineering in microbes

The heterologous production of artemisinin in micro-organisms is another potential solution for the shortage of artemisinin. Plasmid-based expression of *S. cerevisiae* genes encoding the MVA pathway in combination with a codon-optimized ADS in *Escherichia coli* resulted in the estimated produc-tion of amorpha-4,11-diene up to 22.6 to 112.2 mg/L (Martin et al., 2003). The same group reported the engineering of *Saccharomyces cerevisiae* to produce up to 100 mg per liter of AA using an engineered MVA pathway, in combination with ADS and CYP71AV1 (Ro et al., 2006). The conversion of AA into artemisinin was achieved by photochemical transformation involv-ing a singlet-oxygen-induced ene reaction and the addition of triplet oxy-gen, which triggers the reaction cascade that incorporates the endoperox-ide group (Lévesque and Seeberger, 2012). The biosynthetic pathway was further optimized by including an NAD-dependent alcohol dehydrogenase

(ADH1), and a ubiquitous electron transport protein of cytochrome P450s in the ER membrane, cytochrome b5 (CYb5). This further improved AA yield and fermentation titres reached 25 grams per liter of AA in *Saccharomyces cerevisiae* (Paddon et al., 2013). In addition, a chemical source of singlet oxygen was developed to convert DHAA to artemisinin, thus avoiding the need for specialized photochemical equipment (Paddon et al., 2013). The successfully produced semi-synthetic artemisinin is now in commercial production, demonstrating the potential of synthetic biology for the development and production of pharmaceutical compounds (Paddon and Keasling, 2014).

2.3 Transport in artemisinin biosynthesis pathway

Although the metabolic engineering work in plants improved the production of artemisinin, it is still far away from being economically attractive. Upon over-expression of the artemisinin biosynthesis genes In heterologous plants only a very low production of artemisinin was detected. This may result from the subcellular compartmentalization of the artemisinin biosynthesis genes, and the therefore suboptimal operation of the pathway in a heterologous system. For example, DBR2 is a soluble cytosolic protein while CYP71AV1 is embedded in the ER, so the rate of intracellular transport of intermediates between these enzymes may influence the metabolic flow (Olofsson et al., 2011). Thus, it is important to know how these molecules are transported intracellularly, but also intercellularly and/or how they are sequestered in the apoplast, where artemisinin is stored naturally, in the glandular trichomes, in *A. annua*. However, so far there is little knowledge about artemisinin transport. As far as I know there is just one reference about transporters in *A. annua*. After identifying the ABC transporter unigenes in the NCBI database, the transcription levels over the leaf development and under methyl jasmonate or abscisic acid treatments were tested, however, no any transport activity and products information were included in this literature (Zhang et al., 2012).

3 Thesis outline

The objective of my PhD project is to gain knowledge artemisinin biosynthesis pathway and to improve the production of artemisinin using different metabolic engineering strategies. After the general introduction in **Chapter 1**, **Chapter 2** compares the artemisinin biosynthesis genes in two different *A. annua* chemoptypes. The ectopic expression of the artemisinin biosynthesis

pathway in *N. benthamiana* is established. Then two different biosynthesis genes, CYP71AV1, are identified and their contribution to the production difference between the chemotypes investigated. In **Chapter 3**, I explore the sequestration of artemisinin and its biosynthetic pathway intermediates in the apoplast. Two PDRs and three LTPs are identified and isolated from *A. annua*. PDR2 and LTP3 stimulate the sequestering of intermediates DHAA and AA in the apoplast. **Chapter 4** describes how the activity of the flavonoid transcription factor *Rosea1* (*ROS*) stimulates sesquiterpene metabolic flux. Combining ROS with the *A. annua* transporter genes, *PDR2* and *LTP3*, improves sequestration of DHAA and AA in the apolast and the final products artemisinin and arteannuin B are enhanced in *N. benthamiana*. **Chapter 5** explores the role of introns in enhancing sesquiterpene production. Two sesquiterpene synthases, *CST* from *Arabidopsis*, and *ADS* from *A. annua* are included in both stable transgenic and transient expression assays with and without introns. Genomic sesquiterpene synthase overexpression yields higher production of the anticipated products. The results of this thesis are integrated and discussed in a broader perspective in **Chapter 6**. The future perspectives for metabolic engineering of natural secondary metabolites and the importance of transport are discussed as well.

Reference

Abdel-Farid, I.B., Jahangir, M., van den Hondel, C.A.M.J.J., Kim, H.K., Choi, Y.H., and Verpoorte, R. (2009). Fungal infection-induced metabolites in *Brassica rapa*. Plant Science *176*, 608-615.

Abe, I., Rohmer, M., and Prestwich, G.D. (1993). Enzymatic cyclization of squalene and oxidosqualene to sterols and triterpenes. Chemical Reviews *93*, 2189-2206.

Abel, C., Clauss, M., Schaub, A., Gershenzon, J., and Tholl, D. (2009). Floral and insect-induced volatile formation in *Arabidopsis lyrata ssp. petraea*, a perennial, outcrossing relative of *A. thaliana*. Planta *230*, 1-11.

Alam, P., and Abdin, M.Z. (2011). Over-expression of HMG-CoA reductase and amorpha-4,11-diene synthase genes in *Artemisia annua* L. and its influence on artemisinin content. Plant Cell Rep *30*, 1919-1928.

Alfenito, M.R., Souer, E., Goodman, C.D., Buell, R., Mol, J., Koes, R., and Walbot, V. (1998). Functional complementation of anthocyanin sequestration in the vacuole by widely divergent glutathione S-transferases. The Plant Cell *10*, 1135-1149.

Aziz, N., Paiva, N., May, G., and Dixon, R. (2005). Transcriptome analysis of alfalfa glandular trichomes. Planta *221*, 28-38.

Bankaitis, V.A., Aitken, J.R., Cleves, A.E., and Dowhan, W. (1990). An essential role for a phospholipid transfer protein in yeast Golgi function. Nature *347*, 561-562.

Bankaitis, V.A., Ile, K.E., Nile, A.H., Ren, J., Ghosh, R., and Schaaf, G. (2012). Thoughts on Sec14-like nanoreactors and phosphoinositide signaling. Advances in Biological Regulation *52*, 115-121.

Bankaitis, V.A., Malehorn, D.E., Emr, S.D., and Greene, R. (1989). The *Saccharomyces cerevisiae* SEC14 gene encodes a cytosolic factor that is required for transport of secretory proteins from the yeast Golgi complex. The Journal of Cell Biology *108*, 1271-1281.

Bertea, C.M., Freije, J.R., van der Woude, H., Verstappen, F.W.A., Perk, L., Marquez, V., De Kraker, J.W., Posthumus, M.A., Jansen, B.J.M., de Groot, A*., et al.* (2005). Iden-tification of intermediates and enzymes involved in the early steps of artemisinin biosynthesis in *Artemisia annua*. Planta Medica *71*, 40-47.

Bird, D., Beisson, F., Brigham, A., Shin, J., Greer, S., Jetter, R., Kunst, L., Wu, X., Ye-phremov, A., and Samuels, L. (2007). Characterization of *Arabidopsis* ABCG11/WBC11, an ATP binding cassette (ABC) transporter that is required for cuticular lipid secretion. The Plant Journal *52*, 485-498.

Blein, J.-P., Coutos-Thévenot, P., Marion, D., and Ponchet, M. (2002). From elicitins to lipid-transfer proteins: a new insight in cell signalling involved in plant defence mechanisms. Trends in Plant Science *7*, 293-296.

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

Bouwmeester, H.J., Wallaart, T.E., Janssen, M.H.A., van Loo, B., Jansen, B.J.M., Post-humus, M.A., Schmidt, C.O., De Kraker, J.-W., König, W.A., and Franssen, M.C.R. (1999). Amorpha-4,11-diene synthase catalyses the first probable step in artemisi-nin biosynthesis. Phytochemistry *52*, 843-854.

Breman, J.G., Alilio, M.S., and Mills, A. (2004). Conquering the intolerable burden of malaria: what's new, what's needed: a summary. The American Journal of Tropical Medicine and Hygiene *71*, 1-15.

Brown, G.D. (2010). The biosynthesis of artemisinin (Qinghaosu) and the phyto-chemistry of *Artemisia annua* L. (Qinghao). Molecules *15*, 7603-7698.

Brown, G.D., and Sy, L.-K. (2007). In vivo transformations of artemisinic acid in *Arte-misia annua* plants. Tetrahedron *63*, 9548-9566.

Bultreys, A., Trombik, T., Drozak, A., and Boutry, M. (2009). *Nicotiana plumbagini-folia* plants silenced for the ATP-binding cassette transporter gene NpPDR1 show increased susceptibility to a group of fungal and oomycete pathogens. Mol Plant Pathol *10*, 651-663.

Butelli, E., Titta, L., Giorgio, M., Mock, H.-P., Matros, A., Peterek, S., Schijlen,

E.G.W.M., Hall, R.D., Bovy, A.G., Luo, J., *et al.* (2008). Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. Nature Biotechnology *26*, 1301-1308.

Chappell, J., Wolf, F., Proulx, J., Cuellar, R., and Saunders, C. (1995). Is the reaction catalyzed by 3-hydroxy-3-methylglutaryl coenzyme A reductase a rate-limiting step for isoprenoid biosynthesis in plants? Plant Physiology *109*, 1337-1343.

Chatzopoulou, F., Makris, A., Argiriou, A., Degenhardt, J., and Kanellis, A. (2010). EST analysis and annotation of transcripts derived from a trichome-specific cDNA library from *Salvia fruticosa*. Plant Cell Rep *29*, 523-534.

Chen, F., Tholl, D., D'Auria, J.C., Farooq, A., Pichersky, E., and Gershenzon, J. (2003). Biosynthesis and Emission of Terpenoid Volatiles from *Arabidopsis* Flowers. The Plant Cell *15*, 481-494.

Choi, Y.E., Lim, S., Kim, H.-J., Han, J.Y., Lee, M.-H., Yang, Y., Kim, J.-A., and Kim, Y.-S. (2012). Tobacco NtLTP1, a glandular-specific lipid transfer protein, is required for lipid secretion from glandular trichomes. The Plant Journal *70*, 480-491.

Covello, P.S., Teoh, K.H., Polichuk, D.R., Reed, D.W., and Nowak, G. (2007). Functional genomics and the biosynthesis of artemisinin. Phytochemistry *68*, 1864-1871.

Crouzet, J., Roland, J., Peeters, E., Trombik, T., Ducos, E., Nader, J., and Boutry, M. (2013). NtPDR1, a plasma membrane ABC transporter from *Nicotiana tabacum*, is involved in diterpene transport. Plant Molecular Biology *82*, 181-192.

Cunillera, N., Arró, M., Delourme, D., Karst, F., Boronat, A., and Ferrer, A. (1996). *Arabidopsis thaliana* contains two differentially expressed farnesyl-diphosphate synthase genes. Journal of Biological Chemistry *271*, 7774-7780.

De Matteis, M.A., Di Campli, A., and D'Angelo, G. (2007). Lipid-transfer proteins in membrane trafficking at the Golgi complex. Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids *1771*, 761-768.

De Muynck, B., Navarre, C., Nizet, Y., Stadlmann, J., and Boutry, M. (2009). Different subcellular localization and glycosylation for a functional antibody expressed in *Nicotiana tabacum* plants and suspension cells. Transgenic Res *18*, 467-482.

Debeaujon, I., Peeters, A.J.M., Léon-Kloosterziel, K.M., and Koornneef, M. (2001). The *Transparent Testa12* Gene of *Arabidopsis* encodes a multidrug secondary transporter-like protein required for flavonoid sequestration in vacuoles of the seed coat endothelium. The Plant Cell *13*, 853-871.

Deus-Neumann, B., and Zenk, M.H. (1986). Accumulation of alkaloids in plant vacuoles does not involve an ion-trap mechanism. Planta *167*, 44-53.

Dong, L., Miettinen, K., Goedbloed, M., Verstappen, F.W.A., Voster, A., Jongsma, M.A., Memelink, J., Krol, S.v.d., and Bouwmeester, H.J. (2013). Characterization of two geraniol synthases from *Valeriana officinalis* and *Lippia dulcis*: similar activity

but difference in subcellular localization. Metabolic Engineering *20*, 198-211.

Dorsey, J.K., and Porter, J.W. (1968). The inhibition of mevalonic kinase by geranyl and farnesyl pyrophosphates. Journal of Biological Chemistry *243*, 4667-4670.

Doshi, R., Woebking, B., and van Veen, H.W. (2010). Dissection of the conformational cycle of the multidrug/lipidA ABC exporter MsbA. Proteins: Structure, Function, and Bioinformatics *78*, 2867-2872.

Dudareva, N., Klempien, A., Muhlemann, J.K., and Kaplan, I. (2013). Biosynthesis, function and metabolic engineering of plant volatile organic compounds. New Phytologist *198*, 16-32.

Dudareva, N., Negre, F., Nagegowda, D.A., and Orlova, I. (2006). Plant volatiles: recent advances and future perspectives. Critical Reviews in Plant Sciences *25*, 417-440.

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

Durante, M., Caretto, S., Quarta, A., De Paolis, A., Nisi, R., and Mita, G. (2011). β-Cyclodextrins enhance artemisinin production in *Artemisia annua* suspension cell cultures. Appl Microbiol Biotechnol *90*, 1905-1913.

Edstam, M.M., Viitanen, L., Salminen, T.A., and Edqvist, J. (2011). Evolutionary History of the Non-Specific Lipid Transfer Proteins. Molecular Plant *4*, 947-964.

Edwards, R., and Dixon, D.P. (2005). Plant glutathione transferases. In Methods in Enzymology, S. Helmut, and P. Lester, eds. (Academic Press), pp. 169-186.

Edwards, R., Dixon, D.P., and Walbot, V. (2000). Plant glutathione S-transferases: enzymes with multiple functions in sickness and in health. Trends in Plant Science *5*, 193-198.

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

Emami, S., Arumainayagam, D., Korf, I., and Rose, A.B. (2013). The effects of a stimulating intron on the expression of heterologous genes in *Arabidopsis thaliana*. Plant Biotechnology Journal *11*, 555-563.

Estévez, J.M., Cantero, A., Reindl, A., Reichler, S., and León, P. (2001). 1-deoxy-d-xylulose-5-phosphate Synthase, a limiting enzyme for plastidic isoprenoid biosynthesis in plants. Journal of Biological Chemistry *276*, 22901-22909.

Farhi, M., Marhevka, E., Ben-Ari, J., Algamas-Dimantov, A., Liang, Z., Zeevi, V., Edelbaum, O., Spitzer-Rimon, B., Abeliovich, H., Schwartz, B.*, et al.* (2011). Generation of the potent anti-malarial drug artemisinin in tobacco. Nature Biotechnology *29*, 1072-1074.

Faurant, C. (2011). From bark to weed: The history of artemisinin. Parasite : journal de la Société Française de Parasitologie *18*, 215-218.

Francisco, R.M., Regalado, A., Ageorges, A., Burla, B.J., Bassin, B., Eisenach, C., Zarrouk, O., Vialet, S., Marlin, T., Chaves, M.M.*, et al.* (2013). ABCC1, an ATP binding cassette protein from grape berry, transports anthocyanidin 3-O-glucosides. The Plant Cell *25*, 1840-1854.

Gardner, R.G., and Hampton, R.Y. (1999). A highly conserved signal controls degradation of 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) reductase in eukaryotes. Journal of Biological Chemistry *274*, 31671-31678.

Gatta, A.T., Wong, L.H., Sere, Y.Y., Calderón-Noreña, D.M., Cockcroft, S., Menon, A.K., and Levine, T.P. (2015). A new family of START domain proteins at membrane contact sites has a role in ER-PM sterol transport. eLife *4*, e07253.

Georgiev, A.G., Sullivan, D.P., Kersting, M.C., Dittman, J.S., Beh, C.T., and Menon, A.K. (2011). Osh proteins regulate membrane sterol organization but are not required for sterol movement between the ER and PM. Traffic *12*, 1341-1355.

Goerig, M., and Schulte am Esch, J. (1991). Friedrich wilhelm adam sertürner - dem entdecker des morphins zum 150. todestag. Anästhesiol Intensivmed Notfallmed Schmerzther *26*, 492-498.

Goldstein, J.L., and Brown, M.S. (1990). Regulation of the mevalonate pathway. Nature *343*, 425-430.

Goodman, C.D., Casati, P., and Walbot, V. (2004). A multidrug resistance–associated protein involved in anthocyanin transport in *Zea mays*. The Plant Cell *16*, 1812-1826.

Goossens, A., Hakkinen, S.T., Laakso, I., Oksman-Caldentey, K.M., and Inze, D. (2003). Secretion of secondary metabolites by ATP-binding cassette transporters in plant cell suspension cultures. Plant Physiology *131*, 1161-1164.

Hanada, K., Kumagai, K., Yasuda, S., Miura, Y., Kawano, M., Fukasawa, M., and Nishijima, M. (2003). Molecular machinery for non-vesicular trafficking of ceramide. Nature *426*, 803-809.

Harada, E., Kim, J.-A., Meyer, A.J., Hell, R., Clemens, S., and Choi, Y.-E. (2010). Expression profiling of tobacco leaf trichomes identifies genes for biotic and abiotic stresses. Plant and Cell Physiology *51*, 1627-1637.

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

Hemmerlin, A., Harwood, J.L., and Bach, T.J. (2012). A raison d'être for two distinct pathways in the early steps of plant isoprenoid biosynthesis? Progress in Lipid Research *51*, 95-148.

Higgins, C.F., and Linton, K.J. (2004). The ATP switch model for ABC transporters. Nature Structural & Molecular Biology *11*, 918-926.

Holthuis, J.C.M., and Menon, A.K. (2014). Lipid landscapes and pipelines in mem-

brane homeostasis. Nature *510*, 48-57.

Huang, M., Sanchez-Moreiras, A.M., Abel, C., Sohrabi, R., Lee, S., Gershenzon, J., and Tholl, D. (2012). The major volatile organic compound emitted from *Arabidopsis thaliana* flowers, the sesquiterpene (E)-β-caryophyllene, is a defense against a bacterial pathogen. New Phytologist *193*, 997-1008.

Ikram, N.K.B.K., Zhan, X., Pan, X.-W., King, B.C., and Simonsen, H.T. (2015). Stable heterologous expression of biologically active terpenoids in green plant cells. Frontiers in Plant Science *6*, 129.

Ishikawa, T. (1992). The ATP-dependent glutathione S-conjugate export pump. Trends in Biochemical Sciences *17*, 463-468.

Kawano, M., Kumagai, K., Nishijima, M., and Hanada, K. (2006). Efficient trafficking of ceramide from the endoplasmic reticulum to the Golgi apparatus requires a VAMP-associated protein-interacting FFAT motif of CERT. Journal of Biological Chemistry *281*, 30279-30288.

Klein, M., Burla, B., and Martinoia, E. (2006). The multidrug resistance-associated protein (MRP/ABCC) subfamily of ATP-binding cassette transporters in plants. FEBS Letters *580*, 1112-1122.

Kumar, S., Hahn, F.M., Baidoo, E., Kahlon, T.S., Wood, D.F., McMahan, C.M., Cornish, K., Keasling, J.D., Daniell, H., and Whalen, M.C. (2012). Remodeling the isoprenoid pathway in tobacco by expressing the cytoplasmic mevalonate pathway in chloroplasts. Metabolic Engineering *14*, 19-28.

Lange, B.M., Mahmoud, S.S., Wildung, M.R., Turner, G.W., Davis, E.M., Lange, I., Baker, R.C., Boydston, R.A., and Croteau, R.B. (2011). Improving peppermint essential oil yield and composition by metabolic engineering. Proceedings of the National Academy of Sciences of the United States of America *108*, 16944-16949.

Lange, B.M., Wildung, M.R., Stauber, E.J., Sanchez, C., Pouchnik, D., and Croteau, R. (2000). Probing essential oil biosynthesis and secretion by functional evaluation of expressed sequence tags from mint glandular trichomes. Proceedings of the National Academy of Sciences *97*, 2934-2939.

Lascombe, M.-B., Buhot, N., Bakan, B., Marion, D., Blein, J.P., Lamb, C.J., and Prangé, T. (2006). Crystallization of DIR1, a LTP2-like resistance signalling protein from *Arabidopsis thaliana*. Acta Crystallographica Section F: Structural Biology and Crystallization Communications *62*, 702-704.

Law, J.H., and Regnier, F.E. (1971). Pheromones. Annual Review of Biochemistry *40*, 533-548.

LeClere, S., Tellez, R., Rampey, R.A., Matsuda, S.P.T., and Bartel, B. (2002). Characterization of a family of IAA-amino acid conjugate hydrolases from *Arabidopsis*. Journal of Biological Chemistry *277*, 20446-20452.

Lev, S. (2010). Non-vesicular lipid transport by lipid-transfer proteins and beyond. Nature Reviews Molecular Cell Biology *11*, 739-750.

Lévesque, F., and Seeberger, P.H. (2012). Continuous-flow synthesis of the anti-malaria drug artemisinin. Angewandte Chemie International Edition *51*, 1706-1709.

Licciardello, C., D'Agostino, N., Traini, A., Recupero, G.R., Frusciante, L., and Chiusano, M.L. (2014). Characterization of the glutathione S-transferase gene family through ESTs and expression analyses within common and pigmented cultivars of *Citrus sinensis* (L.) Osbeck. BMC Plant Biology *14*, 39-39.

Liu, B., Wang, H., Du, Z., Li, G., and Ye, H. (2011a). Metabolic engineering of artemisinin biosynthesis in *Artemisia annua* L. Plant Cell Rep *30*, 689-694.

Liu, F., Zhang, X., Lu, C., Zeng, X., Li, Y., Fu, D., and Wu, G. (2015). Non-specific lipid transfer proteins in plants: presenting new advances and an integrated functional analysis. Journal of Experimental Botany.

Liu, Q., Majdi, M., Cankar, K., Goedbloed, M., Charnikhova, T., Verstappen, F.W.A., de Vos, R.C.H., Beekwilder, J., van der Krol, S., and Bouwmeester, H.J. (2011b). Reconstitution of the costunolide biosynthetic pathway in yeast and *Nicotiana benthamiana*. PLoS ONE *6*, e23255.

Lombard, J., and Moreira, D. (2011). Origins and early evolution of the mevalonate pathway of isoprenoid biosynthesis in the three domains of life. Molecular Biology and Evolution *28*, 87-99.

Lommen, W.J.M., Schenk, E., Bouwmeester, H.J., and Verstappen, F.W.A. (2006). Trichome dynamics and artemisinin accumulation during development and senescence of *Artemisia annua* leaves. Planta Medica *72*, 336-345.

Lu, X., Zhang, L., Zhang, F., Jiang, W., Shen, Q., Zhang, L., Lv, Z., Wang, G., and Tang, K. (2013). AaORA, a trichome-specific AP2/ERF transcription factor of *Artemisia annua*, is a positive regulator in the artemisinin biosynthetic pathway and in disease resistance to *Botrytis cinerea*. New Phytologist *198*, 1191-1202.

Ma, C., Wang, H., Lu, X., Wang, H., Xu, G., and Liu, B. (2009). Terpenoid metabolic profiling analysis of transgenic *Artemisia annua* L. by comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry. Metabolomics *5*, 497-506.

Mackenzie, S., Howells, A., Cox, G., and Ewart, G. (2000). Sub-cellular localisation of the white/scarlet ABC transporter to pigment granule membranes within the compound eye of *Drosophila Melanogaster*. Genetica *108*, 239-252.

Maes, L., Nieuwerburgh, F.C.W.V., Zhang, Y., Reed, D.W., Pollier, J., Casteele, S.R.F.V., Inzé, D., Covello, P.S., Deforce, D.L.D., and Goossens, A. (2011). Dissection of the phytohormonal regulation of trichome formation and biosynthesis of the antimalarial compound artemisinin in *Artemisia annua* plants. New Phytologist *189*, 176-

189.

Maldonado, A.M., Doerner, P., Dixon, R.A., Lamb, C.J., and Cameron, R.K. (2002). A putative lipid transfer protein involved in systemic resistance signalling in *Arabidopsis*. Nature *419*, 399-403.

Marrs, K.A. (1996). The functions and regulation of glutathione S-transferases in plants. Annual Review of Plant Physiology and Plant Molecular Biology *47*, 127-158.

Marrs, K.A., Alfenito, M.R., Lloyd, A.M., and Walbot, V. (1995). A glutathione S-transferase involved in vacuolar transfer encoded by the maize gene *Bronze-2*. Nature *375*, 397-400.

Martin, V.J.J., Pitera, D.J., Withers, S.T., Newman, J.D., and Keasling, J.D. (2003). Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. Nature Biotechnology *21*, 796-802.

McCaskill, D.G., Martin, D.L., and Scott, A.I. (1988). Characterization of alkaloid uptake by *Catharanthus roseus* (L.) G. Don protoplasts. Plant Physiology *87*, 402-408.

Mesmin, B., Pipalia, N.H., Lund, F.W., Ramlall, T.F., Sokolov, A., Eliezer, D., and Maxfield, F.R. (2011). STARD4 abundance regulates sterol transport and sensing. Molecular Biology of the Cell *22*, 4004-4015.

Miettinen, K., Dong, L., Navrot, N., Schneider, T., Burlat, V., Pollier, J., Woittiez, L., van der Krol, S., Lugan, R., Ilc, T.*, et al.* (2014). The seco-iridoid pathway from *Catharanthus roseus*. Nature Communication *5*.

Mousley, C.J., Tyeryar, K.R., Vincent-Pope, P., and Bankaitis, V.A. (2007). The Sec14-superfamily and the regulatory interface between phospholipid metabolism and membrane trafficking. Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids *1771*, 727-736.

Mueller, L.A., Goodman, C.D., Silady, R.A., and Walbot, V. (2000). AN9, a petunia glutathione S-transferase required for anthocyanin sequestration, is a flavonoid-binding protein. Plant Physiology *123*, 1561-1570.

Ntoumi F, D.A., Mbacham W, et al. (2004). The importance and future of malaria research in Africa. The American journal of tropical medicine and hygiene *71*, IV.

Okada, K., Saito, T., Nakagawa, T., Kawamukai, M., and Kamiya, Y. (2000). Five geranylgeranyl diphosphate synthases expressed in different organs are localized into three subcellular compartments in *Arabidopsis*. Plant Physiology *122*, 1045-1056.

Olofsson, L., Engstrom, A., Lundgren, A., and Brodelius, P. (2011). Relative expression of genes of terpene metabolism in different tissues of *Artemisia annua* L. BMC Plant Biology *11*, 45.

Otsu, C.T., daSilva, I., de Molfetta, J.B., da Silva, L.R., de Almeida-Engler, J., Engler, G., Torraca, P.C., Goldman, G.H., and Goldman, M.H.S. (2004). NtWBC1, an ABC transporter gene specifically expressed in tobacco reproductive organs. Journal of Exper-

imental Botany *55*, 1643-1654.

Paddon, C.J., and Keasling, J.D. (2014). Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. Nature Reviews Microbiology *12*, 355-367.

Paddon, C.J., Westfall, P.J., Pitera, D.J., Benjamin, K., Fisher, K., McPhee, D., Leavell, M.D., Tai, A., Main, A., Eng, D*., et al.* (2013). High-level semi-synthetic production of the potent antimalarial artemisinin. Nature *496*, 528-532.

Panikashvili, D., Savaldi-Goldstein, S., Mandel, T., Yifhar, T., Franke, R.B., Höfer, R., Schreiber, L., Chory, J., and Aharoni, A. (2007). The *Arabidopsis* DESPERADO/AtWBC11 transporter is required for cutin and wax secretion. Plant Physiology *145*, 1345-1360.

Paré, P.W., and Tumlinson, J.H. (1999). Plant volatiles as a defense against insect herbivores. Plant Physiology *121*, 325-332.

Pěnčík, A., Simonovik, B., Petersson, S.V., Henyková, E., Simon, S., Greenham, K., Zhang, Y., Kowalczyk, M., Estelle, M., Zažímalová, E*., et al.* (2013). Regulation of auxin homeostasis and gradients in *Arabidopsis* roots through the formation of the indole-3-acetic acid catabolite 2-oxindole-3-acetic acid. The Plant Cell.

Pii, Y., Pandolfini, T., and Crimi, M. (2010). Signaling LTPs: a new plant LTPs sub-family? Plant Signaling & Behavior *5*, 594-597.

Prinz, W.A. (2007). Non-vesicular sterol transport in cells. Progress in lipid research *46*, 297-314.

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

Renault, H., Bassard, J.-E., Hamberger, B., and Werck-Reichhart, D. (2014). Cytochrome P450-mediated metabolic engineering: current progress and future challenges. Current Opinion in Plant Biology *19*, 27-34.

Ro, D.-K., Paradise, E.M., Ouellet, M., Fisher, K.J., Newman, K.L., Ndungu, J.M., Ho, K.A., Eachus, R.A., Ham, T.S., Kirby, J*., et al.* (2006). Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature *440*, 940-943.

Rodman, J.E., Karol, K.G., Price, R.A., and Sytsma, K.J. (1996). Molecules, morphology, and Dahlgren's expanded order *Capparales*. Systematic Botany *21*, 289-307.

Rodríguez-Concepción, M. (2006). Early steps in isoprenoid biosynthesis: multilevel regulation of the supply of common precursors in plant cells. Phytochem Rev *5*, 1-15.

Rodríguez-Concepción, M., and Boronat, A. (2002). Elucidation of the methylerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids. A met-

abolic milestone achieved through genomics. Plant Physiology *130*, 1079-1089.

Rohloff, J. (2004). Essential oil drugs — terpene composition of aromatic herbs. In Production Practices and Quality Assessment of Food Crops, R. Dris, and S.M. Jain, eds. (Springer Netherlands), pp. 73-128.

Rose, A.B. (2008). Intron-mediated regulation of gene expression. In Nuclear pre-mRNA Processing in Plants, A.N. Reddy, and M. Golovkin, eds. (Springer Berlin Heidelberg), pp. 277-290.

Roytrakul, S., and Verpoorte, R. (2007). Role of vacuolar transporter proteins in plant secondary metabolism: *Catharanthus roseus* cell culture. Phytochem Rev *6*, 383-396.

Russo, E.B. (2011). Taming THC: potential cannabis synergy and phytocannabi-noid-terpenoid entourage effects. British Journal of Pharmacology *163*, 1344-1364.

Rydén, A.-M., Ruyter-Spira, C., Quax, W.J., Osada, H., Muranaka, T., Kayser, O., and Bouwmeester, H. (2010). The molecular cloning of dihydroartemisinic aldehyde re-ductase and its implication in artemisinin biosynthesis in *Artemisia annua*. Planta Medica *76*, 1778-1783.

Samuel, D., Liu, Y.-J., Cheng, C.-S., and Lyu, P.-C. (2002). Solution structure of plant nonspecific lipid transfer protein-2 from rice (*Oryza sativa*). Journal of Biological Chemistry *277*, 35267-35273.

Sandermann Jr, H. (1992). Plant metabolism of xenobiotics. Trends in Biochemical Sciences *17*, 82-84.

Sasabe, M., Toyoda, K., Shiraishi, T., Inagaki, Y., and Ichinose, Y. (2002). cDNA cloning and characterization of tobacco ABC transporter: *NtPDR1* is a novel elicitor-respon-sive gene. FEBS Letters *518*, 164-168.

Schilmiller, A.L., Miner, D.P., Larson, M., McDowell, E., Gang, D.R., Wilkerson, C., and Last, R.L. (2010). Studies of a biochemical factory: tomato trichome deep expressed sequence tag sequencing and proteomics. Plant Physiology *153*, 1212-1223.

Sheehan, D., Meade, G., Foley, V.M., and Dowd, C.A. (2001). Structure, function and evolution of glutathione transferases: implications for classification of non-mam-malian members of an ancient enzyme superfamily. Biochemical Journal *360*, 1-16.

Shen, Q., Chen, Y.F., Wang, T., Wu, S.Y., Lu, X., Zhang, L., Zhang, F.Y., Jiang, W.M., Wang, G.F., and Tang, K.X. (2012). Overexpression of the cytochrome P450 mono-oxygenase (cyp71av1) and cytochrome P450 reductase (cpr) genes increased arte-misinin content in *Artemisia annua* (*Asteraceae*). Genetics and Molecular Research *11*, 3298-3309.

Shi, M., and Xie, D. (2014). Biosynthesis and metabolic engineering of anthocyanins in *Arabidopsis thaliana*. Recent Patents on Biotechnology *8*, 47-60.

Simanshu, D.K., Kamlekar, R.K., Wijesinghe, D.S., Zou, X., Zhai, X., Mishra, S.K., Mo-

lotkovsky, J.G., Malinina, L., Hinchcliffe, E.H., Chalfant, C.E.*, et al.* (2013). Non-vesicular trafficking by a ceramide-1-phosphate transfer protein regulates eicosanoids. Nature *500*, 463-467.

Simonnet, X., Quennoz, M., and Carlen, C. (2008). New *Artemisia annua* hybrids with high artemisinin content. ISHS Acta Horticulturae *769*, 371-373.

Singh, B., and Sharma, R.A. (2015). Plant terpenes: defense responses, phylogenetic analysis, regulation and clinical applications. 3 Biotech *5*, 129-151.

Small, I., Wintz, H., Akashi, K., and Mireau, H. (1998). Two birds with one stone: genes that encode products targeted to two or more compartments. Plant molecular biology *38*, 265-277.

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

Staniek, A., Bouwmeester, H., Fraser, P.D., Kayser, O., Martens, S., Tissier, A., van der Krol, S., Wessjohann, L., and Warzecha, H. (2013). Natural products – modifying metabolite pathways in plants. Biotechnology Journal *8*, 1159-1171.

Strasser, R., Altmann, F., and Steinkellner, H. (2014). Controlled glycosylation of plant-produced recombinant proteins. Current Opinion in Biotechnology *30*, 95-100.

Stukkens, Y., Bultreys, A., Grec, S., Trombik, T., Vanham, D., and Boutry, M. (2005). NpPDR1, a pleiotropic drug resistance-type ATP-binding cassette transporter from *Nicotiana plumbaginifolia*, plays a major role in plant pathogen defense. Plant Physiology *139*, 341-352.

Sun, Y., Li, H., and Huang, J.-R. (2012). Arabidopsis TT19 functions as a aarrier to transport anthocyanin from the cytosol to tonoplasts. Molecular Plant *5*, 387-400.

Talisuna, A.O., Bloland, P., and D'Alessandro, U. (2004). History, dynamics, and public health importance of malaria parasite resistance. Clinical Microbiology Reviews *17*, 235-254.

Tang, K., Shen, Q., Yan, T., and Fu, X. (2014). Transgenic approach to increase artemisinin content in *Artemisia annua* L. Plant Cell Rep *33*, 605-615.

Teoh, K.H., Polichuk, D.R., Reed, D.W., and Covello, P.S. (2009). Molecular cloning of an aldehyde dehydrogenase implicated in artemisinin biosynthesis in *Artemisia annua*. Botany *87*, 635-642.

Teoh, K.H., Polichuk, D.R., Reed, D.W., Nowak, G., and Covello, P.S. (2006). *Artemisia annua* L. (*Asteraceae*) trichome-specific cDNAs reveal CYP71AV1, a cytochrome P450 with a key role in the biosynthesis of the antimalarial sesquiterpene lactone artemisinin. FEBS Letters *580*, 1411-1416.

Tholl, D., and Lee, S. (2011). Terpene Specialized Metabolism in *Arabidopsis thali-*

ana. The Arabidopsis Book / American Society of Plant Biologists *9*, e0143.

Ting, H.-M., Wang, B., Rydén, A.-M., Woittiez, L., van Herpen, T., Verstappen, F.W.A., Ruyter-Spira, C., Beekwilder, J., Bouwmeester, H.J., and van der Krol, A. (2013). The metabolite chemotype of *Nicotiana benthamiana* transiently expressing artemisinin biosynthetic pathway genes is a function of CYP71AV1 type and relative gene dosage. New Phytologist *199*, 352-366.

Van Herpen, T.W.J.M., Cankar, K., Nogueira, M., Bosch, D., Bouwmeester, H.J., and Beekwilder, J. (2010). *Nicotiana benthamiana* as a production platform for artemisinin precursors. PLoS ONE *5*, e14222.

Vickers, C.E., Bongers, M., Liu, Q., Delatte, T., and Bouwmeester, H. (2014). Metabolic engineering of volatile isoprenoids in plants and microbes. Plant, Cell & Environment *37*, 1753-1775.

Vranová, E., Coman, D., and Gruissem, W. (2013). Network analysis of the MVA and MEP pathways for isoprenoid synthesis. Annual Review of Plant Biology *64*, 665-700.

Wallaart, T.E., Bouwmeester, H.J., Hille, J., Poppinga, L., and Maijers, N.C.A. (2001). Amorpha-4,11-diene synthase: cloning and functional expression of a key enzyme in the biosynthetic pathway of the novel antimalarial drug artemisinin. Planta *212*, 460-465.

Wallaart, T.E., Pras, N., Beekman and, A.C., and Quax, W.J. (2000). Seasonal variation of artemisinin and its biosynthetic orecursors in plants of *Artemisia annua* of different geographical origin: proof for the existence of chemotypes. Planta Medica *66*, 57-62.

Wang, G., Tian, L., Aziz, N., Broun, P., Dai, X., He, J., King, A., Zhao, P.X., and Dixon, R.A. (2008). Terpene biosynthesis in glandular trichomes of hop. Plant Physiology *148*, 1254-1266.

Wang, N.-J., Lee, C.-C., Cheng, C.-S., Lo, W.-C., Yang, Y.-F., Chen, M.-N., and Lyu, P.-C. (2012). Construction and analysis of a plant non-specific lipid transfer protein database (nsLTPDB). BMC Genomics *13*, S9.

Witzany, G. (2006). Plant communication from biosemiotic perspective: differences in abiotic and biotic signal perception determine content arrangement of response behavior. context determines meaning of meta-, inter- and intraorganismic plant signaling. Plant Signaling & Behavior *1*, 169-178.

Wright, L.P., Rohwer, J.M., Ghirardo, A., Hammerbacher, A., Ortiz-Alcaide, M., Raguschke, B., Schnitzler, J.-P., Gershenzon, J., and Phillips, M.A. (2014). Deoxyxylulose 5-Phosphate Synthase Controls Flux through the Methylerythritol 4-Phosphate Pathway in Arabidopsis. Plant Physiology *165*, 1488-1504.

Wu, S., Schalk, M., Clark, A., Miles, R.B., Coates, R., and Chappell, J. (2006). Redirec-

tion of cytosolic or plastidic isoprenoid precursors elevates terpene production in plants. Nature Biotechnology *24*, 1441-1447.

Wu, W., Yuan, M., Zhang, Q., Zhu, Y., Yong, L., Wang, W., Qi, Y., and Guo, D. (2011). Chemotype-dependent metabolic response to methyl jasmonate elicitation in *Artemisia annua*. Planta Medica *77*, 1048-1053.

Yeats, T.H., and Rose, J.K.C. (2008). The biochemistry and biology of extracellular plant lipid-transfer proteins (LTPs). Protein Science : A Publication of the Protein Society *17*, 191-198.

Yu, F., and De Luca, V. (2013). ATP-binding cassette transporter controls leaf surface secretion of anticancer drug components in *Catharanthus roseus*. Proceedings of the National Academy of Sciences of the United States of America *110*, 15830-15835.

Yu, Z.-X., Li, J.-X., Yang, C.-Q., Hu, W.-L., Wang, L.-J., and Chen, X.-Y. (2012). The Jasmonate-Responsive AP2/ERF Transcription Factors AaERF1 and AaERF2 Positively Regulate Artemisinin Biosynthesis in Artemisia annua L. Molecular Plant *5*, 353-365.

Zhan, X., Zhang, Y.-H., Chen, D.-F., and Simonsen, H.T. (2014). Metabolic engineering of the moss *Physcomitrella patens* to produce the sesquiterpenoids patchoulol and α/β-santalene. Frontiers in Plant Science *5*, 636.

Zhang, D., Liang, W., Yin, C., Zong, J., Gu, F., and Zhang, D. (2010). OsC6, encoding a lipid transfer protein, is required for postmeiotic anther development in rice. Plant Physiology *154*, 149-162.

Zhang, L., Jing, F., Li, F., Li, M., Wang, Y., Wang, G., Sun, X., and Tang, K. (2009). Development of transgenic *Artemisia annua* (Chinese wormwood) plants with an enhanced content of artemisinin, an effective anti-malarial drug, by hairpin-RNA-mediated gene silencing. Biotechnology and Applied Biochemistry *52*, 199-207.

Zhang, L., Lu, X., Shen, Q., Chen, Y., Wang, T., Zhang, F., Wu, S., Jiang, W., Liu, P., Zhang, L.*, et al.* (2012). Identification of putative *Artemisia annua* ABCG transporter unigenes related to artemisinin yield following expression analysis in different plant tissues and in response to methyl jasmonate and abscisic acid treatments. Plant Mol Biol Rep *30*, 838-847.

Zhang, Y., Nowak, G., Reed, D.W., and Covello, P.S. (2011). The production of artemisinin precursors in tobacco. Plant Biotechnology Journal *9*, 445-454.

Zhang, Y., Teoh, K.H., Reed, D.W., Maes, L., Goossens, A., Olson, D.J.H., Ross, A.R.S., and Covello, P.S. (2008). The molecular cloning of artemisinic aldehyde Δ11(13) reductase and its role in glandular trichome-dependent biosynthesis of artemisinin in *Artemisia annua*. Journal of Biological Chemistry *283*, 21501-21508.

Zhao, J., and Dixon, R.A. (2010). The 'ins' and 'outs' of flavonoid transport. Trends in Plant Science *15*, 72-80.

Zhou, J., Wang, C., Yoon, S.-H., Jang, H.-J., Choi, E.-S., and Kim, S.-W. (2014). Engineering *Escherichia coli* for selective geraniol production with minimized endogenous dehydrogenation. Journal of Biotechnology *169*, 42-50.

Chapter 2

The metabolite chemotype of *Nicotiana benthamiana* transiently expressing artemisinin biosynthetic pathway genes is a function of CYP71AV1 type and relative gene dosage

Hieng-Ming Ting[1], Bo Wang[1], Anna-Margareta Rydén[1,2], Lotte Woittiez[1], Teun van Herpen[1], Francel W.A. Verstappen[1], Carolien Ruyter-Spira[1], Jules Beekwilder[1,2], Harro J. Bouwmeester[1] and Alexander van der Krol[1]

[1]Laboratory of Plant Physiology, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands.
[2]Plant Research International, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands.

Abstract

The plant *Artemisia annua* that produces the anti-malaria compound arte-misinin, occurs as high artemisinin production (HAP) and low artemisinin production (LAP) chemotypes. Understanding the molecular basis of *A. annua* chemotype may help optimising artemisinin biosynthesis in heterol-ogous production platforms. We present the first systematic comparison of artemisinin biosynthesis genes to determine factors that contribute to arte-misinic acid (AA) or dihydroartemisinic acid (DHAA) chemotype of agro-in-filtrated leaves with ADS, *CYP71AV1/AMO*, *DBR2* and *ALDH1*. Results show that the enzyme activity of DBR2 and ALDH1 from the two chemo-types does not differ. The Amorphadiene Oxidase from HAP (AMOHAP) showed reduced activity compared to that from LAP chemotype (AMO-LAP), which relates to a seven amino acid N-terminal extension in AMO-LAP compared to AMOHAP. The GFP fusion of both proteins show equal localization to the ER, but AMOLAP may be more stable. Product profile characterisation by LC-QTOF-MS/MS, UPLC-MRM-MS and GC-MS of tran-sient expression in *Nicotiana benthamiana* show that AMOLAP not only dis-played a higher enzyme activity but also affected the ratio of end products (e.g. leaf chemotype), which could be mimicked by reduced gene dosage of AMOLAP in the pathway. However, expression in combination with the *DBR2* and *ALDH1* also resulted in a qualitatively different product profile ('chemotype') when DBR2 infiltration dosage was diluted, shifting satu-rated (dihydro) branch toward unsaturated branch and of the pathway.

Keywords

Artemisia annua, artemisinin, CYP71AV1, *Nicotiana benthamiana*, transient expression.

Introduction

Based on the content of artemisinin and its precursors, two chemotypes of *A. annua*, can be distinguished; the low-artemisinin production (LAP) chemotype and a high-artemisinin production (HAP) chemotype (Wallaart *et al.*, 2000). Both chemotypes contain artemisinin and arteannuin B, but the HAP chemotype has a relatively high content of artemisinin and its presumed precursor dihydroartemisinic acid (DHAA), while the LAP chemotype has a high content of arteannuin B and its presumed precursor artemisinic acid (AA).

The artemisinin biosynthesis pathway has been largely elucidated and the genes required for production of dihydroartemisinic acid, the most likely precursor of artemisinin (*ADS, CYP71AV1, DBR2* and *ALDH1*) have all been described (Bouwmeester *et al.*, 1999, Rydén *et al.*, 2010, Teoh *et al.*, 2009, Teoh *et al.*, 2006, Zhang *et al.*, 2008). Artemisinin is a sesquiterpene lactone endoperoxide, which is synthesized in the cytosol from the general isoprenoid precursors IPP and DMAPP. These are converted to FPP and the first committed step in the artemisinin biosynthetic pathway is the cyclization of FPP to amorpha-4,11-diene (AD) by amorphadiene synthase (Fig. 1) (Bouwmeester *et al.*, 1999, Mercke *et al.*, 2000). In the subsequent step, AD is oxidized by the cytochrome P450 enzyme, CYP71AV1/AMO, to artemisinic alcohol (AAOH), artemisinic aldehyde (AAA) and artemisinic acid (AA) (Fig. 1) (Ro *et al.*, 2006, Teoh *et al.*, 2006). However, the latter mainly occurs in the LAP chemotype. In the HAP chemotype only very little of the AAA is converted to AA, as most of the AAA is converted to dihydroartemisinic aldehyde (DHAAA) by DBR2, the enzyme that reduces the exocyclic double bond of AAA (Fig. 1) (Bertea *et al.*, 2005, Zhang *et al.*, 2008). Supposedly, DHAAA is subsequently oxidized by alcohol dehydrogenase ALDH1 to the final intermediate dihydroartemisinic acid (DHAA) (Bertea *et al.*, 2005, Zhang *et al.*, 2008). The conversion of DHAA to artemisinin, is believed to be a non-enzymatic and spontaneous photo-oxidation reaction (Sy&Brown, 2002, Wallaart *et al.*, 1999). Similarly, in the LAP chemotype, AA is likely spontaneously converted to arteannuin B.

Recently we reported on the (transient) reconstruction of the artemisinin biosynthetic pathway in *Nicotiana benthamiana* leaves, resulting in up to 39.5 mg.kg^{-1} FW of AA (van Herpen et al., 2010). In the present work we analyse the role of DBR2, ALDH1 and CYP71AV1 in determining the 'chemotype' (as defined by the AA and DHAA ratio) of *N. benthamiana* leaves agro-infiltrated with artemisinin biosynthesis genes. Results show that the chemotype is a function of the *CYP71AV1* type and relative dosage of *DBR2* and *ALDH1*.
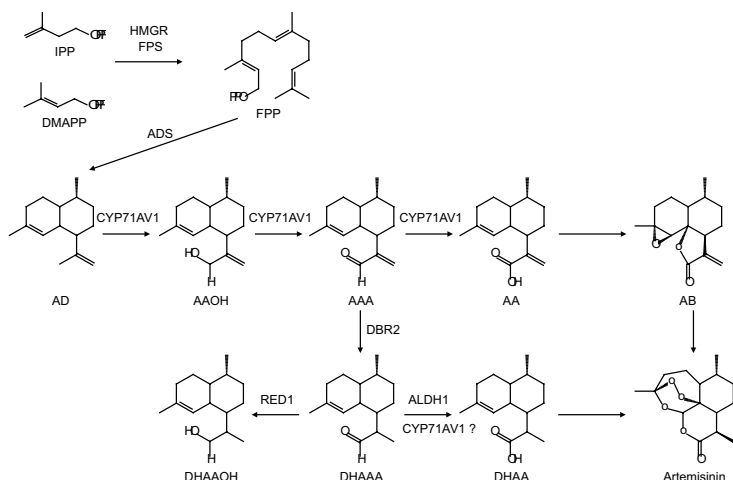
Figure 1. Artemisinin biosynthetic pathway in *Artemisia annua*. IPP: isopentenyl diphosphate; DMAPP: dimethylallyl diphosphate; FPP: farnesyl diphosphate; AD: amorpha-4,11-diene; AAOH: artemisinic alcohol; AAA: artemisinic aldehyde; AA: artemisinic acid; AB: arteannuin B; DHAAOH: dihydroartemisinic alcohol; DHAAA: dihydroartemisinic aldehyde; DHAA: dihydroartemisinic acid; FPS: farnesyl diphosphate synthase; HMGR: 3-hydroxy-3-methylglutaryl-CoA reductase; ADS: amorphadiene synthase; CYP71AV1: amorphadiene oxidase; DBR2: artemisinic aldehyde double-bond reductase; RED1: dihydroartemisinic aldehyde reductase 1; ALDH1: aldehyde dehydrogenase 1. Broken arrows indicate the involvement of more than one step (Nguyen *et al.*, 2011).

Materials and Methods

Cloning of the ADS+FPS+HMGR expression construct AmFH.

Cloning of the *AmFH* expression construct which contains *ADS* with CoxIV mitochondrial targeting signal, mitochondrial targeted *FPS* and a cytosolic (truncated) *HMGR* under control of the CaMV35S promoter has been described previously (van Herpen *et al.*, 2010).

Identification and cloning of CYP71AV1 from HAP and LAP chemotypes

The *AMOHAP* EST sequence was identified in the sequence database of an *Artemisia annua* L. (HAP chemotype) glandular trichome cDNA library (Bertea *et al.*, 2006). This sequence has been deposited in the GenBank database (JQ254992). The full length sequence of *AMOHAP* was obtained by RACE PCR (Clontech, Mountain View, CA, USA). We note that the race PCR did not yield any sequences that indicated the presence of a *AMOLAP*

version in our cDNA library. Subsequently, the full length coding region was amplified from *Artemisia annua* trichome cDNA by PCR using Phusion polymerase (Finnzymes, Espoo, Finland) using primers 1+2 (Table S1). The full length *AMOLAP* coding region was cloned by primers 3+4. After confirmation of the correct sequences both the *AMOHAP* and *AMOLAP* coding region were isolated from pGEMT/AMOHAP and pGEMT/AMOLAP using BamHI and KpnI for cloning into the yeast expression vector pYEDP60 (Pompon *et al.*, 1996), resulting in pYEDP60-AMOHAP and pYEDP60-AMOLAP.

For the cloning into a plant binary expression vector, the genes were first introduced into ImpactVectorC3.1 (http://www.wageningenur.nl/en/ show/Productie-van-farmaceutische-en-industriele-eiwitten-door-planten.htm). For this a Kpn1 site was introduced into C3.1, resulting in C3.1/ Kpn1. Hereto, two oligo's (GATCCATTTCGGTACCAATTAGC and GGC-CGCTAATTGGTACCGAAATG) were hybridized, kinase-treated and ligated into C3.1 digested with BamHI and NotI. The BamHI/KpnI fragment of pGEMT/AMOHAP and pGEMT/AMOLAP was isolated and ligated into the vector C3.1/Kpn1 between the CaMV35S promoter and Rbcs1 terminator. The resulting plasmids C3.1/AMOHAP and C3.1/AMOLAP were digested with AscI and PacI and the full gene sequence was cloned into the AscI and PacI site of the pBinPlus binary vector (van Engelen *et al.*, 1995).

Identification and cloning of DBR2 from HAP and LAP chemotypes

We cloned a *DBR2HAP* from *Artemisia annua* HAP chemotype (Bertea *et al.*, 2006) and *DBR2LAP* was isolated from *Artemisia annua* LAP chemotype from Iran (Table S2). Both *DBR2* cDNA sequences were isolated by RT-PCR from cDNA constructed from RNA extracted from *Artemisia annua* flowers isolated from either chemotype. For amplification of the *DBR2* sequence, primers 5+6 were used (Table S1), of which sequences were based on the published *DBR2* sequence from a HAP chemotype (Zhang *et al.*, 2008). These primer sets were also able to amplify a *DBR2* sequence from the LAP chemotype. The primers used introduce BamHI and NotI restriction sites which were used for cloning into ImpactVectorpIV1A_2.1 (http://www.wageningenur.nl/ en/show/Productie-van-farmaceutische-en-industriele-eiwitten-door-planten.htm). The resulting pIV1A_2.1/DBR2 was digested with AscI and PacI and the full gene was cloned into the AscI and PacI sites of the pBinPlus binary vector (van Engelen *et al.*, 1995). *DBR2* sequences identified here have been deposited in the GenBank database (KC505370, JX898526 and JX898527).

Identification and cloning of ALDH1 from HAP and LAP chemotypes

ALDH1 sequences were isolated by RT-PCR on cDNA from floral RNA isolated from of the *Artemisia annua* HAP and LAP chemotype using primers 7+8 (Table S1). The sequences of these primers were based on the published *ALDH1* sequence from a HAP chemotype (Teoh *et al.*, 2009) and could amplify an *ALDH1* sequence from both HAP and LAP cDNA. Inspection of the coding sequence revealed that there were no specific amino acid residue differences in ALDH1 from HAP and LAP (both identical to *ALDH1* GenBank: FJ809784). The *ALDH1* cDNA was cloned into pBinPlus binary vector (van Engelen *et al.*, 1995).

Cloning of CYP71AV1 GFP reporter constructs

For the GFP fusion protein expression constructs (NtermAMOHAP:GFP, NtermAMOLAP:GFP, AMOHAP:GFP and AMOLAP:GFP): the N-terminal domains of *AMOHAP* (first 43 codons) and *AMOLAP* (first 50 codons) were amplified using primers 9+10 and 11+12, respectively. After digestion with BamHI and KpnI the fragments were cloned into C3.1/Kpn1 to form C3.1/NtermAMOHAP and C3.1/NtermAMOLAP.

The GFP coding sequence was amplified by PCR using primers 13+14 (Table S1) using pBin-Egfp as template (http://www.wageningenur.nl/en/show/Productie-van-farmaceutische-en-industriele-eiwitten-door-planten.htm). After digesting with KpnI and NotI, GFP was subcloned into plasmid C3.1/NtermAMOHAP, C3.1/NtermAMOLAP, C3.1/AMOHAP and C3.1/AMOLAP to form 35S-NtermAMOHAP:GFP, 35S-NtermAMOLAP:GFP, 35S-AMOHAP:GFP and 35S-AMOLAP:GFP.

To remove the stop codon of *AMOHAP* and *AMOLAP* and fuse them in-frame to the ATG start-codon of the GFP coding sequence in the 35S-AMOHAP:GFP and 35S-AMOLAP:GFP constructs, the 3'ends of the *AMOLAP* and *AMOHAP* were PCR amplified using Phusion DNA polymerase (Finnzymes) by primers 15+16, and 17+18, respectively.

The PCR products and plasmids 35S-AMOHAP:GFP and 35S-AMOLAP:GFP were digested with KpnI and EcoRI, followed by gel purification. The PCR products were ligated into the plasmids and transformed into *Escherichia coli*. Positive colonies were analyzed and sequenced to confirm that inserts were correct.

Transient expression in leaves of *Nicotiana benthamiana*

Agro-infiltration for transient expression in leaves of *Nicotiana benthamiana* Domin was done as described (van Herpen *et al.*, 2010). Briefly, individual *Agrobacterium tumefaciens* strains with different expression constructs (or empty vector as control) were co-infiltrated into *Nicotiana benthamiana* leaves using a syringe without needle. After seven days of transient expression, leaves were harvested for chemical analysis. In each set of experiments the total dosage of *Agrobacterium tumefaciens* between treatments was the same by diluting with *Agrobacterium tumefaciens* with empty vector where necessary. We note that leaves infiltrated with *AmFH+AMOLAP* developed necrotic lesions, indicating that at this time certain compounds started to accumulate to toxic levels in the infiltrated leaves.

CYP71AV1 subcellular localization studies

For subcellular localization of the AMO:GFP fusion proteins, *Arabidopsis thaliana* (L.) Heynh protoplasts were isolated and transfected with expression constructs 35S-NtermAMOHAP:GFP, 35S-NtermAMOLAP:GFP, 35S-AMOHAP:GFP and 35S-AMOLAP:GFP, based on a published protocol (Yoo *et al.*, 2007). As reference for ER subcellular localization the ER-YFP construct was used (Aker *et al.*, 2006). After transfection, protoplasts were analyzed by a Carl-Zeiss Confocal Scanning Laser Microscopy, with excitation of GFP at 488 nm and YFP at 514 nm. The fluorescence was detected via a band pass filter (GFP: 505-530 nm, YFP: 535-590 nm). Chlorophyll was detected using a 650 nm long pass filter.

Analysis of non-volatile metabolites by LC-QTOF-MS/MS

Seven days after agro-infiltration of *Nicotiana benthamiana*, the infiltrated leaves were harvested and immediately frozen in liquid nitrogen and ground to a fine powder. From each infiltrated leaf, 100 mg of powder was extracted in 300 µl methanol: formic acid (1000:1, v/v). Non-volatile compounds from the infiltrated leaves were analyzed by LC-QTOF-MS as described (van Herpen *et al.*, 2010).

Data were processed using the protocol for untargeted metabolomics of plant tissues as described (De Vos *et al.*, 2007, van Herpen *et al.*, 2010). Briefly, LC-QTOF-MS data were analyzed using Masslynx 4.0 (Waters) and processed using MetAlign version 1.0 (www.metAlign.nl) for base-

line correction, noise elimination and subsequent spectral data alignment (De Vos *et al.*, 2007). The processing parameters of MetAlign for LC-QTOF-MS data were set to analyze from scan numbers 60-2590 (corresponding to retention time 1.15-49.16 min) with a maximum amplitude of 25,000. After MetAlign processing, masses were clustered using the Multivariate Mass Spectra Reconstruction (MMSR) approach (Tikunov *et al.*, 2005) to elucidate which mass signals originate from the same metabolite. The mass signal intensity differences between treatments were compared using the student's t-test. Mass-directed LC-QTOF-MS/MS analysis for further elucidation of metabolite identities was done on differential compounds with signal intensities higher than 500 ion counts per scan.

Quantification of artemisinin precursors by UPLC-MRM-MS

Targeted analysis of artemisinin precursors in agro-infiltrated *Nicotiana benthamiana* leaves was performed with a Waters Xevo tandem quadrupole mass spectrometer equipped with an electrospray ionization source and coupled to an Acquity UPLC system (Waters) as described (Kohlen *et al.*, 2011) with some modifications. For details and instrument settings see supplemental data (Methods S1).

Analysis of volatile metabolites by GC-MS

Extracts were analysed by GC-MS using a gas chromatograph (7890A; Agilent, Amstelveen, the Netherlands) equipped with a 30-m x 0.25-mm i.d., 0.25-mm film thickness column with 5-m guard column (Zebron ZB5-MS; Phenomenex, Utrecht, The Netherlands) and a mass selective detector (model 5965c, Agilent). The GC was programmed at an initial temperature of 80°C for 1min, with a ramp of 5°C min^{-1} to 235°C and then a ramp of 25 °C min^{-1} to 280°C with a final time of 5 min. The injection port temperature was 250°C, and the He inlet pressure was controlled with electronic pressure control to achieve a constant column flow of 1.0 mL min^{-1}. 1 µl of the extracts was injected in split mode with a split flow set at 9 ml min^{-1}. Scanning was performed from 45 to 450 atomic mass units.

Glycosidase treatment

Viscozyme L (Sigma) was used as glycosidase treatment to hydrolyze hexose-conjugated compounds for subsequent quantification us-

ing GC-MS. Hereto, 200 mg infiltrated leaf material from each treatment was incubated in 1 ml citrate phosphate buffer, pH 5.4 containing 200 µl of Viscozyme L as previously described (van Herpen *et al.*, 2010).

Glutathione conjugation assay

In vitro conjugation of metabolites to glutathione by glutathione transferase activity (GST) was performed as described (Liu *et al.*, 2011). In brief, glutathione (GSH) (150 mM) in 7 µl potassium phosphate buffer (100 mM; pH 6.5), and 30 mM of artemisinin precursor (AAA, AAOH, AA, DHAAOH, DHAAA and DHAA) in 7 µl ethanol were added to 200 µl potassium phosphate buffer (100 mM; pH 6.5). The reaction was initiated by adding 7 µl of glutathione transferase (GST) (1g L^{-1}, in 100 mM KH$_2$PO$_4$ potassium phosphate buffer; pH 6.5) into the mixture. The controls were complete assay mixtures without GST enzyme or either of the substrates. After 15 min incubation at room temperature, samples were cooled to -20°C until LC-QTOF-MS analysis.

Results

Comparison of artemisinin biosynthesis protein sequences from HAP and LAP chemotypes reveals only relevant differences for CYP71AV1

Because the difference between the HAP and LAP chemotypes must arise after the ADS step in the biosynthesis pathway, and because no expression differences were found for *ADS* genes in HAP and LAP chemotypes (Maes *et al.*, 2011), here we focussed on analysis of putative differences in biosynthesis genes downstream of *ADS* (e.g. *CYP71AV1, DBR2, ALDH1*) in search of an explanation for the two different chemotypes of *A. annua.* For this purpose *CYP71AV1, DBR2* and *ALDH1* were isolated from an *A. annua* HAP and LAP chemotypes and the encoded protein sequences compared.

CYP71AV1: Analysis of the different *CYP71AV1* sequences that have been deposited in GenBank shows the occurrence of two major types of *CYP71AV1*, encoding two proteins which differ by a seven amino acids extension at the N-terminus of the protein (Fig. 2). The long version of CYP71AV1 (which we refer to as AMOLAP) has been isolated from *A. annua* Tanzania (Sandeman seed), which is a LAP chemotype (Ro et al., 2006). The other long version of CYP71AV1 (which we refer to as AMOLAP.1) was isolated from a different *A. annua* LAP chemotype (Kim et al., 1992) (Soon-Un Kim,

personal communication). Two versions of the CYP71AV1 (here referred to as AMOHAP and AMOHAP.1) were cloned from two different HAP chemotypes (Bertea *et al.*, 2006, Teoh *et al.*, 2006). The alignment of the AMOHAP and AMOLAP variants shows that none of the other single amino acid substitutions between the different AMOLAP and AMOHAP sequences are specific to the long or the short version of CYP71AV1 and therefore likely do not play a role in determining the LAP or HAP chemotypes (Fig. 2). RACE-PCR was used to analyse multiple *CYP71AV1* 5'sequences amplified from RNA isolated from a HAP chemotype and only *AMOHAP* 5'sequences were found (Fig. S1). Variation in the 5'untranslated region (nt 56-60) could be an indication that two different alleles of *AMOHAP* are present in this chemotype, both translating into the short AMOHAP. The RACE sequence data were consistent with the recently published sequence of the *CYP71AV1* promoter cloned from an *A. annua* HAP chemotype (Wang *et al.*, 2011).

DBR2: We cloned *DBR2* from the *A. annua* HAP chemotype (here referred to as *DBR2HAP.1*) and the sequence we obtained was similar to the recently described *DBR2* (here referred to as *DBR2HAP*) (Zhang *et al.*, 2008),

```
                GenBank No.
AMOLAP     ABB82944  MKSILKAMAL SLTTSIALAT ILLFVYKFAT RSKSTKKSLP EPWRLPIIGH MHHLIGTTPH RGVRDLARKY GSLMHLQLGE VPTIVVSSPK WAKEILTTYD 100
AMOLAP.1   ACF74516  MKSILKAMAL SLTTSIALAT ILLFVYEFAT RSKSTKKSLP EPWRLPIIGH MHHLIGTTPH RGVRDLARKY GSLMHLQLGE VPTIVVSSPK WAKEILTTYD 100
           ABI31728  MKSILKAMAL SLTTSIALAT ILLFVYKFAT RSKSTKKSLP EPWRLPIIGH MHHLIGTTPH RGVRDLARKY GSLMHLQLGE VPTIVVSSPK WAKEVLTTYD 100
AMOHAP     AFP19100  -------MAL SLTTSIALAT ILLFVYKFAT RSKSTKKSLP EPWRLPIIGH MHHLIGTTPH RGVRDLARKY GSLMHLQLGE VPTIVVSSPK WAKEILTTYD 93
AMOHAP.1   ABC41927  -------MAL SLTTSIALAT ILLFVYKFAT RSKSTKKSLP EPWRLPIIGH MHHLIGTTPH RGVRDLARKY GSLMHLQLGE VPTIVVSSPK WAKEILTTYD 93
           ABE57266  -------MAL SLTTSIALAT ILLFVYKFAT RSKSTKKSLP EPWRLPIIGH MHHLIGTTPH RGVRDLARKY GSLMHLQLGE VPTIVVSSPK WAKEILTTYD 93
           ABM88788  -------MAL SLTTSIALAT ILLFVYKFAT RSKSTKKSLP EPWRLPIIGH MHHLIGTTPH RGVRDLARKY GSLMHLQLGE VPTIVVSSPK WAKEILTTYD 93
           ABG49366  -------MAL SLTTSIALAT ILFTIPH... RSKSNKNSLP EPWRLPIIGH MHHLIGTIPH RGVMDLARKY GSLMHLQLGE VSTIVVSSPK WAKEILTTYD 93

AMOLAP     ABB82944  ITFANRPETL TGEIVLYHNT DVVLAPYGEY WRQLRKICTL ELLSVKKVKS FQSLREEECW NLVQEIKASG SGRPVNLSEN VFKLIATILS RAAFGKGIKD 200
AMOLAP.1   ACF74516  ISFANRPETL TGEIVLYHNT DVVLAPYGEY WRQLRKICTS ELLSVKKVKS FQSLREEECW NLVQEIKASG SGRPVNLSEN VFKLIATILS RAAFGKGIKD 200
           ABI31728  ITFANRPETL TGEIVLYHNT DVVLAPYGEY WRQLRKICTL ELLSVKKVKS FQSLREEECW NLVQEIKASG SGRPVNLSEN VFKLIATILS RAAFGKGIKD 200
AMOHAP     AFP19100  ITFANRPETL TGEIVLYHNT DVVLAPYGEY WRQLRKICTL ELLSVKKVKS FQSLREEECW NLVQEIKASG SGRPVNLSEN VFKLIATILS RAAFGKGIKD 193
AMOHAP.1   ABC41927  ITFANRPETL TGEIVLYHNT DVVLAPYGEY WRQLRKICTL ELLSVKKVKS FQSLREEECW NLVQEIKASG SGRPVNLSEN IFKLIATILS RAAFGKGIKD 193
           ABE57266  ITFANRPETL TGEIVLYHNT DVVLAPYGEY WRQLRKICTL ELLSVKKVKS FQSLREEECW NLVQEIKASG SGRPVNLSEN VFKLIATILS RAAFGKGIKD 193
           ABM88788  ISFANRPETL TGEIVLYHNT DVVLAPYGEY WRQLRKICTL ELLSVKKVKS FQSLREEECW NLVQEIKASG SGRPVNLSEN VFKLIATILS RAAFGKGIKD 193
           ABG49366  ITFANRPETL TGEIVAYHNT DIVLAPYGEY WRQLRKLCTL ELLSVKKVKS FQSLREEECW NLVQEIKAAG SGRPVNLSEN IFKLIATILS RAAFGKGIKD 193

AMOLAP     ABB82944  QKELTEIVKE ILRQTGGFDV ADIFPSKKFL HHLSGKRARL TSLRKKIDNL IDNLVAEHTV NTSSKTNETL LDVLLRLKDS AEFPLTSDNI KAIILDMFGA 300
AMOLAP.1   ACF74516  QKELTEIVKE ILRQTGGFDV ADIFPSKKFL HHLSGKRARL TSLRKKIDNL IDNLVAEHTV NTSSKTNETL LDVLLRLKDS AEFPLTSDNI KAIILDMFGA 300
           ABI31728  QKELTEIVKE ILRQTGGFDV ADIFPSKKFL HHLSGKRARL TSLRKKIDNL IDNLVAEHTV NTSSKTNETL LDVLLRLKDS AEFPLTSDNI KAIILDMFGA 300
AMOHAP     AFP19100  QKELTEIVKE ILRQTGGFDV ADIFPSKKFL HHLSGKRARL TSLRKKIDNL IDNLVAEHTV NTSSKTNETL LDVLLRLKDS AEFPLTSDNI KAIILDMFGA 293
AMOHAP.1   ABC41927  QKELTEIVKE ILRQTGGFDV ADIFPSKKFL HHLSGKRARL TSLRKKIDNL IDNLVAEHTV NTSSKTNETL LDVLLRLKDS AEFPLTSDNI KAIILDMFGA 293
           ABE57266  QKELTEIVKE ILRQTGGFDV ADIFPSKKFL HHLSGKRARL TSLRKKIDNL IDNLVAEHTV NTSSKTNETL LDVLLRLKDS AEFPLTSDNI KAIILDMFGA 293
           ABM88788  QKELTEIVKE ILRQTGGFDV ADIFPSKKFL HHLSGKRARL TSLRKKIDNL IDNLVAEHTV NTSSKTNETL LDVLLRLKDS AEFPLTSDNI KAIILDMFGA 293
           ABG49366  QKEPTEIVKE ILRQTGGFDV ADIFPSKKFL HHLSGKRARL TSIHMKLDNL IDNLVAEHTV NTSSKTNETL LDVLLRLKDS AEFPLTADNV KAIILDMFGA 293

AMOLAP     ABB82944  GTDTSSSTIE WAISELIKCP KAMEKVQAEL RKALNGKEKI HEEDIQELSY LNMVIKETLR LHPPLPLVLP RECRQPVNLA GYNIPNKTKL IVNVFAINRD 400
AMOLAP.1   ACF74516  GTDTSSSTIE WAISELIKCP KAMEKVQAEL RKALNGKEKI HEEDIQELSY LNMVIKETLR LHPPLPLVLP RECRQPVNLA GYNIPNKTKL IVNVFAINRD 400
           ABI31728  GTDTSSSTIE WAIPELIKCP KAMEKVQAEL RKALNGKEKI HEEDIQELSY LNMVIKETLR LHPPLPLVLP RECRQPVNLA GYNIPNKTKL IVNVFAINRD 400
AMOHAP     AFP19100  GTDTSSSTIE WAISELIKCP KAMEKVQAEL RKALNGKEKI HEEDIQELSY LNMVIKETLR LHPPLPLVLP RECRQPVNLA GYNIPNKTKL IVNVFAINRD 393
AMOHAP.1   ABC41927  GTDTSSSTIE WAISELIKCP KAMEKVQAEL RKALNGKEKI HEEDIQELSY LNMVIKETLR LHPPLPLVLP RECRQPVNLA GYNIPNKTKL IVNVFAINRD 393
           ABE57266  GTDTSSSTIE WAISELIKCP KAMEKVQAEL RKALNGKEKI HEEDIQELSY LNMVIKETLR LHPPLPLVLP RECRQPVNLA GYNIPNKTKL IVNVFAINRD 393
           ABM88788  GTDTSSSTIE WAISELIKCP KAMEKVQAEL RKALNGKEKI HEEDIQELSY LNMVIKETLR LHPPLPLVLP RECRQPVNLA GYNIPNKTKL IVNVFAINRD 393
           ABG49366  GTDTSSATIE WAISELIKCP RAMEKVQAEL RKALKGKEKI HEEDIQELSY LNLVIKETLR LHPPLPLVTP RECRQPVNLA GYNIPNKTKL IVNVFAINRD 393

AMOLAP     ABB82944  PEYWKDAEAF IPERFENSSA TVMGAEYEYL PFGAGRRMCP GAALGLANVQ LPLANILYHF NWKLPNGVSY DQIDMTESSG ATMQRKTELL LVPSF 495
AMOLAP.1   ACF74516  PEYWKDAEAF IPERFENSSA TVMGAEYEYL PFGAGRRMCP GAALGLANVQ LPLANILYHF NWKLPNGVSY DQIDMTESSG ATMQRKTELL LVPSF 495
           ABI31728  PEYWKDAEAF IPERFENSSA TVMGAEYEYL PFGAGRRMCP GAALGLANVQ LPLANILYHF NWKLPNGVSY DQIDMTESSG ATMQRKAELL LVPSF 495
AMOHAP     AFP19100  PEYWKDAEAF IPERFENSSA TVMGAEYEYL PFGAGRRMCP GAALGLANVQ LPLANILYHF NWKLPNGVSY DQIDMTESSG ATMQRKTELL LVPSF 488
AMOHAP.1   ABC41927  PEYWKDAEAF IPERFENSSA TVMGAEYEYL PFGAGRRMCP GAALGLANVQ LPLANILYHF NWKLPNGVSY DQIDMTESSG ATMQRKTELL LVPSF 488
           ABE57266  PEYWKDAEAF IPERFENSSA TVMGAEYEYL PFGAGRRMCP GAALGLANVQ LPLANILYHF NWKLPNGVSY DQIDMTESSG ATMQRKTELL LVPSF 488
           ABM88788  PEYWKDAEAF IPERFENSSA TVMGAEYEYL PFGAGRRMCP GAALGLANVQ LPLANILYHF NWKLPNDVSY DQIDMTESSG ATMQRKTELL LVPSF 488
           ABG49366  PEYWKDAETF IPERFENSST TVMGAEYEYL PFGAGRRMCP GSALGLANVQ LPLANILYHF NWKLPNGASY DQIDMAESFG ATVQRKTELL LVPSF 488
```

Figure 2. Alignment of the AMOLAP and AMOHAP deduced amino acid sequences. All amino acid sequences used in the alignment were retrieved from GenBank (GenBank number given in front of each sequence). Two of the sequences were from confirmed LAP chemotypes: ABB82944, referred to as AMOLAP (Ro *et al.*, 2006) and ACF74516, referred to as AMOLAP.1 (Kim *et al.*, 1992). Two of the sequences were from confirmed HAP chemotypes: AFP19100, referred to as AMOHAP (Bertea *et al.*, 2006) and ABC41927, referred to as AMOHAP.1 (Teoh *et al.*, 2006). For the other sequences deposited in GenBank the chemotype of the plant from which they were isolated was not given.

with the exception of a one amino acid difference (Fig. 3). Using primers based on *DBR2HAP* we isolated two variants of *DBR2* from an *A. annua* LAP chemotype (see methods). Five of the eleven clones showed few amino acid differences with the published *DBR2HAP* (here referred to as *DBR2LAP.1*). However, in the remaining six clones (here referred to as *DBR2LAP*), the encoded protein sequence showed a number of amino acid residue differences with DBR2HAP, including two additional amino acids in position 295 (Fig. 3).

ALDH1: Cloning of *ALDH1* from *A. annua* has been described (Teoh *et al.*, 2009). We isolated *ALDH1* from both the HAP and LAP chemotypes. Alignment of the AA-sequence showed no differences between the two proteins (data not shown).

Both AMOHAP and AMOLAP are localized to the Endoplasmic Reticulum

Since AMOLAP and AMOHAP only consistently differ in their N-terminal amino acid residues which supposedly encode the ER anchoring domain (Fig. 2), we investigated whether this difference causes altered subcellular targeting or difference in protein stability. Expression constructs encoding either full-length protein-GFP fusions or truncated N-terminal domain-GFP fusions

```
LAP   DBR2LAP    MSEKPTLFSA YKMGNFNLSH RVVLAPMTRC RAINAIPNEA LVEYYRQRST AGGFLITEGT MISPSSAGFP HVPGIFTKEQ VEGWKKVVDA AHKEGAVIFC 100
HAP   DBR2HAP    MSEKPTLFSA YKMGNFNLSH RVVLAPMTRC RAINAIPNEA LVEYYRQRST AGGFLITEGT MISPSSAGFP HVPGIFTKEQ VEGWKKVVDA AHKEGAVIFC 100
HAP   DBR2HAP.1  MSEKPTLFSA YKMGKFNLSH RVVLAPMTRC RAINAIPNEA LVEYYRQRST AGGFLITEGT MISPSSAGFP HVPGIFTKEQ VEGWKKVVDA AHKEGAVIFC 100

LAP   DBR2LAP    QLWHVGRASH KVYQPGGAAP ISSTSKPISK KWEILLPDAT YGTYPEPRPL AANEILEVVE DYRVAAINAI EAGFDGIEIH GAHGYLLDQF MKDGINDRTD 200
HAP   DBR2HAP    QLWHVGRASH QVYQPGGAAP ISSTSKPISK KWEILLPDAT YGTYPEPRPL AANEILEVVE DYRVAAINAI EAGFDGIEIH GAHGYLLDQF MKDGINDRTD 200
HAP   DBR2HAP.1  QLWHVGRASH QVYQPGGAAP ISSTSKPISK KWEILLPDAT YGTYPEPRPL AANEILEVVE DYRVAAINAI EAGFDGIEIH GAHGYLLDQF MKDGINDRTD 200

LAP   DBR2LAP    EYGGSLENRC KFILQVVQAV SAAIGTDRVG IRISPAIDHT DAMDSDPRSL GLAVIERLNK LQFKLGSRLA YLHVTQPRYT ADGHGQTEAG ANGSDHEEEV 300
HAP   DBR2HAP    EYGGSLENRC KFILQVVQAV SAAIGTDRVG IRISPAIDHT DAMDSDPRSL GLAVIERLNK LQFKLGSRLA YLHVTQPRYT ADGHGQTEAG ANGS--EEEV 298
HAP   DBR2HAP.1  EYGGSLENRC KFILQVVQAV SAAIGTDRVG IRISPAIDHT DAMDSDPRSL GLAVIERLNK LQFKLGSRLA YLHVTQPRYT ADGHGQTEAG ANGS--EEEV 298

LAP   DBR2LAP    AQLMKTWRGA YVGTFICCGG YTRELGLQAV AQGDADLVAF GRYFISNPDL VLRLKLNAPL NRYDRATFYT HDPVVGYTDY PSLDQGSLL 389
HAP   DBR2HAP    AQLMKTWRGA YVGTFICCGG YTRELGLQAV AQGDADLVAF GRYFVSNPDL VLRLKLNAPL NRYDRATFYT HDPVVGYTDY PSLDKGSLL 387
HAP   DBR2HAP.1  AQLMKTWRGA YVGTFICCGG YTRELGLQAV AQGDADLVAF GRYFVSNPDL VLRLKLNAPL NRYDRATFYT HDPVVGYTDY PSLDKGSLL 387
```

Figure 3. Alignment of the DBR2LAP and DBR2HAP deduced amino acid sequences. All amino acid sequences used in the alignment were retrieved from GenBank: DBR2LAP (JX898527), DBR2LAP.1 (KC505370), DBR2HAP (ACH61780) and DBR2HAP.1 (JX898526). DBR2LAP and DBR2LAP.1 were cloned from LAP, DBR2HAP was cloned from HAP (Zhang *et al.*, 2008) and DBR2HAP.1 was cloned from HAP (Bertea *et al.*, 2006). Alignment shows that one variant of DBR2LAP.1 is similar to DBR2HAP with only two AA difference. The second variant of the DBR2LAP has a two amino acids insertion at the C-terminus. No difference was detected in *in planta* activity of DBR2HAP and DBR2LAP (Figs 5b, 5c, S3b).

of AMOLAP and AMOHAP were transiently expressed in *A. thaliana* protoplasts. Confocal microscopy of *Arabidopsis* protoplasts co-transfected with the full length AMOLAP protein fused to GFP (AMOLAP:GFP) showed co-localization of the GFP fluorescence signal with the fluorescence signal of the ER marker (ER:YFP) (Fig. 4a). Similarly, full length AMOHAP protein fused to

GFP (AMOHAP:GFP) also showed co-localization with the ER marker (ER:Y-FP) (Fig. 4c). In addition, the truncated N-terminal portion of AMOLAP and AMOHAP were fused to GFP. When transfected into *Arabidopsis* protoplasts the NtermAMOLAP:GFP and NtermAMOHAP:GFP both showed co-localization with the ER:YFP ER marker (Figs. 4c,d). Localization experiments with 35S expression constructs may lead to artefacts such as cytosolic localization when ER import is saturated, however this was not observed in these experiments. Combined, the results demonstrate that AMOLAP and AMOHAP do not differ in subcellular targeting, as both proteins localize to the ER. Although the fluorescence signal varies between transfection assays, there were indications that AMOLAP may be more stable. For instance we do find a higher fluorescence signal for NtermAMOLAP:GFP than for the NtermAMOHAP:GFP.
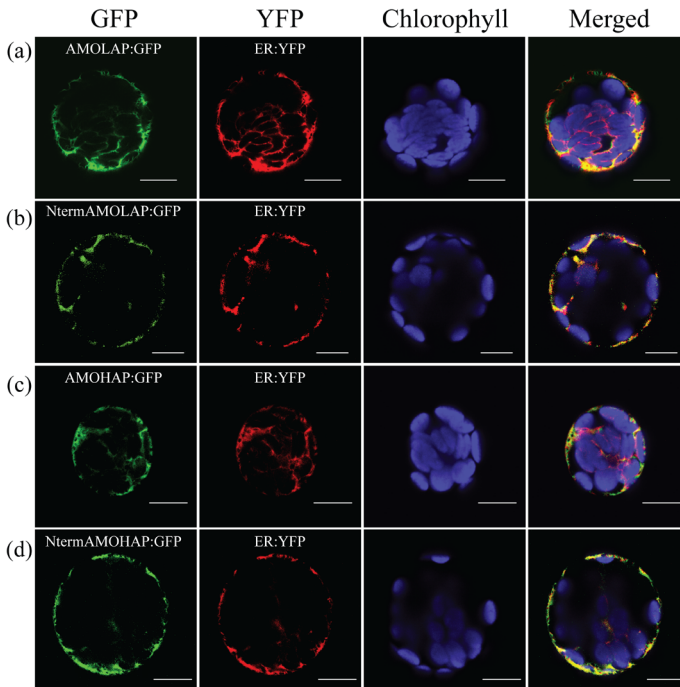
Figure 4. Subcellular localization of AMOLAP and AMOHAP. Confocal microscopy analysis of *Arabidopsis* protoplasts co-transfected with (a) AMOLAP:GFP + ER:YFP; (b) NtermAMOLAP:GFP + ER:YFP; (c) AMOHAP:GFP + ER:YFP; (d) NtermAMOHAP:GFP + ER:YFP. Artificial colors were given to GFP fluorescence (green), YFP fluorescence (red), and auto fluorescence of chloroplasts (purple). Merging of the pictures results in a yellow color for GFP-YFP overlap. The scale bar = 10 µm.

Different product profiles in planta from AMOLAP and AMOHAP

Both the AMOLAP and the AMOHAP enzymes have been characterised in a yeast expression system and both were shown to be able to produce AAOH, AAA and AA from AD (Ro *et al.*, 2006, Teoh *et al.*, 2006). However, a direct comparison of variants AMOLAP and AMOHAP in the same expression system has not been performed until now. The different *in planta* expression studies using heterologous plant hosts (tobacco and *N. benthamiana*) have only been reported for the longer *AMOLAP* (van Herpen *et al.*, 2010, Zhang *et al.*, 2011). To test the effect of the seven AA extension of the AMOLAP protein we made two expression constructs, both based on the *AMOHAP* sequence but in one construct we introduced the seven AA extension to the protein sequence as found in AMOLAP, thus limiting the difference between the two forms to the N-terminal extension. The activity of these AMOLAP and AMOHAP genes was subsequently compared *in planta* by co-expression with *ADS*, using transient expression in *N. benthamiana* leaves. To achieve high levels of artemisinin precursor production, *ADS* was expressed with a mito-chondrial targeting signal, and overexpression was combined with a mito-chondrial targeted *FPS* and a truncated, cytosolic form of *HMGR*. *ADS*, *FPS* and *HMGR* were combined into a single 2A expression construct (*AmFH*) as described before (van Herpen *et al.*, 2010). Each expression construct (*AmFH* and *AMOLAP* or *AMOHAP*) was introduced into *A. tumefaciens* and *N. benthamiana* leaves were infiltrated with *AmFH+AMOLAP* or *AmFH+AMOHAP*.

Particularly with *AmFH+AMOLAP*, infiltrated leaves developed symptoms of necrosis around seven days post infiltration (Fig. S2), suggesting the production of a toxic compound. Necrosis symptoms were stronger in leaves infiltrated with *AmFH+AMOLAP* than with *AmFH+AMOHAP*, suggesting that the products of both treatments may not be the same. Because necrosis started to appear after seven days, in all our experiments the leaves were harvested at day seven instead of day ten after infiltration, as previously done (van Herpen *et al.*, 2010).

Analysis of free products: To quantify the products from the ADS and AMOLAP/HAP transient enzyme activity in the infiltrated *N. benthamiana*, leaves were extracted with aqueous methanol for UPLC-MRM-MS analysis. Intriguingly, the distribution over the entire product range was different in leaves infiltrated with *AmFH+AMOLAP* and *AmFH+AMOHAP* (Table 1). Leaves infiltrated with *AmFH+AMOLAP* produced predominantly AA, while in leaves infiltrated with *AmFH+AMOHAP* AA levels were 50-fold lower. However, leaves infiltrated with *AmFH+AMOHAP* contained 3-fold higher levels of AAOH and AAA than leaves expressing *AmFH+AMOLAP* (Table 1).

Also DHAAOH, DHAAA and DHAA were detected in the *AmFH+AMOLAP* leaf samples (Table 1), suggesting the presence of an endogenous *N. benthamiana* enzyme with carbon double bond reducing activity (catalysing the conversion of AAA to DHAAA just as DBR2 in *A. annua*), and enzymes similar to *A. annua* RED1 (catalysing the formation of DHAAOH from DHAAA) and *A. annua* ALDH1 (catalysing the conversion of DHAAA to DHAA) (Fig. 1). Free DHAAOH levels were higher in *AmFH+AMOLAP* compared to *AmFH+AMOHAP* infiltrated leaves, suggesting that formation of DHAAOH is not directly related to the free DHAAA, AAOH or AAA levels in leaves, which were lower in *AmFH+AMOLAP*. DHAA was only detected in *AmFH+AMOLAP* infiltrated leaves.

Table 1. Unconjugated artemisinin precursors produced in *Nicotiana benthamiana* as identified and quantified by UPLC-MRM-MS

(ng g^{-1} FW)	AmFH+AMO-LAP	AmFH+AMO-HAP	AmFH+AMO-LAP+DBR2	AmFH+AMO-HAP+DBR2
AAOH	16709 ± 3977	46946 ± 5692	3596 ± 1247	16554 ± 3233
AAA	5501 ± 1486	19564 ± 6314	781 ± 246	3418 ± 890
AA	3969 ± 1391	74 ± 15	53 ± 16	ND
DHAAOH	1821 ± 576	597 ± 19	87972 ± 15014	57289 ± 9455
DHAAA a	(5966 ± 1646)	(8067 ± 1341)	(220347 ± 65373)	(124787±26145)
DHAA	17 ± 8	ND	838 ± 517	25 ± 9

AAOH: artemisinic alcohol; AAA: artemisinic aldehyde; AA: artemisinic acid; DHAAOH: dihydroartemisinic alcohol; DHAAA: dihydroartemisinic aldehyde; DHAA: dihydroartemisinic acid. ND: not detectable. [a] The values for DHAAA are shown in brackets as they represent peak intensities and not concentrations.

Analysis of glycosylated products: Previous results showed that most of the products of the ADS and AMOLAP activity in *N. benthamiana* agro-infiltration are present as glycosylated conjugates, mainly of AA (van Herpen *et al.*, 2010). Therefore, leaf material was also analysed by LC-QTOF-MS. *N. benthamiana* leaves co-expressing *AmFH+AMOLAP* indeed contain AA-12-β-di-glucoside, as previously reported (van Herpen *et al.*, 2010). However, in addition several other AA-glycoside conjugates were detected, including conjugates with additional hexose units as well as malonylated hexoses. Also for the glycosylated products the distribution over the entire product range was different between *AmFH+AMOLAP* and *AmFH+AMOHAP* (Fig. 5a, Table 2 (Page 80)). Leaves infiltrated with *AmFH+AMOLAP* produced more AA conjugates, while leaves infiltrated with *AmFH+AMOHAP* produced more AAOH conjugates (Table 2, Fig. S3a). For both treatments also several DHAAOH

and DHAA conjugates with hexose and malonyl groups were detected, but no DHAAA conjugates (Table 2). Table 2 shows the mass fragmentation profiles of the detected products and their putative identification. MS/MS analysis was used to further confirm product identity and an example of the identification of one of the DHAA-hexose conjugates is shown in Fig. S5.

To quantify the levels of glycosylated products, samples were treated with a mix of glycosidases (Viscozyme L) and deglycosylated products were quantified using GC-MS. Note that the Viscozyme treatment only cleaves hexose conjugates but not malonylated hexose conjugates (Fig. S4). Results show that leaves infiltrated with *AmFH+AMOLAP* contained *c.* 40 mg.kg^{-1} FW of AA [consistent with the previously reported 39.5 mg.kg^{-1} FW of AA (van Herpen *et al.*, 2010)], while the sensitivity of the GC-MS was not sufficient to detect any AA in leaves infiltrated with *AmFH+AMOHAP* (Table 3). GC-MS analysis after Viscozyme treatment confirmed that AAOH was the major glycosylated product in leaves infiltrated with *AmFH+AMOHAP* (as was suggested by Table 2) at 24 mg.kg^{-1} FW.

Table 3. Artemisinin precursors in *Nicotiana benthamiana* agro-infiltrated with artemisinin biosynthetic pathway genes.

(mg kg^{-1} FW)	AmFH+AMO-LAP	AmFH+AMO-HAP	AmFH+AMO-LAP+DBR2	AmFH+AMO-HAP+DBR2
AAOH	8.1 ± 1.6	24.0 ± 3.5	5.1 ± 0.5	9.4 ± 2.1
AAA	1.6 ± 0.1	1.6 ± 0.1	ND	ND
AA	39.9 ± 9.8	ND	ND	ND
DHAAOH	1.6 ± 0.2	2.0 ± 0.2	42.9 ± 14.9	22.8 ± 8.6
DHAAA	ND	ND	4.0 ± 2.3	1.3 ± 0.6
DHAA	ND	ND	7.3 ± 2.2	ND

Agro-infiltrated leaves were treated with glycosidase (Viscozyme L.) and hydrolysed metabolites extracted and analysed by GC-MS. AAOH: artemisinic alcohol; AAA: artemisinic aldehyde; AA: artemisinic acid; DHAAOH: dihydroartemisinic alcohol; DHAAA: dihydroartemisinic aldehyde; DHAA: dihydroartemisinic acid; ND: not detectable. Results are means ± S.D. of three co-infiltrated leaves.

No difference in DBR2 activity from HAP and LAP *A. annua* chemotypes

We compared the activity of the two variants of DBR2 by comparing product profiles of *N. benthamiana* leaves infiltrated with *AmFH+AMOLAP* in combination with either *DBR2HAP* or *DBR2LAP*. In addition we tested the

two DBR2 variants in combination with *AmFH+AMOHAP*. Analysis of the conjugated products show that there is no difference in product profile between *DBR2HAP* and *DBR2LAP* (Fig. 5b, S3b, Table S3), indicating that
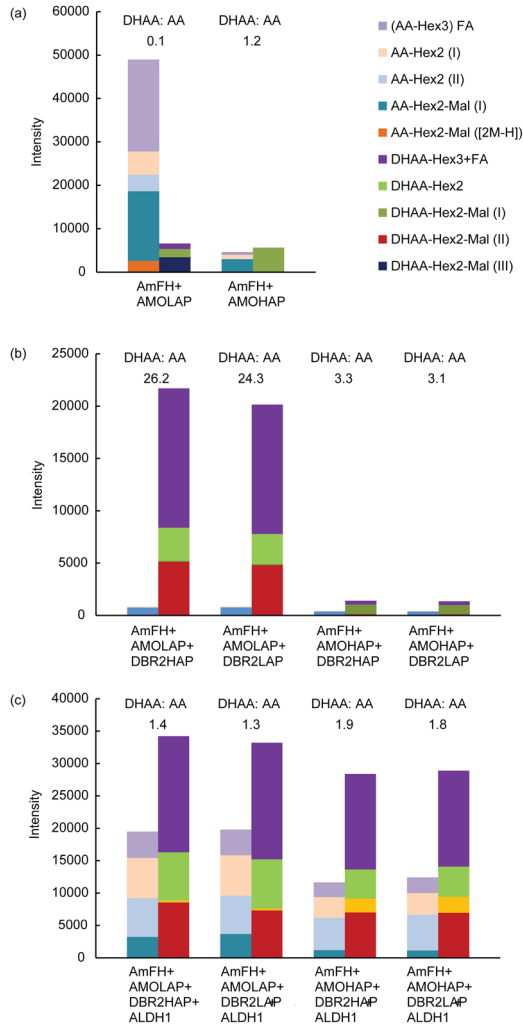

Figure 5. Artemisinic acid (AA) and dihydroartemisinic acid (DHAA) related compounds in leaves of *N. benthamiana* agro-infiltrated with artemisinin biosynthesis genes as identified by LC-QTOF-MS. (a) Agro-infiltrated leaves with *AmFH+AMOLAP* and *AmFH+AMOHAP*. For all products identified by LC-QTOF-MS see Table 2. (b) Agro-infiltrated leaves with *AmFH+AMOLAP/HAP+DBR2LAP/HAP*. For all products identified by LC-QTOF-MS see Table S3. (c) Agro-infiltrated leaves with *AmFH+AMOLAP/HAP+DBR2LAP/HAP+ALDH1*. For all products identified by LC-QTOF-MS see Table S3. Data represent peak intensities for each of the compounds in LC-QTOF-MS analysis. Peak intensities are the mean of three agro-infiltrated leaves.

the two forms of DBR2 do not differ in enzymatic activity. The co-infiltration with *DBR2* relieved the necrosis symptoms caused by expression of *AmFH+AMOLAP* or *AmFH+AMOHAP* alone (Fig. S2), suggesting that additional DBR2 enzyme activity lowered the level of the product(s) that cause necrosis. Product analysis in leaves agro-infiltrated with *AmFH+AMOLAP+DBR2* or *AmFH+AMOHAP+DBR2* showed that DBR2 activity resulted in a significant increase in DHAAOH, DHAAA and DHAA levels (Table 1) and this is also clear from LC-QTOF-MS analysis that shows a strong increase in DHAAOH and DHAA conjugates to hexose and malonyl groups (Table 2).

The analysis of deglycosylated extracts by GC-MS confirmed that co-expression of *DBR2* increased the levels of DHAAOH, DHAAA and DHAA at the expense of AAOH, AAA and AA levels (Table 3). The total yield of DHAA in leaves agro-infiltrated with *AmFH+AMOLAP+DBR2* as released by glycosidase treatment was *c.* 7.3 mg.kg^{-1} FW while DHAA in leaves agro-infiltrated with *AmFH+AMOHAP+DBR2* was below the level of detection by GC-MS. Combined, these results show that DBR2 further enhances the double bond reduction of the CYP71AV1 products that is also already catalysed by endogenous tobacco reductase activity. In addition, endogenous tobacco glycosyl and malonyl transferases modify these double-bond-reduced products which leads to DHAAOH and DHAA conjugates.

Increased DHAA and AA by combining ADS, AMO and DBR2 with ALDH1

As described above, no differences were found between the ALDH1 protein sequence from LAP and HAP *A. annua* chemotypes. To test how the addition of *ALDH1* activity affects the product profile of the artemisinin HAP and LAP biosynthesis pathway, leaves were agro-infiltrated with *AmFH+AMOLAP+DBR2+ALDH1* or *AmFH+AMOHAP+DBR2+ALDH1*. After seven days leaves were extracted and products were profiled by LC-QTOF-MS. The levels of conjugated DHAAOH products significantly decreased when *ALDH1* was added to *AmFH+AMOLAP+DBR2* (Table S3, Fig. S3b), coinciding with a substantial increase in glycosylated AA and DHAA product levels (Fig. 5c). This suggests that ALDH1 may be more efficient in the conversion of AAA to AA and DHAAA to DHAA than AMOLAP and AMOHAP as already suggested by the work of Teoh (Teoh *et al.*, 2009). Although the level of the presumed direct precursor of artemisinin (DHAA) was substantially increased by ALDH1 (*c.* 13-fold by adding *ALDH1* to *AmFH+AMOLAP+DBR2* and *c.* 110-fold by adding *ALDH1* to *AmFH+AMOHAP+DBR2*), no artemis-

inin could be detected in *N. benthamiana* by UPLC-MRM-MS (Table 1, 4).

Qualitative effects on product profile by AMOLAP dosage

The comparison of the total product levels produced by the reconstituted pathway with AMOLAP or AMOHAP suggests that AMOLAP has a higher enzyme activity than AMOHAP, in combination with a different product profile (Fig. 5a, Table 2). This difference could not be related to different subcellular localization (Fig. 4). To test if differences in relative enzyme activity within the pathway can affect the product profile in a qualitative way, we tested the effect of different dilutions of AMOLAP in combination with the rest of the biosynthesis pathway (*ADS+AMOLAP+DBR2HAP*). This was achieved by diluting the *Agrobacterium* strain carrying the *AMOLAP* expression construct with a suspension of an *Agrobacterium* strain carrying an empty expression vector to keep the total *Agrobacterium* dosage for infiltration the same. Results show that reduction of the AMOLAP agro-infiltration dosage to 1/2 had only little effect, but that dilution up to 1/10 strongly decreased DHAA-glycoside production (Figs. 6a, S3c; Table S4). The AA-glycoside levels were also reduced but to a much lower extent.

More AAA-related glutathione-conjugate from AMOHAP than from AMOLAP

Recently we described the reconstruction of the biosynthetic pathway of costunolide in *N. benthamiana* (Liu *et al.*, 2011) in which it was shown that the exocyclic carbon double bond of costunolide conjugates to glutathione (GSH). Since some of the products of ADS (AD) and AMOLAP/HAP enzyme activities also contain such an exocyclic double bond (AAA), but lack the hydroxyl or acid group used for glycosylation in AAOH and AA, we specifically looked for GSH conjugates of artemisinin intermediates in *N. benthamiana* leaves agro-infiltrated with the biosynthetic pathway genes. A putative GSH-conjugated compounds was detected by LC-QTOF-MS, both in negative mode (m/z=542.25) and positive mode (m/z=544.24) (Fig. S6). The level of the GSH-conjugate (m/z=542.25) was higher in *AmFH+AMOHAP* than in *AmFH+AMOHAP+DBR2* agro-infiltrated leaves (Table 2). We tested the artemisinin biosynthetic pathway intermediates (AAA, AAOH, DHAAOH, AA, DHAA and DHAAA) in *in vitro* reactions for spontaneous or glutathione-S-transferase (GST) driven GSH conjugation. Only AAA formed an AAA-GSH conjugate similar to that extracted from agro-infiltrated leaves express-
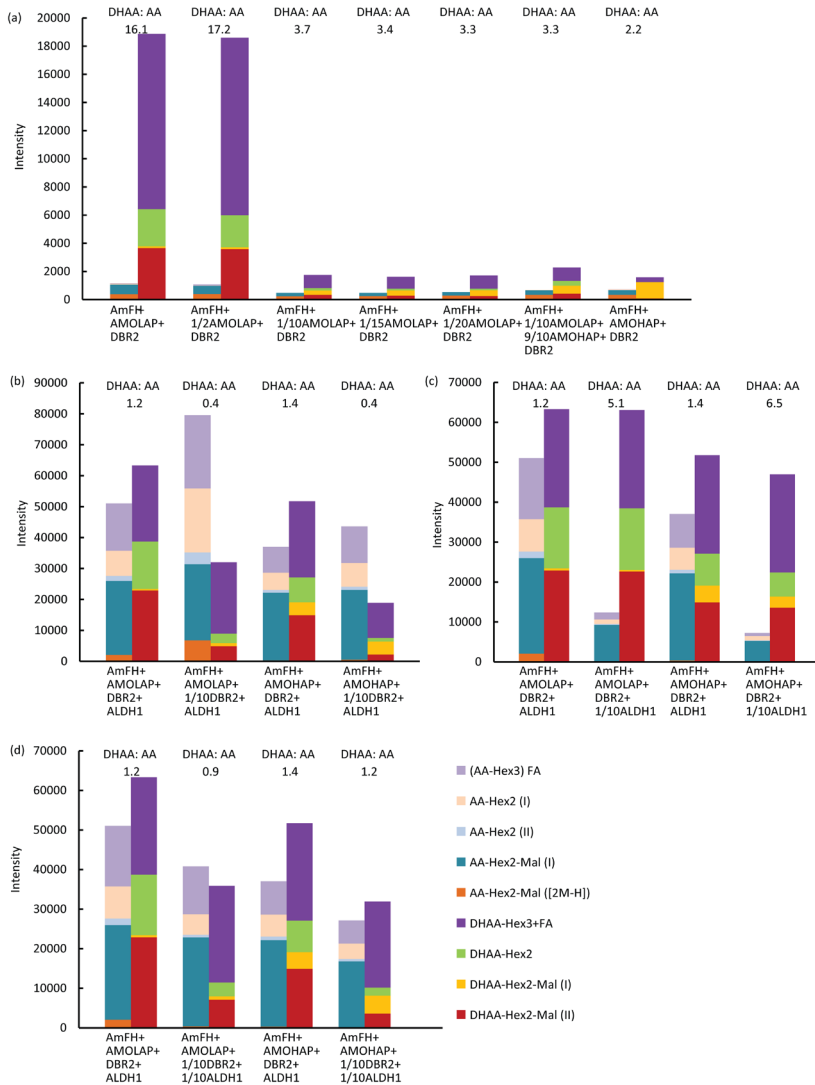
Figure 6. Artemisinic acid (AA) and dihydroartemisinic acid (DHAA) conjugated compounds in leaves of *N. benthamiana* agro-infiltrated with AMOLAP/HAP and different dosage of DBR2 and ALDH1. (a) Comparison AA and DHAA conjugated compounds in agro-infiltrated leaves with dilution of AMOLAP/HAP. For all products identified by LC-QTOF-MS see Table S4. (b) Comparison AA and DHAA conjugated compounds in agro-infiltrated leaves with dilution of DBR2. For all products identified by LC-QTOF-MS see Table S5. (c) Comparison AA and DHAA conjugated compounds in agro-infiltrated leaves with dilution of ALDH1. For all products identified by LC-QTOF-MS see Table S5. (d) Comparison AA and DHAA conjugated compounds in agro-infiltrated leaves with dilution of DBR2 and ALDH1. For all products identified by LC-QTOF-MS see Table S5. Data represent peak intensities for each of the compounds in LC-QTOF-MS analysis.

ing the pathway genes (*m/z*=526.22) (Fig. S6 and S7). The mass of the major GSH conjugate formed *in planta* is 18 D higher, suggesting an additional two protons and one oxygen atom, which could be explained by hydroxylation of the endocyclic double bond in AAA (Fig. S6). The level of the putative AAA glutathione conjugate was higher in leaves infiltrated with *AmFH+AMO-HAP* than in leaves infiltrated with *AmFH+AMOLAP* (Table 2). Dilution of AMOLAP resulted in an increase in the level of the AAA glutathione conjugate, although not reaching the level of the full dosage of AMOHAP (Fig. 7).
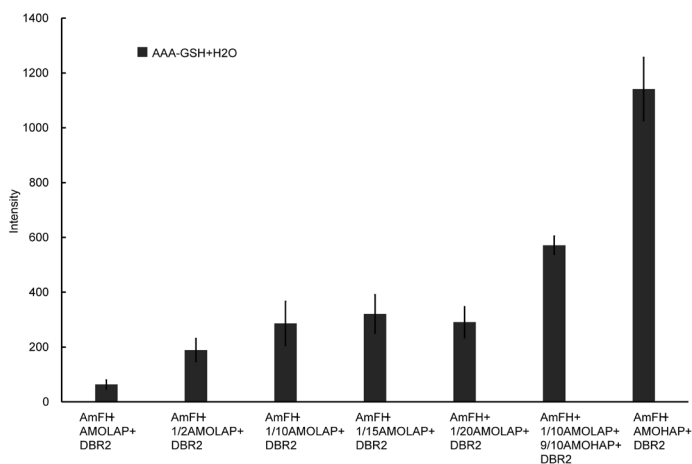
Effects on product profile by DBR2 or ALDH1 dosage

Figure 7. Artemisinic aldehyde **(**AAA) conjugated compounds in leaves of *N. benthamiana* agro-infiltrated with dilution of AMOLAP/HAP. Data represent peak intensities for each of the compounds in LC-QTOF-MS analysis. Peak intensities are the mean ± S.D. of three agro-infiltrated leaves.

We also tested the effect of *DBR2* and *ALDH1* gene dosage on DHAA:AA related product ratio by testing dilutions of DBR2 and ALDH1 in combination with AMOLAP or AMOHAP. Product analysis by LC-QTOF-MS of glycosylated products showed that with full DBR2 and ALDH1 agro-infiltration dosage the ratio of DHAA:AA was more skewed towards the DHAA branch of the pathway. However, when the DBR2 infiltration dosage was diluted 10-fold, the ratio of DHAA:AA decreased, shifting the pathway activity more towards the AA branch, with relatively little effect on the total product level (Fig. 6b, Fig. 1).

In contrast, 10-fold dilution of ALDH1 agro-infiltration dosage resulted in a strong decrease in glycosylated AA conjugates, while the glycosylated

DHAA conjugates were hardly affected (Fig. 6c), suggesting that ALDH1 has a preference for the DHAAA substrate over the AAA substrate (Fig. 1).

When both DBR2 and ALDH1 agro-infiltration dosage were diluted 10-fold the DHAA:AA related products ratio came close to one, with slight preference for the DHAA branch of the pathway with AMOHAP and slight preference for the AA branch of the pathway with AMOLAP (Fig. 6d). Combined these results suggest that the agro-infiltrated leaf 'chemotype' is determined by a combination of the AMOLAP/AMOHAP catalytic effectivity in combination with especially *DBR2* gene dosage (Figs. 6, S3d; Table S5).

Discussion

Here we have compared the different proteins from the artemisinin biosynthesis pathway encoded by genes isolated from high and from low artemisinin producing *A. annua* chemotypes. The expression levels of artemisinin biosynthesis genes (*ADS, AMO, DBR2* and *ALDH1*) were recently analysed in *A. annua* HAP and LAP chemotypes showing that the chemotype identities of *A. annua* did not correlate with differences in expression level of these genes in the absence of stress (Maes *et al.*, 2011). Nevertheless, total product yield in 5-week-old *A. annua* leaves is higher in LAP than HAP chemotypes (Maes *et al.*, 2011). Therefore, differences in protein activity rather than differences in gene expression level may account for differences in chemotype. In the present work, two *DBR2* variants were identified (Fig. 3) but characterization of the *in planta* activity showed they have equal activity (Fig. 5b, S3b). No difference was found for the ALDH1 amino acid sequence from *A. annua* HAP and LAP chemotypes. Therefore, the only consistent difference in artemisinin biosynthesis proteins in the branched pathway after ADS is in the AMOHAP and AMOLAP from the *A. annua* HAP and LAP chemotype, respectively (Fig. 2).

In the LAP chemotype CYP71AV1 consistently is seven amino acids longer than CYP71AV1 from the HAP chemotype (Fig. 2). When expressed together with *ADS* (+*FPS*+*HMGR*) the total product yield (based on the cumulative levels of AAOH, AAA, AA, DHAAOH and DHAA released from conjugated products) in leaves co-infiltrated with *AMOLAP* was approximately twice as high as in leaves co-infiltrated with *AMOHAP* (Table 3). This suggests a lower enzyme activity for AMOHAP than for AMOLAP (Fig. 5a). AMOHAP and AMOLAP seem to anchor equally well to the ER membrane (Fig. 4). The observed difference in efficiency may be caused by a different stability of the two proteins. Nevertheless, dilution of AMOLAP gene dosage in agro-infiltration experiments did not fully mimic the AMOHAP phenotype (Figs. 6a,

S3c). In theory, the lower efficiency of AMOHAP could also be due to a less efficient interaction with the cytochrome P450 reductase (CPR) (Lengler *et al.*, 2006) . Alignment of the AMOLAP and AMOHAP protein sequence with other related sesquiterpene oxidases like germacrene A oxidase (GAO) from several different Asteraceae (Nguyen *et al.*, 2010) shows that the N-terminal extension in AMOLAP is the exception. However, the CPR interaction domain does not map to the N-terminus (Sevrioukova *et al.*, 1999). Finally, it may be that the substrate entry or release of AMOHAP is compromised.

Variable 'chemotype' of leaves expressing artemisinin biosynthesis genes

We defined the 'chemotype' of the *N. benthamiana* agro-infiltrated leaves based on the DHAA and AA glycoside conjugates (Fig. 5). We note that the peak intensity is a relative quantification as detection efficiency (e.g. ionisation) may differ between compounds. However, comparison of the relative quantifications based on peak intensities from LC-QTOF-MS and the absolute quantification of DHAA and AA products by UPLC-MRM-MS or GC-MS (Table 1-3) indicates a good correlation between the two analytical techniques. Expression of *ADS+AMOHAP* results in a HAP chemotype (more DHAA than AA) and expression of *ADS+AMOLAP* in a LAP chemotype (more AA than DHAA) (Fig. 5a). However, when *DBR2* is included, the chemotype for both the combination *ADS+AMOHAP+DBR2* and *ADS+AMOLAP+DBR2* is changed to a HAP chemotype with relatively higher DHAA level (Fig. 5b). In all these combinations, the overall yield of DHAA was always lower for *AMOHAP* (*c.* 15-fold). When *ALDH1* was included, the chemotype remained that of HAP, but the relative yield of the gene combinations that include AMO-HAP was substantially increased and was now comparable to that of the gene combinations with AMOLAP (Fig. 5c). Because in the agro-infiltration assay in *N. benthamiana* addition of new genes to the pathway leads to a reduction in the relative dosage of the other genes infiltrated into the leaf, we tested whether the relative gene dosage affects the product profile. Lowering the dosage of AMOLAP with similar dosage of ADS and DBR2 resulted in a profile more closely related to that of AMOHAP, but also resulted in lower product yield of AA and DHAA conjugates (Fig. 6a). Lowering the relative dosage of DBR2 resulted in a reversion of the infiltrated leaf chemotype from HAP to LAP (Fig. 6b), while lowering the relative dosage of ALDH1 did not change the chemotype, but did decrease AA conjugates more than the DHAA conjugates (Fig. 6c). When both DBR2 and ALDH1 dosage were reduced, AMOHAP resulted in a more HAP related chemotype, while the com-

bination with AMOLAP resulted in a more LAP related chemotype (Fig. 6d).

N. benthamiana enzyme activities both enhance and limit the DHAA 'chemotype'

AMOHAP seems to be less efficient in the conversion to AA, presumably resulting in an early release of AAA. Indeed free AAA and an AAA gluta-thione conjugate were present at higher levels when the pathway was expressed in combination with AMOHAP than with AMOLAP (Table 1, Fig. 7). The AAA released by, particularly, AMOHAP may subsequently be substrate for double-bond reductases (endogenous from *N. benthamiana* or the co-expressed *DBR2*) to produce DHAAA and DHAAA derived products (Table 1, 2 and 3). The endogenous DBR2-like activity is far from saturating, as introduction of *A. annua DBR2* greatly enhanced the conversion to DHAAOH, DHAAA and DHAA (Table 1). The early release of AAA by AMOHAP also reveals the activity of an endogenous *N. benthamiana* reductase, similar to the *A. annua* aldehyde reductase (RED1) that catalyses the conversion of DHAAA to DHAAOH (Bertea *et al.*, 2005, Rydén *et al.*, 2010). Indeed, the presence of a RED1-like activity in tobacco was previously demonstrated through feeding experiments: tobacco leaves supplied with DHAAA and AAA form DHAAOH and AAOH, respectively (Zhang *et al.*, 2011). This suggests that in *N. benthamiana* the elevated pools of AAOH and DHAAOH (and glycosylated derivatives) may be the result of a reverse product flux from AAA back to AAOH and DHAAA to DHAAOH (Fig. 1). If this is the case the affinity of AMOHAP for the AAOH substrate may be underestimated.

Table 4. Artemisinic acid (AA) and dihydroartemisinic acid (DHAA) produced in agro-infiltrated *Nicotiana benthamiana* as identified and quantified by UPLC-MRM-MS.

(ng g^{-1} FW)	AmFH+AMOLAP+DBR2+ALDH1	AmFH+AMOHAP+DBR2+ALDH1
AA	8836 ± 1730	2589 ± 563
DHAA	10792 ± 341	2756 ± 547

 Results are means ± S.D. of three co-infiltrated leaves.

DHAA (conjugates) were detected in the leaves infiltrated with *AmFH+AMOLAP*, This suggests the presence of an endogenous aldehyde dehydrogenase, similar to ALDH1 from *A. annua*, which can produce DHAA from DHAAA (Teoh *et al.*, 2009). Alternatively, the low levels of DHAA could be the result of AMOLAP catalysing oxidation of DHAAA. Work in yeast showed that AMOLAP is far less effective in the conversion of DHAAA to DHAA than in the conversion

of AAA to AA (Teoh *et al.*, 2009). If we assume that also AMOHAP catalyses this step less effectively, expression of *ALDH1* in a HAP background should further enhance the production of DHAA in the HAP chemotype. Indeed, our data show that product flow towards DHAA increased *c.* 110-fold when expression of *AmFH+AMOHAP+DBR2* was combined with *ALDH1* (Table 4).

Figure 8. Schematic representation of the artemisinin precursor biosynthetic pathway in *Nicotiana benthamiana*. NbDBR2-like: *Nicotiana benthamiana* artemisinic aldehyde double-bond reductase-like enzyme; NbRED1-like: *Nicotiana benthamiana* dihydroartemisinic aldehyde reductase 1-like enzyme; NbALDH1-like: *Nicotiana benthamiana* aldehyde dehydrogenase 1-like enzyme; NbGTase: *Nicotiana benthamiana* glycosyl transferase; NbMTase: *Nicotiana benthamiana* malonyl transferase; NbGSTase: *Nicotiana benthamiana* glutathione-S-transferase; Enzyme names in round brackets indicates low activity of the enzyme for the corresponding step.

Conjugating activities limit precursor pool for artemisinin production

Most of the products formed upon agro-infiltration of the artemisinin pathway genes are present in the form of hexose and/or hexose/malonyl conjugates. This explains why a previous heterologous expression study in tobacco detected only limited levels of free AD, AAOH, DHAAOH and no AA or DHAA (Zhang *et al.*, 2011). The deglycosylation experiments show that the conjugated forms such as AA and DHAA accumulate to levels up to 8 to 10-fold higher than the corresponding free forms (Tables 1-3). The glycosylation/conjugation is most likely a response to the production of potentially toxic compounds, as leaves expressing *AmFH+AMOLAP/HAP* showed signs of necrosis (Fig. S2). Necrosis was stronger in leaves containing higher levels of free (and conjugated) AA and indeed presence of *DBR2* reduced both the necrotic phenotype and the free and conjugated AA levels (Fig. S2, Table 1). In addition to the AA/DHAA glycosides, AAA glutathione conjugates were detected. All these conjugating activities limit accumulation of DHAA, the direct precursor of artemisinin (Fig. 8). Interestingly, in extracts from *A. annua* HAP flowers, no glycosides of AAOH, AA, DHAAOH and DHAA or glutathione conjugates of AAA were detected (Fig. S8). This indicates that the cells of *A. annua* that produce artemisinin either do not have competing glycosyl/malonyl/GSH transferase activity or, perhaps more likely, that product flux in these specialised *A. annua* cells is protected from such competing activities.

Concluding remarks

Our results show that the chemical profile of agro-infiltrated *N. benthamiana* leaves is a function of both type and relative dosage of the expression constructs. Results still do not fully explain the difference in HAP and LAP chemotypes found in *A. annua*. The expression level of the different biosynthesis genes do not differ between *A. annua* chemotypes (Maes *et al.*, 2011) and therefore, at present, the difference in CYP71AV1 catalytic efficiency between HAP and LAP is the only identified factor that contributes to this difference in chemotype. However, other, as yet unidentified, factors in *A. annua* may further contribute to the chemical difference in the HAP and LAP varieties.

Acknowledgements

Reference

Aker J, Borst JW, Karlova R, de Vries S. 2006. The *Arabidopsis thaliana* AAA protein CDC48A interacts in vivo with the somatic embryogenesis receptor-like kinase 1 receptor at the plasma membrane. *Journal of Structural Biology* 156: 62-71.

Bertea CM, Freije JR, van der Woude H, Verstappen FW, Perk L, Marquez V, De Kraker JW, Posthumus MA, Jansen BJ, de Groot A, Franssen MC, Bouwmeester HJ. 2005. Identification of intermediates and enzymes involved in the early steps of artemisinin biosynthesis in *Artemisia annua*. *Planta Medica* 71: 40-47.

Bertea CM, Voster A, Verstappen FW, Maffei M, Beekwilder J, Bouwmeester HJ. 2006. Isoprenoid biosynthesis in *Artemisia annua*: cloning and heterologous expression of a germacrene A synthase from a glandular trichome cDNA library. *Archives of Biochemistry and Biophysics* 448: 3-12.

Bouwmeester HJ, Wallaart TE, Janssen MH, van Loo B, Jansen BJ, Posthumus MA, Schmidt CO, De Kraker JW, Konig WA, Franssen MC. 1999. Amorpha-4,11-diene synthase catalyses the first probable step in artemisinin biosynthesis. *Phytochemistry* 52: 843-854.

De Vos R, Moco S, Lommen A, Keurentjes J, Bino RJ, Hall RD. 2007. Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. *Nature Protocols* 2: 778-791.

Dieckhaus CM, Fernández-Metzler CL, King R, Krolikowski PH, Baillie TA. 2005. Negative ion tandem mass spectrometry for the detection of glutathione conjugates. *Chemical Research in Toxicology* 18: 630-638.

Kim NC, Kim JG, Kim S, Lim H, Hahm T. 1992. Production of secondary metabolites by tissue culture of *Artemisia annua* L. *Journal of the Korean Agricultural Chemical Society* 35: 99-105.

Kohlen W, Charnikhova T, Liu Q, Bours R, Domagalska MA, Beguerie S, Verstappen F, Leyser O, Bouwmeester H, Ruyter-Spira C. 2011. Strigolactones are transported through the xylem and play a key role in shoot architectural response to phosphate deficiency in nonarbuscular mycorrhizal host Arabidopsis. *Plant Physiology* 155: 974-987.

Lengler J, Omann M, Duvier D, Holzmuller H, Gregor W, Salmons B, Gunzburg WH, Renner M. 2006. Cytochrome P450 reductase dependent inhibition of cytochrome P450 2B1 activity: Implications for gene directed enzyme prodrug therapy. *Biochemical Pharmacology* 72: 893-901.

Liu Q, Majdi M, Cankar K, Goedbloed M, Charnikhova T, Verstappen FWA, de Vos RCH, Beekwilder J, van der Krol S, Bouwmeester HJ. 2011. Reconstitution of the Costunolide Biosynthetic Pathway in Yeast and Nicotiana benthamiana. *PLoS ONE* 6: e23255.

Maes L, Van Nieuwerburgh F, Zhang Y, Reed D, Pollier J, Vande Casteele S, Inze D, Covello P, Deforce DLD L, Goossens A. 2011. Dissection of the phytohormonal regulation of trichome formation and biosynthesis of the antimalarial compound artemisinin in *Artemisia annua* plants. *New Phytologist* 189: 176 - 189.

Mercke P, Bengtsson M, Bouwmeester HJ, Posthumus MA, Brodelius PE. 2000. Molecular cloning, expression, and characterization of amorpha-4,11-diene synthase, a key enzyme of artemisinin biosynthesis in *Artemisia annua* L. *Archives of Biochemistry and Biophysics* 381: 173-180.

Nguyen DT, Göpfert JC, Ikezawa N, MacNevin G, Kathiresan M, Conrad J, Spring O, Ro DK. 2010. Biochemical conservation and evolution of germacrene A oxidase in Asteraceae. *Journal of Biological Chemistry* 285: 16588-16598.

Nguyen KT, Arsenault PR, Weathers PJ. 2011. Trichomes + roots + ROS = artemisinin: regulating artemisinin biosynthesis in *Artemisia annua* L. *In Vitro Cellular & Developmental Biology Plant* 47: 329-338.

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

Ro DK, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J, Chang MC, Withers ST, Shiba Y, Sarpong R, Keasling JD. 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440: 940-943.

Rydén AM, Ruyter-Spira C, Quax WJ, Osada H, Muranaka T, Kayser O, Bouwmeester H. 2010. The molecular cloning of dihydroartemisinic aldehyde reductase and its implication in artemisinin biosynthesis in *Artemisia annua*. *Planta Medica* 76: 1778-1783.

Sevrioukova IF, Li H, Zhang H, Peterson JA, Poulos TL. 1999. Structure of a cytochrome P450–redox partner electron-transfer complex. *Proceedings of the National*

Academy of Sciences 96: 1863-1868.

Sy L-K, Brown GD. 2002. The mechanism of the spontaneous autoxidation of dihydroartemisinic acid. *Tetrahedron* 58: 897-908.

Teoh KH, Polichuk DR, Reed DW, Covello PS. 2009. Molecular cloning of an aldehyde dehydrogenase implicated in artemisinin biosynthesis in *Artemisia annua*. *Botany* 87: 635-642.

Teoh KH, Polichuk DR, Reed DW, Nowak G, Covello PS. 2006. *Artemisia annua* L. (Asteraceae) trichome-specific cDNAs reveal CYP71AV1, a cytochrome P450 with a key role in the biosynthesis of the antimalarial sesquiterpene lactone artemisinin. *FEBS Letters* 580: 1411-1416.

Tikunov Y, Lommen A, De Vos C, Verhoeven HA, Bino RJ, Hall RD, Bovy AG. 2005. A novel approach for nontargeted data analysis for metabolomics. Large-scale profiling of tomato fruit volatiles. *Plant Physiology* 139: 1125-1137.

van Engelen FA, Molthoff JW, Conner AJ, Nap JP, Pereira A, Stiekema WJ. 1995. pBIN-PLUS: an improved plant transformation vector based on pBIN19. *Transgenic Research* 4: 288-290.

van Herpen TW, Cankar K, Nogueira M, Bosch D, Bouwmeester HJ, Beekwilder J. 2010. *Nicotiana benthamiana* as a production platform for artemisinin precursors. *PLoS ONE* 5: e14222.

Wallaart TE, Pras N, Beekman AC, Quax WJ. 2000. Seasonal variation of artemisinin and its biosynthetic precursors in plants of *Artemisia annua* of different geographical origin: proof for the existence of chemotypes. *Planta Medica* 66: 57-62.

Wallaart TE, van Uden W, Lubberink HG, Woerdenbag HJ, Pras N, Quax WJ. 1999. Isolation and identification of dihydroartemisinic acid from *Artemisia annua* and its possible role in the biosynthesis of artemisinin. *Journal of Natural Products* 62: 430-433.

Wang Y, Yang K, Jing F, Li M, Deng T, Huang R, Wang B, Wang G, Sun X, Tang KX. 2011. Cloning and characterization of trichome-specific promoter of *cpr71av1* gene involved in artemisinin biosynthesis in *Artemisia annua* L. *Molecular Biology* 45: 751-758.

Yoo SD, Cho YH, Sheen J. 2007. Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nature Protocols* 2: 1565-1572.

Zhang Y, Nowak G, Reed DW, Covello PS. 2011. The production of artemisinin precursors in tobacco. *Plant Biotechnology Journal* 9: 445-454.

Zhang Y, Teoh KH, Reed DW, Maes L, Goossens A, Olson DJ, Ross AR, Covello PS. 2008. The molecular cloning of artemisinic aldehyde $\Delta 11(13)$ reductase and its role in glandular trichome-dependent biosynthesis of artemisinin in *Artemisia annua*. *Journal of Biological Chemistry* 283: 21501-21508.

Supporting Information

Methods S1. Details of fractionation and UPLC-MRM-MS analysis.

Chromatographic separation was obtained on an Acquity UPLC BEH C18 column (100 × 2.1 mm, 1.7 μm; Waters) by applying a water/acetonitrile gradient to the column, starting from 5% (v/v) acetonitrile in water for 1.25 min and rising to 50% (v/v) acetonitrile in water in 2.35 min, followed by an increase to 90% (v/v) acetonitrile in water in 3.65 min, which was maintained for 0.75 min before returning to 5% acetonitrile in water using a 0.15 min gradient. Finally, the column was equilibrated for 1.85 min using this solvent composition. Column temperature and flow rate were 50°C and 0.5mL min^{-1}, respectively. Injection volume was 5 μl. The mass spectrometer was operated in positive electrospray ionization mode. Cone and desolvation gas flows were set to 50 and 1000 L h^{-1}, respectively. The capillary voltage was set at 3.0 kV, the source temperature at 150°C, and the desolvation temperature at 650°C. The cone voltage was optimized for individual artemisinin precursors using the Waters IntelliStart MS Console. Argon was used for fragmentation by collision-induced dissociation in the ScanWave collision cell. Multiple Reaction Monitoring (MRM) was used for identification and quantification by comparing retention times and MRM mass transitions with that of authentic AAOH, AAA, AA, DHAAOH, DHAAA, DHAA and artemisinin standards. MRM transitions for artemisinin precursors were optimized using the Waters IntelliStart MS Console as follows:

	Transitions (m/z)		
AAOH	221.16>203.27	221.16>147.09	
AAA	219.16>145.08	219.16>201.20	219.16>159.09
AA	235.16>189.22	235.16>199.25	235.16>217.21
DHAAOH	223.22>205.27	223.22>95.07	223.22>109.13
DHAAA	221.16>105.14	221.16>203.20	
DHAA	237.16>163.17	237.16>81.01	237.16>107.12
Artemisinin	283.19>219.21	283.19>247.19	283.19>265.22

Table S1. Primers used in this study.

No. primer	Sequence 5' to 3'*	Restriction site
1	5'CGGCGGATCCATGGCACTCTCACTGACCACTTCCA 3'	Bam HI
2	5'CGGCGGTACCCTAGAAACTTGGAACGAGTAACAACT-CAGCCTTTC 3'	Kpn I

3	5'CGGCGGATCCATGAAGAGTATACTAAAAGCAATGG-CACTCTCACTGACCAC 3'	Bam HI
4	5'CGGCGGTACCCTAGAAACTTGGAACGAGTAACAACT-CAGTCTTTC 3'	Kpn I
5	5' ATGGATCCGTCTGAAAAACCAACCTTG3'	Bam HI
6	5' TATAGCGGCCGCTAGAGGAGTGACCCTT3'	Not I
7	5'GACAAATCTAGAAAGATGAGCTCAGGAGCTAATG-GAAG 3'	XbaI
8	5'CACAAAGCGGCCGCTTAAAGCCACGGGGAATCATA 3'	Not I
9	5'ATGGATCCATGAAGAGTATACTAAAAGCAATGG-CACTCTCACTGA3'	Bam HI
10	5'TATGGTACCGTGACCAA TAATGGGAAGCCGCCAT 3'	Kpn I
11	5'ATGGATCCATGGCACTCTCACTGACCACTTC-CATTGCTCT3'	Bam HI
12	5'TATGGTACCGTGACCAATAATGGGAAGTCGCCATG-GCTC3'	Kpn I
13	5'GCCGGTACCATGAGCAAGGGCGAGGAGCTGTTCA3'	Kpn I
14	5'TTAATTGCGGCCGCTTATTCCACATCTTTATACAGCTC-GTCCATG3'	Not I
15	5'GACAGTGCTGAATTCCCATTAACATCTGATAACATTA-AAG3'	EcoR I
16	5'CTCGCCCTTGCTCATGGTACCCAAGAAACTTGGAAC-GAGT3'	Kpn I
17	5'ACAGTGCTGAATTCCCATTAACATCTGATAACATTA-AAGC3'	EcoR I
18	5'CCTCGCCCTTGCTCATGGTACCCAAGAAACTTGGAAC-GAG3'	Kpn I

* Restriction sites are shown in italics.

Table S2. Unconjugated artemisinin precursors as identified and quantified by UPLC-MRM-MS in *Artemisia anuua*.

(µg mg-1 FW)	A. annua-HAP	A. annua -LAP
AA	0.16	0.35
DHAA	8.14	0.55
Artemisinin	0.95	0.04

AA: artemisinic acid, DHAA: dihydroartemisinic acid.

```
        RACE-1   -TCTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACGCGGGGAACCATCAATTGTATACTAAAAGCAATGGCACTCTCACTGACCAC  99
        RACE-2   -TCTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACGCGGGGAACCATCAATTGTATACTAAAAGCAATGGCACTCTCACTGACCAC  99
        RACE-3   --TTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACGCGGGG----ATCAATTGTATACTAAAAGCAATGGCACTCTCGCTGACCAC  94
HAP     RACE-4   -TCTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACGCGGGG----ATCAATTGTATACTAAAAGCAATGGCACTCTCACTGACCAC  95
        RACE-5   ----------------------TAAGCAGTGGTATCAACGCAGAGTACG-GGGGAACCATCAATTGTATACTAAAAGCAATGGCACTCTCACTGACCAC  76
        RACE-6   TGGTATCAACGCAGAGTACGCGGGGAAGCAGTGGTATCAACGCAGAGTACGCGGGGAACCATCAATTGTATACTAAAAGCAATGGCACTCTCACTGACCAC 100
        RACE-7   ----------------------TAAGCAGTGGTATCAACGCAGAGTACGCGGGG----ATCAATTGTATACTAAAAGCAATGGCACTCTCACTGACCAC  73
        AMOLAP   --------------------------------------------------------------ATGAAGAGTATACTAAAAGCAATGGCACTCTCACTGACCAC  41
        AMOHAP.1 -----------------------------------------------------------------GAGTATACTAAAAGCAATGGCACTCTCACTGACCAC  36
        AMOHAP.5 -TATTCAATTCCTATATCAATCCCATACAACAATATAAACAATCTCATCATTTGCGACCATAAAGAGTATACTAAAAGCAATGGCACTCTCACTGACCAC  99

        RACE-1   TTCCATTGCTCTTGCCACTATCCTCTTCTTCGTTTACAAGTTCGCTACTCGTTCCAAATCCAACAAAAACAGCCTTCCTGAGCCATGGCGACTTCCCATT 199
        RACE-2   TTCCATTGCTCTTGCCACTATCCTCTTCTTCGTTTACAAGTTCGCTACTCGTTCCAAATCCAACAAAAACAGCCTTCCTGAGCCATG-CGACTTCCCATT 198
        RACE-3   TTCCATTGCTCTTGCCACTATCCTCTTCTTCGTTTACAAGTTCGCTACTCGTTCCAAATCCAACAAAAACAGCCTTCCTGAGCCATGGCGACTTCCCATT 194
HAP     RACE-4   TTCCATTGCTCTTGCCACTATCCTCTTCTTCGTTTACAAGTTCGCTACTCGTTCCAAATCCAACAAAAACAGCCTTCCTGAGCCATGGCGACTTCCCATT 195
        RACE-5   TTCCATTGCTCTTACCACTATCCTCTTCTTCGTTTACAAGTTCGCTACTCGTTCCAAATCCAACAAAAACAGCCTTCCTGAGCCATGGCGACTTCCCATT 176
        RACE-6   TTCCATTGCTCTTGCCACTATCCTCTTCTTCGTTTACAAGTTCGCTACTCGTTCCAAATCCAACAAAAACAGCCTTCCTGAGCCATGGCGACTTCCCATT 200
        RACE-7   TTCCATTGCTCTTGCCACTATCCTCTTCTTCGTTTACAAGTTCGCTACTCGTTCCAAATCCAACAAAAACAGCCTTCCTGAGCCATGGCGACTTCCCATT 173
        AMOLAP   TTCCATTGCTCTTGCAACGATCCTTTTGTTCGTTTACAAGTTCGCTACTCGTTCCAAATCCACCAAAAAAAGCCTTCCTGAGCCATGGCGGCTTCCCATT 141
        AMOHAP.1 TTCCATTGCTCTTGCAACGATCCTTTTGTTCGTTTACAAGTTCGCTACTCGTTCCAAATCCACCAAAAAAAGCCTTCCTGAGCCATGGCGACTTCCCATT 136
        AMOHAP.5 TTCCATTGCTCTTGCAACGATCCTTTTGTTCGTTTACAAGTTCGCTACTCGTTCCAAATCCACCAAAAAAAGCCTTCCTGAGCCATGGCG---------- 189

        RACE-1   ATTGGTCACATGCATCAGTTGATTGGTACAATACCACATCGTGGGGTTATGGATTTAGCCAGAAAGTATGGATCTTTGATGCATTTGCAACTTGGTGAAG 299
        RACE-2   ATTGGTCACATGCATCAGTTGATTGGTACAATACCACATCGTGGGGTTATGGATTTAGCCAGAAAGTATGGATCTTTGATGCATTTGCAACTTGGTGAAG 298
        RACE-3   ATTGGTCACATGCATCAGTTGATTGGTACAATACCACATCGTGGGGTTATGGATTTAGCCAGAAAGTATGGATCTTTGATACATTTGCAACTTGGTGAAG 294
HAP     RACE-4   ATTGGTCACATGCATCACTTGATTGGTACAATACCACATCGTGGGGTTATGGATTTAGCCAGAAAGTATGGATCTTTGATGCATTTGCAACTTGGTGAAG 295
        RACE-5   ATTGGTCACATGCATCAGTTGATTGGTACAATACCACATCGTGGGGTTATGGATTTAGCCAGAAAGTATGGATCTTTGATGCATTTGCAACTTGGTGAAG 276
        RACE-6   ATTGGTCACATGCATCACTTGATTGGTACAATACCACATCGTGGGGTTATGGATTTAGCCAGAAAGTATGGATCTTTGATGCATTTGCAACTTGGTGAAG 300
        RACE-7   ATTGGTCACATGCATCACTTGATTGGTACAATACCACATCGTGGGGTTATGGATTTAGCCAGAAAGTATGGATCTTTGATGCATTTGCAACTTGGTGAAG 273
        AMOLAP   ATTGGTCACATGCATCACTTGATTGGTACAACGCCACATCGTGGGGTTAGGGATTTAGCCAGAAAGTATGGATCTTTGATGCATTTGCAACTTGGTGAAG 241
        AMOHAP.1 ATTGGTCACATGCATCACTTGATTGGTACAACGCCACATCGTGGGGTTAGGGATTTAGCCAGAAAGTATGGATCTTTGATGCATTTACAGCTTGGTGAAG 236
        AMOHAP.5 -------------------------------------------------------------------------------------------------- 189

        RACE-1   TTTCAACAATCGTGGTGTGTCATCTCCGAAATGGGCTAAAGAGAT 342
        RACE-2   TTTCAACAATCGTGGTGTCATCTCCGAAATGGGCTAAAGAGAT 341
        RACE-3   TTTCAACAATCGTGGTGTCATCTCCGAAATGGGCTAAAGAGAT 337
HAP     RACE-4   TTTCAACAATCGTGGTGTCATCTCCGAAATGGGCTAAAGAGAT 338
        RACE-5   TTTCAACAATCGTGGTGTCATCTCCGAAATGGGCTAAAGAGAT 319
        RACE-6   TTTCAACAATCGTGGTGTCATCTCCGAAATGGGCTAAAGAGAT 343
        RACE-7   TTTCAACAATCGTGGTGTCATCTCCGAAATGGGCTAAAAGAGAT 316
        AMOLAP   TTCCAACAATCGTGGTGTCATCTCCGAAATGGGCTAAAGAGAT 284
        AMOHAP.1 TTCCAACAATCGTGGTGTCATCTCCGAAATGGGCTAAAGAGAT 279
        AMOHAP.5 ---------------------------------------- 189
```

Figure S1. Alignment of the 5'RACE sequences of *AMOHAP* in Clustal W. Seven 5'RACE clones from a HAP chemotype are aligned with *AMOLAP* (DQ268763), *AMOHAP.1* (DQ315671) and the recently published *CYP71AV1* promoter isolated from an *Artemisia annua* HAP chemotype, here referred to as *AMOHAP.5* (Wang *et al.*, 2011). The start codons of *AMOLAP* and *AMOHAP* are indicated by red lines. Note that *AMOLAP* has an ATG at the same position as the start codon of *AMOHAP* but is preceded by another in-frame ATG.
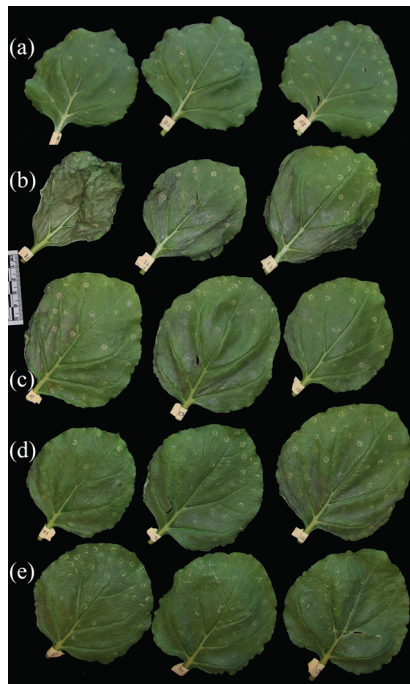
Figure S2. Necrotic symptoms in *Nicotiana benthamiana*. Leaves agro-infiltrated with (a) pBin empty vector, (b) *AmFH+AMOLAP*, (c) *AmFH+AMOLAP+DBR2*, (d) *AmFH+AMOHAP*, and (e) *AmFH+AMOHAP+DBR2* seven days after infiltration.
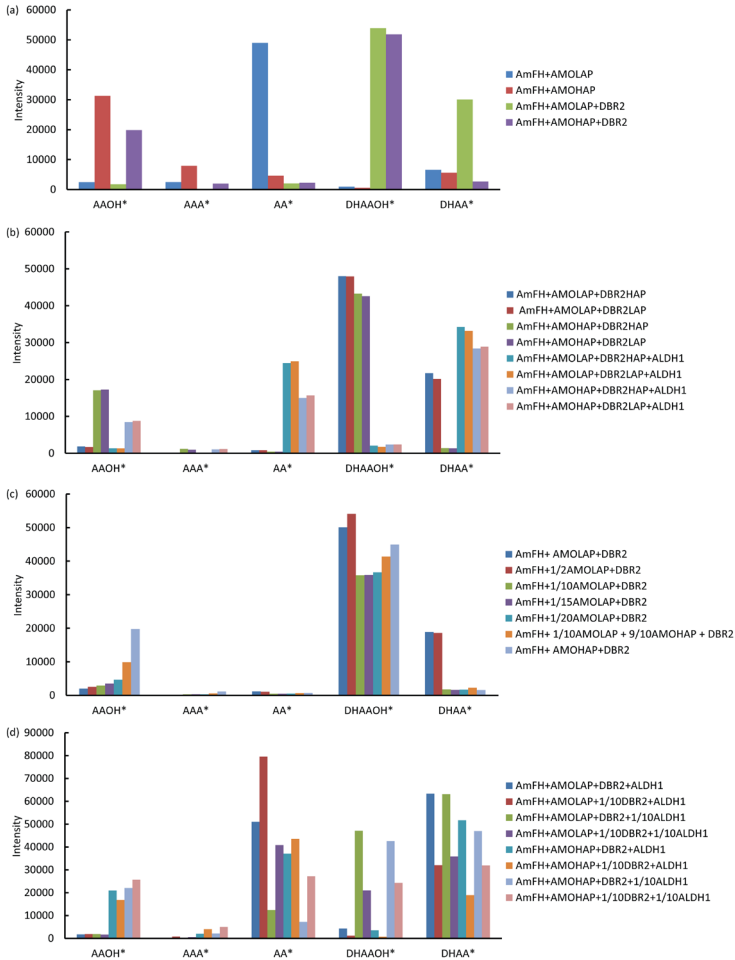

Figure S3. All conjugated compounds grouped according to free precursor detected in agro-infiltrated *Nicotiana benthamiana*. (a) Agro-infiltrated leaves with *AmFH+AMO-LAP/HAP* and *AmFH+AMOLAP/HAP+DBR2*. For all products identified by LC-QTOF-MS see Table 2. (b) Agro-infiltrated leaves with *AmFH+AMOLAP/HAP+DBR2LAP/HAP+ALDH1*. For all products identified by LC-QTOF-MS see Table S3. (c) Agro-infiltrated leaves with dilution of AMOLAP/HAP. For all products identified by LC-QTOF-MS see Table S4. (d) Agro-infiltrated leaves with dilution of DBR2 and ALDH1. For all products identified by LC-QTOF-MS see Table S5. Data shown with peak intensity which analysed by LC-QTOF-MS. Peak intensities are the mean of three agro-infiltrated leaves. (*: different conjugated compounds grouped according to free precursor)
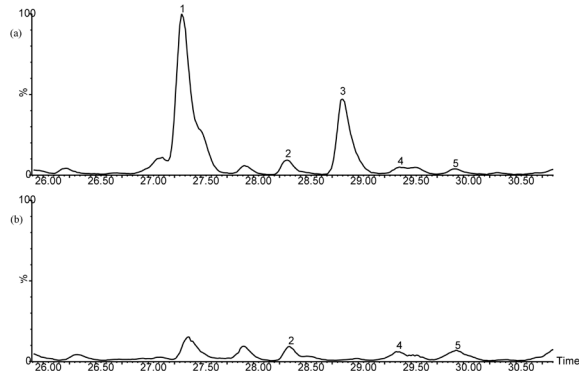
Figure S4. Malonylated glycosylated compounds are not cleaved by Viscozyme L. (a) LC-QTOF-MS chromatogram (total ion count) of methanol extract of *Nicotiana benthamiana* leaves infiltrated with *AmFH+AMOLAP+DBR2* without Viscozyme-treatment. (b) LC-QTOF-MS chromatogram (total ion count) of methanol extract of *Nicotiana benthamiana* leaves infiltrated with *AmFH+AMOLAP+DBR2* after Viscozyme-treatment. Peak 1: DHAA-Hex2, Peak 2: DHAA-Hex2-Mal, Peak 3: DHAAOH-Hex2, Peak 4: AA-Hex2-Mal, Peak 5: DHAAOH-Hex2-Mal. (DHAA: dihydroartemisinic acid; DHAAOH: dihydroartemisinic alcohol; AA: artemisinic acid; Hex: compound conjugated with hexose; Mal: compound conjugated with malonate)
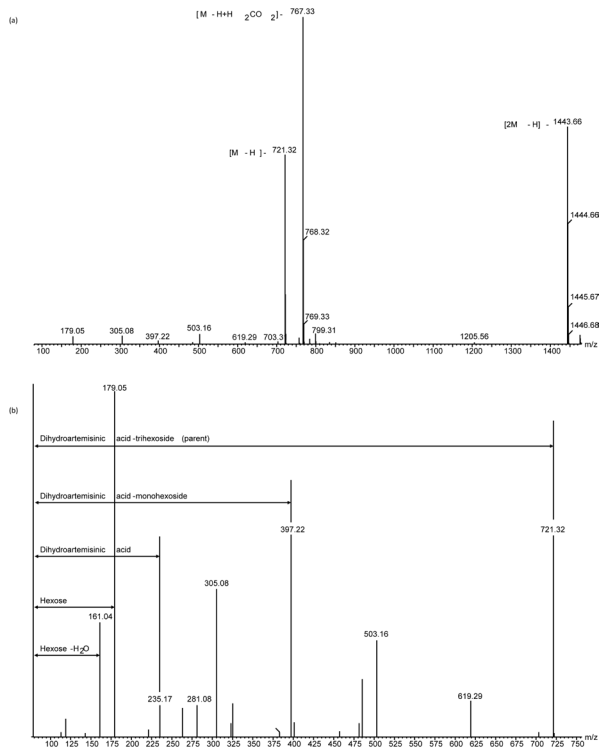
Figure S5. MS/MS spectrum and MS/MS fragmentation of DHAA-Hex3. (a) MS/MS spectrum of the compound eluting at 23.87 min, DHAA-Hex3 (m/z 721.32;[M-H]$^-$). (b) MS/MS fragmentation of mass 721.32 eluting at 23.87 min. (DHAA: dihydroartemisinic acid, Hex: compound conjugated with hexose)
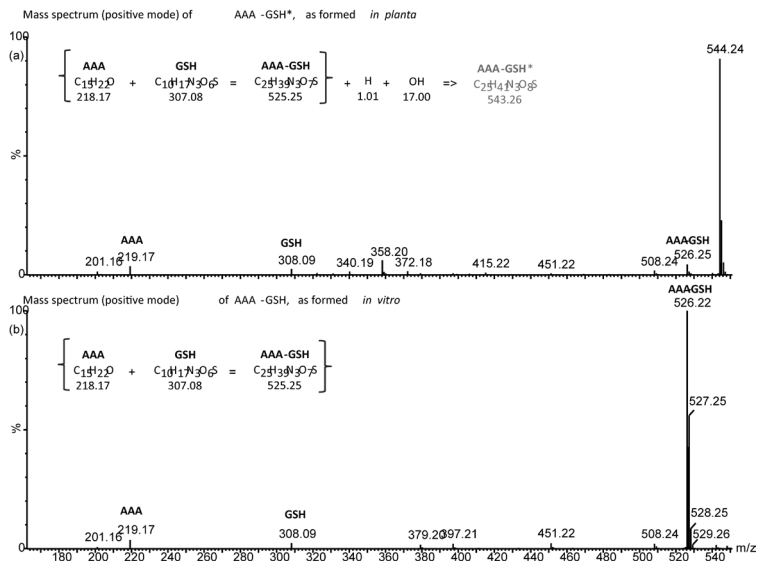
Figure S6. Comparison of *in vitro* and *in planta* formed AAA-GSH conjugates. (a) Mass spectrum (positive mode) of AAA-GSH*, as formed *in planta* (Table 2). The mass suggests a AAA-GSH derivative with 18 D mass increase, suggesting an additional OH and H. This is most easily explained by a reduction of the endocyclic double bond in AAA and hydroxylation. (b) Mass spectrum (positive mode) of AAA-GSH, as formed *in vitro* (see Fig. S6c). (AAA: artemisinic aldehyde, GSH: glutathione)
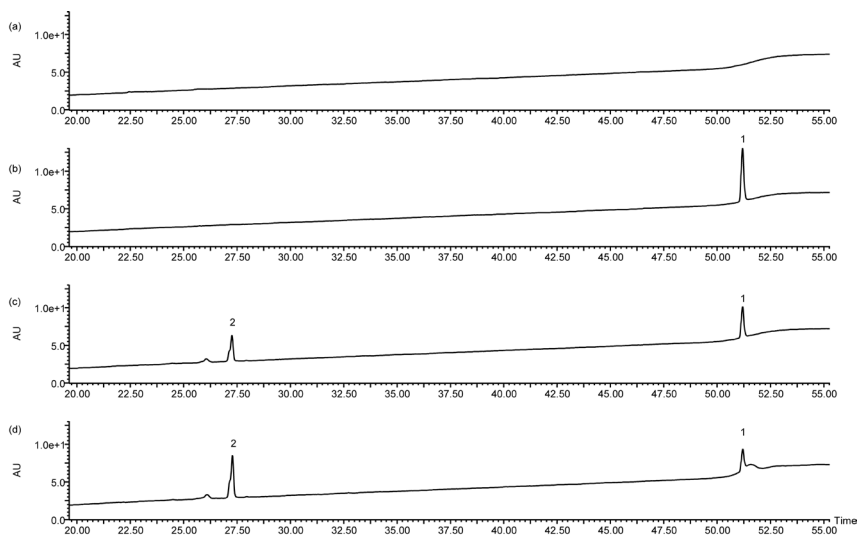
Figure S7. *In vitro* conjugation of artemisinic aldehyde (AAA) with glutathione. LC-QTOF-MS analysis of (a) Glutathione (GSH) incubated with Glutathione S transferase (GST), (b) AAA+GST, (c) AAA+GSH and (d)AAA+GSH+GST. Peak 1: AAA, Peak 2: AAA-GSH conjugate.
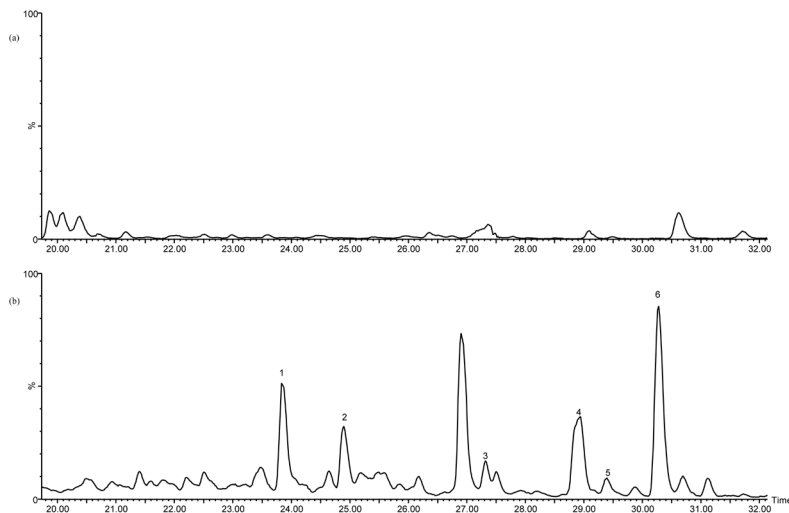
Figure S8. Absence of artemisinin precursor conjugates in *Artemisia annua* HAP chemotype. (a) LC-QTOF-MS chromatogram (total ion count) of methanol extract of flowers of *Artemisia annua* HAP chemotype. (b) LC-QTOF-MS chromatogram (total ion count) of methanol extract of *Nicotiana benthamiana* leaves infiltrated with *AmFH+AMO-LAP+DBR2*. Peak 1: DHAA-Hex3, Peak 2: DHAAOH-Hex3, Peak 3: DHAA-Hex2, Peak 4: DHAAOH-Hex2, Peak 5: AA-Hex2-Mal, Peak 6: DHAAOH-Hex2-Mal. (DHAA: dihydroartemisinic acid; DHAAOH: dihydroartemisinic alcohol; AA: artemisinic acid; Hex: compound conjugated with hexose; Mal: compound conjugated with malonate)

Table S3 Peak intensity of conjugated artemisinin precursors produced in agro-infiltrated *Nicotiana benthamiana* leaves with *AmFH+AMOLAP/HAP+DBR2LAP/HAP* and *AmFH+AMOLAP/HAP+DBR2LAP/HAP+ALDH1* were analysed by LC-QTOF-MS.

Putative ID [++]	Intensity [+]							
	AmFH+AMOLAP +DBR2HAP	AmFH+AMOLAP +DBR2LAP	AmFH+AMOHAP +DBR2HAP	AmFH+AMOHAP +DBR2LAP	AmFH+AMOLAP +DBR2HAP+ALDH1	AmFH+AMOLAP +DBR2LAP+ALDH1	AmFH+AMOHAP +DBR2HAP+ALDH1	AmFH+AMOHAP +DBR2LAP+ALDH1
AAOH-Hex2	229	220	1383	1312	158	156	1448	1545
AAOH-Hex2-Mal (I) [a]	1042	933	1350	1339	1029	997	1737	1716
AAOH-Hex2-Mal (II)	579	525	13994	14298	157	158	4589	4801
AAOH-Hex2-Mal (III)	ND	ND	339	313	2	ND	682	718
AAA-GSH-H_2O [b]	68	67	1168	968	98	98	1055	1135
(AA-Hex3) FA [c]	10	10	ND	ND	4923	5101	3318	3231
AA-Hex2 (I) [d]	77	62	48	50	4067	3969	2285	2441
AA-Hex2 (II)	ND	ND	3	3	6212	6210	3157	3320
AA-Hex2-Mal (I)	710	729	355	353	5969	5912	5008	5525
AA-Hex2-Mal ([2M-H]⁻)	32	29	21	22	3249	3703	1214	1150
(DHAAOH-Hex3) FA	8994	9033	6724	6254	205	188	302	340
DHAAOH-Hex2 (I)	12703	12612	10560	10243	261	249	330	346
DHAAOH-Hex2-Mal (I)	24565	24544	24521	24534	1374	1039	1637	1598
DHAAOH-Hex2-Mal (II)	1761	1749	1485	1529	215	238	100	117
(DHAA-Hex3) FA	13317	12376	366	364	17940	17983	14736	14812
DHAA-Hex2	3160	2876	23	18	7441	7599	4506	4660
DHAA-Hex2-Mal (I)	100	102	890	867	284	268	2133	2453
DHAA-Hex2-Mal (II)	5123	4800	113	95	8563	7343	7025	6975

[+] Peak intensities are the mean of three agro-infiltrated leaves. [++] AAOH: artemisinic alcohol; AA: artemisinic acid; AAA: artemisinic aldehyde; DHAAOH: dihydroartemisinic alcohol; DHAA: dihydroartemisinic acid; DHAAA: dihydroartemisinic aldehyde; ND: not detectable. [a] Hex: compound conjugated with hexose; Mal: compound conjugated with malonate; (I~III): different isobaric forms (i.e. identical accurate mass, but different retention times). [b] GSH: glutathione. [c] (FA): formic acid adduct. [d] AA-Hex2: The structure of artemisinic acid-12-β-diglucoside was confirmed by NMR (van Herpen *et al.*, 2010).

Table S4 Peak intensity of conjugated artemisinin precursors produced in agro-infiltrated *Nicotiana benthamiana* leaves with *AmF-H+AMOLAP/HAP+DBR2* compared with different dilution of *AMOLAP/HAP* were analysed by LC-QTOF-MS.

Putative ID[++]	Intensity[+]						
	AmFH +AMOLAP +DBR2	AmFH +1/2AMOLAP +DBR2	AmFH +1/10AMOLAP +DBR2	AmFH +1/15AMOLAP +DBR2	AmFH +1/20AMOLAP +DBR2	AmFH +1/10AMOLAP +9/10AMOHAP +DBR2	AmFH +AMOHAP +DBR2
AAOH-Hex2	220	313	253	309	309	714	1449
AAOH-Hex2-Mal (I)[a]	1057	919	655	596	511	802	1324
AAOH-Hex2-Mal (II)	ND	ND	10	23	52	29	332
AAOH-Hex2-Mal (III)	717	1293	1956	2564	3774	8335	16637
AAA-GSH-H$_2$O[b]	64	189	286	321	291	572	1142
(AA-Hex3) FA[c]	39	80	ND	ND	ND	ND	3
AA-Hex2 (I)[d]	60	17	ND	ND	ND	24	46
AA-Hex2 (II)	ND	ND	ND	ND	ND	ND	9
AA-Hex2-Mal (I)	681	579	237	225	243	325	332
AA-Hex2-Mal ([2M-H]⁻)	390	405	237	248	282	339	341
(DHAAOH-Hex3) FA	11759	10811	3940	4173	3933	5254	7919
DHAAOH-Hex2 (I)	12104	16755	6741	6614	7434	10300	11256
DHAAOH-Hex2-Mal (I)	24571	24573	24137	24253	24468	24564	24558
DHAAOH-Hex2-Mal (II)	1640	1943	946	849	814	1247	1194
(DHAA-Hex3) FA	12449	12604	934	844	938	947	345
DHAA-Hex2	2632	2276	176	127	113	338	ND
DHAA-Hex2-Mal (I)	122	131	302	374	408	568	1152
DHAA-Hex2-Mal (II)	3664	3588	337	276	253	425	85

+ Peak intensities are the mean of three agro-infiltrated leaves. ++AAOH: artemisinic alcohol; AAA: artemisinic aldehyde; AA: artemisinic acid; DHAAOH: dihydroartemisinic alcohol; DHAAA: dihydroartemisinic aldehyde; DHAA: dihydroartemisinic acid. ND: not detectable. [a]Hex: compound conjugated with hexose; Mal: compound conjugated with malonate; (I~III): different isobaric forms (i.e. identical accurate mass, but different retention times). [b]GSH: glutathione. [c](FA): formic acid adduct. [d]AA-Hex2: The structure of artemisinic acid-12-β-diglucoside was confirmed by NMR (van Herpen *et al.*, 2010).

Table S5 Peak intensity of conjugated artemisinin precursors produced in agro-infiltrated *Nicotiana benthamiana* leaves with *AmFH+AMOLAP/HAP+DBR2+ALDH1* with different dosage of *DBR2* and *ALDH1* were analysed by LC-QTOF-MS.

Putative ID[++]	Intensity[+]							
	AmFH+AMOLAP +DBR2 +ALDH1	AmFH+AMOLAP +1/10DBR2 +ALDH1	AmFH+AMOLAP +DBR2 +1/10ALDH1	AmFH+AMOLAP +1/10DBR2 +1/10ALDH1	AmFH+AMOHAP +DBR2 +ALDH1	AmFH+AMOHAP +1/10DBR2 +ALDH1	AmFH+AMOHAP +DBR2 +1/10ALDH1	AmFH+AMOHAP +1/10DBR2 +1/10ALDH1
AAOH-Hex2	222	274	217	194	2785	2498	2467	3696
AAOH-Hex2-Mal (I)[a]	935	1071	920	688	1467	1538	1210	1484
AAOH-Hex2-Mal (II)	568	531	741	753	16398	12526	18127	20135
AAOH-Hex2-Mal (III)	ND	ND	ND	ND	316	227	237	369
AAA-GSH-H_2O[b]	180	842	194	570	2018	4006	2127	4995
(AA-Hex3) FA[c]	15298	23704	1764	12096	8416	11777	733	5854
AA-Hex2 (I)[d]	8126	20699	1147	5176	5527	7701	1098	3906
AA-Hex2 (II)	1631	3766	203	643	928	1025	116	567
AA-Hex2-Mal (I)	23925	24628	9145	22476	21836	22578	5174	16747
AA-Hex2-Mal ([2M-H]⁻)	2061	6791	136	403	337	516	93	79
(DHAAOH-Hex3) FA	1247	379	15221	8577	1038	221	10567	7269
DHAAOH-Hex2 (I)	497	166	6180	1977	520	109	7001	2915
DHAAOH-Hex2-Mal (I)	2288	439	24376	9952	1838	414	23860	13496
DHAAOH-Hex2-Mal (II)	285	188	1316	529	123	8	1179	601
(DHAA-Hex3) FA	24592	23082	24615	24399	24600	11333	24555	21761
DHAA-Hex2	15312	3079	15467	3509	8019	1191	6035	2015
DHAA-Hex2-Mal (I)	512	984	361	837	4198	4153	2812	4525
DHAA-Hex2-Mal (II)	22905	4891	22649	7127	14913	2221	13561	3621

[+] Peak intensities are the mean of three agro-infiltrated leaves. [++] AAOH: artemisinic alcohol; AAA: artemisinic aldehyde; AA: artemisinic acid; DHAAOH: dihydroartemisinic alcohol; DHAAA: dihydroartemisinic aldehyde; DHAA: dihydroartemisinic acid. ND: not detectable. [a] Hex: compound conjugated with hexose; Mal: compound conjugated with malonate; (I–III): different isobaric forms (i.e. identical accurate mass, but different retention times). [b] GSH: glutathione. [c] (FA): formic acid adduct. [d] AA-Hex2: The structure of artemisinic acid-12-β-diglucoside was confirmed by NMR (van Herpen *et al.*, 2010).

Table S6. Metabolite profiling of *Artemisia annua* plants. (Data retrieved from Maes *et al.*, 2011).

(µg g^{-1} FW)	*A. annua*-HAP (5-wk)	*A. annua*-LAP (5-wk)
AAOH*	12	9
AAA	13	4
AA	9	1050
AB	11	320
DHAAOH	10	9
DHAAA	188	35
DHAA	450	55
AN	600	80
Total	1293	1562

AAOH: artemisinic alcohol; AAA: artemisinic aldehyde; AA: artemisinic acid; AB: arteannuin B; DHAAOH: dihydroartemisinic alcohol; DHAAA: dihydroartemisinic aldehyde; DHAA: dihydroartemisinic acid; AN: artemisinin.

Table 2 Conjugated artemisinin precursors produced in agro-infiltrated *Nicotiana benthamiana* leaf extracts. Non-volatile metabolites with mass intensity higher than 500 in LC-QTOF-MS, which were significantly increased in leaves agro-infiltrated with *AmFH+AMO-LAP*, *AmFH+AMOHAP*, *AmFH+AMOHAP+DBR2*, or *AmFH+AMOHAP+DBR2* were targeted for analysis by LC-QTOF-MS/MS fragmentation.

Ret (min)	Detected Mass(D)*	MS-MS fragments	Mol form	ΔMass (ppm)	Putative ID[++]	Intensity*			
						AmFH +AMOLAP	AmFH +AMOHAP	AmFH +AMOLAP +DBR2	AmFH +AMOHAP +DBR2
27.85	543.2793	381[M-Hex-H]⁻	$C_{27}H_{44}O_{11}$	1.3	AAOH-Hex2	319±57	3715±610	264±51	1753±823
28.87	629.2810	585[M-CO₂-H]⁻,543[M-Mal-H]⁻, 381[M-Mal-Hex-H]⁻	$C_{30}H_{46}O_{14}$	0	AAOH-Hex2-Mal (I) [a]	1096±288	2239±124	861±43	1378±246
29.35	629.2810	585[M-CO₂-H]⁻,543[M-Mal-H]⁻, 381[M-Mal-Hex-H]⁻	$C_{30}H_{46}O_{14}$	0	AAOH-Hex2-Mal (II)	606±37	24251±76	627±107	16215 ±5097
29.69	629.2810	585[M-CO₂-H]⁻,543[M-Mal-H]⁻, 381[M-Mal-Hex-H]⁻	$C_{30}H_{46}O_{14}$	0	AAOH-Hex2-Mal (III)	439±198	1109±104	ND	515±313
23.80	542.2536	272,254,210,179,143,128 **	$C_{25}H_{41}N_3O_8S$	1.0	AAA-GSH-H₂O [b]	2462±323	7891±65	130±22	1970±395
24.02	765.3181	719[M-H]⁻,395[M-2Hex-H]⁻, 233[M-3Hex-H]⁻	$C_{34}H_{54}O_{19}$	0.1	(AA-Hex3) FA [c]	21190 ±1730	612±35	100±80	ND
27.88	557.2598	395[M-Hex-H]⁻,233[M-2Hex-H]⁻	$C_{27}H_{42}O_{12}$	0	AA-Hex2 (I) [d]	5403 ±1218	917±71	121±61	371±186
28.48	557.2598	395[M-Hex-H]⁻,233[M-2Hex-H]⁻	$C_{27}H_{42}O_{12}$	0	AA-Hex2 (II)	3773±882	110±21	28±15	ND
29.38	643.2602	599[M-CO₂-H]⁻,395[M-Mal-Hex-H]⁻, 233[M-Mal-2Hex-H]⁻	$C_{30}H_{44}O_{15}$	0.5	AA-Hex2-Mal (I)	16021 ±320	2972±162	1652±327	1783±139
29.78	1287.5282	643[M-H]⁻,599[M-CO₂-H]⁻, 395[M-Mal-Hex-H]⁻,233[M-Mal-2Hex-H]⁻	$C_{60}H_{88}O_{30}$	4.5	AA-Hex2-Mal (2M-H)⁻	2611 ±1114	ND	129±13	82±9
24.86	753.3545	707[M-H]⁻,545[M-Hex-H]⁻, 383[M-2Hex-H]⁻,221[M-3Hex-H]⁻	$C_{34}H_{58}O_{18}$	0.9	(DHAAOH-Hex3) FA	145±46	85±8	14248±670	11573 ±206
28.80	545.2962	383[M-Hex-H]⁻,221[M-2Hex-H]⁻	$C_{27}H_{46}O_{11}$	1.8	DHAAOH-Hex2 (I)	343±89	194±46	14321 ±2617	14968 ±5018
30.27	631.2966	587[M-CO₂-H]⁻,545[M-Mal-H]⁻, 383[M-Mal-Hex-H]⁻,221[M-Mal-2Hex-H]⁻	$C_{30}H_{48}O_{14}$	1.8	DHAAOH-Hex2-Mal (I)	479±116	266±12	24475±120	24394 ±181
30.75	631.2966	587[M-CO₂-H]⁻,545[M-Mal-H]⁻, 383[M-Mal-Hex-H]⁻,221[M-Mal-2Hex-H]⁻	$C_{30}H_{48}O_{14}$	0	DHAAOH-Hex2-Mal(II)	ND	18±2	855±156	919±378
23.87	767.3338	721[M-H]⁻,397[M-2Hex-H]⁻, 235[M-3Hex-H]⁻	$C_{34}H_{56}O_{19}$	1.6	(DHAA-Hex3) FA	1237±110	ND	23356 ±1554	861±185
27.28	559.2755	397[M-Hex-H]⁻,235[M-2Hex-H]⁻	$C_{27}H_{44}O_{12}$	0.3	DHAA-Hex2	ND	ND	5188±2039	120±54
26.58	645.2759	601[M-CO₂-H]⁻,397[M-Mal-Hex-H]⁻, 235[M-Mal-2Hex-H]⁻	$C_{30}H_{46}O_{15}$	0	DHAA-Hex2-Mal (I)	1912±244	5539±174	236±14	1645±256
28.92	645.2759	601[M-CO₂-H]⁻,397[M-Mal-Hex-H]⁻, 235[M-Mal-2Hex-H]⁻	$C_{30}H_{46}O_{15}$	0	DHAA-Hex2-Mal (II)	ND	20±6	1303±528	76±35
29.83	645.2759	601[M-CO₂-H]⁻,397[M-Mal-Hex-H]⁻, 235[M-Mal-2Hex-H]⁻	$C_{30}H_{46}O_{15}$	0	DHAA-Hex2-Mal (III)	3424±549	64±8	ND	ND

AA: artemisinic acid, DHAA: dihydroartemisinic acid.

Chapter 3

The *A. annua* Lipid Transfer Protein AaLTP3 prevents reflux of dihydroartemisinic acid that was transported to the apoplast and increases production of artemisinin in *Nicotiana benthamiana*

Bo Wang[1,4], Arman Beyraghdar Kashkooli[1,4], Adrienne Sallets[2], Hieng-Ming Ting[1], Norbert C.A. de Ruijter[3], Marc Boutry[2], Harro Bouwmeester[1], Alexander R. van der Krol[1]

[1]Laboratory of Plant Physiology, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands.
[2]Institut des Sciences de la Vie, Université catholique de Louvain , Croix du Sud, 4-5, Box L7.07.14, 1348 Louvain-la-Neuve, Belgium.
[3]Laboratory of Cell Biology, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands.
[4]These authors contributed equally to this paper

Abstract

Metabolic engineering in plants for the production of high value terpenes is currently limited by our understanding of transport and sequestration of the heterologous products. For instance, although all biosynthesis genes for the production of artemisinin have been identified, ectopic expression of these genes in *Nicotiana benthamiana* yielded mostly glycosylated pathway intermediates and only very little free (dihydro)artemisinic acid [(DH)AA] or artemisinin. Here we tested three *Lipid Transfer Protein* (*LTP*) genes from *Artemisia annua* for their role in the sequestration of free artemisinin pathway (AN-PW) products – produced by co-expression of the AN-PW genes in the apoplast of *N. benthamiana* leaves. Ectopically produced non-glycosylated AN-PW products like (DH)AA accumulate in the apoplast, indicative of a low intrinsic transport activity in *N. benthamiana* leaf cells. The amount of apoplastic (DH)AA is slightly enhanced by the transient co-expression of an ABC membrane transporter, *AaPDR2*, while co-expression of three *AaLTPs in combination with* AN-PW+*AaPDR2*, results in even higher levels of free (DH)AA in the apoplast. Subsequent assays with individual LTPs show that AaLTP3 is most effective for accumulation of (DH)AA in the apoplast. When (DH)AA is infiltrated into the apoplast of *N. benthamiana* leaves it is very rapidly taken up by and converted inside the cells to (DH)AA-glycosides, indicating that active retention in the apoplast is needed. The rapid intracellular glycosylation of infiltrated (DH)AA was reduced by 27% by transient expression of *AaPDR2*, by 74% by expression of *AaLTP3* and by 82% by the combined expression of *AaPDR2* and *AaLTP3*. Free DHAA and AA were converted to artemisinin and arteannuin B, respectively, in necrotic *N. benthamiana* leaves and co-expression of AN-PW+*AaLTP3*+*AaPDR2* resulted in the highest yield of artemisinin and arteanuin B.

Keywords

Introduction

Artemisinin is a highly oxygenated sesquiterpene containing a unique 1,2,4-trioxane ring structure produced by the plant *Artemisa annua*. Artemisinin is used for the treatment of malaria and research of the past decade has led to the identification of all biosynthesis steps that are required for the production of the artemisinin precursor dihydroartemisinic acid (DHAA) (Bouwmeester et al., 1999; Cheng et al., 2004; Covello et al., 2007; Teoh et al., 2009; Teoh et al., 2006; Zhang et al., 2008). The final conversion of DHAA to artemisinin is thought to be the result of non-enzymatic oxidation (Brown, 2010).

DHAA accumulates in the subcuticular space of capitate glandular secretory trichomes on the surface of leaves and flowers of *A. annua* (Duke et al., 1994). These glandular trichomes consist of 10-celled structures with two basal cells, two stalk cells, and three pairs of secretory cells and a subcuticular sac (Duke and Paul, 1993). The subapical cells contain chloroplasts and seem to have a function in the overall metabolism of glandular trichomes (Olsson et al., 2009). The apical cells contain extensive ER membrane structures, typical of secretory cells, and the genes encoding the artemisinin biosynthesis enzymes (*ADS, ADO, DBR2, ALDH1*) are expressed in these apical cells (Olsson et al., 2009; Soetaert et al., 2013). The biosynthesis of the direct precursor of artemisinin (DHAA) takes place inside the apical cells, while artemisinin is formed outside the cells in the subcuticular space, possibly upon rupture of the cuticle of the trichome (Olsson et al., 2009). This implies that DHAA is transported out of the apical secretory cells into the subcuticular space. Studies with *A. annua* cell suspensions which produce artemisinin have shown that most of the produced artemisinin is in the medium (Durante et al., 2001), which could support that DHAA after being transported out of the cell is rapidly converted in the cell wall or medium.

When the artemisinin pathway (AN-PW) genes are transiently expressed in *N. benthamiana* pathway intermediates (artemisinic alcohol (AAOH), artemisinic aldehyde (AAA), artemisinic acid (AA), dihydroartemisinic alcohol (DHAAOH); dihydroartemisinic aldehyde (DHAAA); dihydroartemisinic acid (DHAA)) are produced but they are present in glycosylated form presumably present in the cytosol and/or vacuole (Ting et al., 2013). Only low levels of non-glycosylated, free AN-PW intermediates are detected, suggesting that the capacity of *N. benthamiana* leaf cells to sequester these free compounds away from the glycosyl transferases is very limited. This raises the question whether sequestration of AN-PW intermediates can be stimulated by co-expression of the AN-PW genes with oth-

er genes from *A. annua* which may encode relevant transport proteins.

Here we characterized the activity of *A. annua* LTPs which may be involved in the sequestration of DHAA into the apoplast. LTPs are small, basic proteins which are characterized in *in-vitro* assays by their ability to transfer lipids between membranes. LTPs contain a hydrophobic cavity which may bind lipophilic molecules, but the exact function of LTP proteins *in vivo* has not been fully elucidated. Most LTP proteins contain a signal peptide sequence for secretion to the apoplast and extracellular localization of LTP proteins has been confirmed by LTP-GFP localization studies (Huang et al., 2013; Kader, 1997; Yeats and Rose, 2008). Suggestions for a putative role of LTPs in terpene transport come from the fact that LTP genes show very high transcriptional activity in glandular trichomes that are also known to produce terpenes (peppermint: (Lange et al., 2000), alfalfa: (Aziz et al., 2005), Artemisia annua: (Bertea et al., 2006), hop: (Wang et al., 2008), Salvia fruticosa: (Chatzopoulou et al., 2010), tomato: (Schilmiller et al., 2010) and tobacco: (Harada et al., 2010)). Moreover, recently it was shown that over-expression of NtLTP1 in transgenic Nicotiana tabacum resulted in increased secretion of diterpenes from the glandular trichomes (Choi et al., 2012).

The *AaLTP* genes were used in a transient expression assay together with AN-PW genes and the previously characterized ABC transporters for AN-PW intermediates, *AaPDR1* and *AaPDR2* (Boutry, personal communication), to study the effect on apoplastic accumulation of (DH)AA in leaves of agroinfiltrated *N. benthamiana*. When AN-PW genes are expressed in *N. benthamiana* leaves, only small amounts of free (non-glycosylated) (DH)AA are detected in leaf extracts. Here we show that free AA and DHAA and other non-glycosylated AN intermediates are in the apoplast, suggesting the presence of a low intrinsic transport activity for these compounds in *N. benthamiana*. Co-expression of the AN-PW genes with either the two *AaPDR* genes or the three *AaLTP* genes had only a small effect on the total level of free AN-PW intermediates in the apolplast. However, co-expression of the AN-PW genes with both the *PDR* and *LTP* genes resulted in a significant increase in free AN-PW intermediates in the apoplast. Using the transient expression system we demonstrate different specificity in transport activity for the two *PDR* genes and we show specificity in LTP function, with co-expression of *PDR2+LTP3* being the most effective combination to accumulate free AN-PW intermediates into the apoplast. Combined, the results reveal new opportunities to enhance production of valuable medicinal terpenoids in heterologous production systems.

Materials and methods

Transient expression in *N. benthamiana* leaves

N. benthamiana was grown in soil in 7 cm pots in greenhouse kept at 22 ºC. The agro-infiltration was in 4th leaves of 5 week old *N. benthamiana* plants. Agroinfiltration of *N. benthamiana* leaves was done as described (van Herpen et al., 2010). The final OD_{600} value was adjusted to 0.5 with the agroinfiltration buffer, including 10 mM MES (2-(N-morpholino) ethanesulfonic acid, Duchefa Biochemie), 10 mM $MgCl_2$ and 100 μM acetosyringone (4'-hydroxy-3',5'-di-methoxyacetophenone, Sigma). then individual *A. tumefaciens* strains transformed with different expression constructs were co-infiltrated into *N. benthamiana* leaves using a 1 ml syringe without needle. PW means AN-PW genes including *V5*, *ADO*, *DBR2*, *ALDH1* (Ting et al., 2013; van Herpen et al., 2010). Since HMGR is a key rat-limiting step in the sesquiterpene biosynthesis pathway, we add the HMGR to boost the AN production in *N. benthamiana*. PW* means AN-PW genes with an extra *HMGR* in the agroinfiltration. The relative dosage of each *A. tumefaciens* strain between treatments within the same experiment was kept the same by diluting with an *A. tumefaciens* strain carrying an empty vector construct (EV). *A. tumefaciens* harboring TBSV p19 (P19) was included to maximize the protein production by suppression of gene silencing (Voinnet et al., 2003). Infiltrated leaves were harvested at the days indicated (varying from 3-18 days post agroinfiltration (dpi)).

A. annua RT-PCR gene expression analysis

The high-artemisinin production (HAP) chemotype *A. annua* were grown in the greenhouse chamber with 12 h light at 20ºC and 12 h darkness at 18 ºC. The shoot apex and leaves from top (Leaf1) to bottom (Leaf10) were collected from 2 month old *A. annua* plants and pooled per leaf number for RNA extraction. From each pooled and frozen grinded sample 100 mg was extracted for gene expression analysis. RNA was isolated from the frozen grinded samples and DNase I (Invitrogen) was used to remove remaining genomic DNA. Reverse transcription was carried out with the SuperScript First-Strand Synthesis System for the cDNA synthesis (Invitrogen). Real-time quantitative PCR was performed using the iQ5 RT-PCR (BioRad). Primers used are listed in Table S1. The *A. annua* β-actin gene was used for the normalization (Lu et al., 2013; Olofsson et al., 2011). ΔΔCT was calculated as follows: ΔCT=CT(Target)-CT(Actin), ΔΔCT was normalized using ΔCT. The fold change value was calculated by $2^{-\Delta\Delta CT}$.

Cloning of the different expression vectors

Cloning of the LTP overexpression genes

The full-length *AaLTP1*, *AaLTP2* and *AaLTP3* were amplified from cDNA made from RNA isolated from *A. annua* flowers by RACE PCR (Clontech, USA) (Bertea et al., 2006). The primer pairs AaLTP1-F/AaLTP1-R, AaLTP2-F/AaLTP2-R and AaLTP3-F/AaLTP3-R (Table S1) introduce a NcoI site at the start codon and a NotI restriction site after the stop codon of each LTP coding sequence. The PCR fragments were cloned into pIV1A2.1 vector (www.impactvector.com) between the CaMV35S promoter and RBCS1 terminator fragment. pIV1A2.1/AaLTP1, pIV1A2.1/AaLTP2, pIV1A2.1/AaLTP3 were cloned into binary vector pBinPlus by LR reaction (Invitrogen) and the resulting pBIN/AaLTP1, pBIN/AaLTP2, and pBIN/AaLTP3 were transformed into the *A. tumefaciens* AGL-O which was used for the agroinfiltration experiments.

Cloning of the LTP-GFP and LTP-RFP fusion protein genes.

For the LTP-GFP fusion protein studies we first made an entry vector with EGFP. The EGFP coding sequence was amplified by PCR using primers EGFP_c-term-F/EGFP_c-term-R (Table S1) by using pBinEGFP as template (http://www.wageningenur.nl/en/show/Productie-van-farmaceutische-en-industriele-eiwitten-door-planten.htm). After digesting with NotI and SacI, GFP was cloned into ImpactVector pIV1A_2.1 (www.impactvector.com) under the control of CaMV 35S promoter to generate an entry vector pIV1A_2.1/EGFP. To remove the stop codon of AaLTP1, AaLTP2 and AaLTP3, DNA fragments were amplified using the primer pairs AaLTP1_gfp-F/AaLTP1_gfp-R, AaLTP2_gfp-F/AaLTP2_gfp-R and AaLTP3_gfpF/AaLTP3_gfp-R (Table S1) by using the pIV1A_2.1/AaLTP1, pIV1A_2.1/AaLTP2 and pIV1A_2.1/AaLTP3 vector as template, respectively. The PCR products were digested with BamHI and NotI and subcloned into vector pIV1A_2.1/EGFP. The resulting pIV1A_2.1/AaLTP1-GFP, pIV1A_2.1/AaLTP2-GFP and pIV1A_2.1/AaLTP3-GFP containing the CaMV 35S promoter were cloned into the pBinPlus binary vector (Invitrogen) using LR recombination (van Engelen et al., 1995). The pBinPlus constructs containing the 35S:AaLTP1-EGFP, 35S:AaLTP2-EGFP and 35S:AaLTP3-EGFP were transferred to *A. tumefaciens* AGL-O using electroporation.

For the LTP-RFP fusion protein genes the LTP coding sequence without stop-codon were cloned in frame in TOPO vector (Invitrogen). The resulting TOPO/AaLTP1, TOPO/AaLTP2 and TOPO/AaLTP3 were cloned under control

of CaMV 35S promoter into the pK7RWG2 binary vector with the RFP coding sequence (http://gateway.psb.ugent.be/) using LR recombination (Invitrogen)

Metabolite analysis using LC- MS

For the AN-PW free compound analysis, the agro-infiltrated leaves were harvested, grinded in liquid nitrogen. Frozen grinded samples (100 mg) were extracted in 300 ul methanol using 10 min sonication, after which the samples were centrifuged at 18,625 x g at 4 °C. The supernatant was collected and filtered through 0.45 µm PTEE-membrane filter and diluted 10-fold with 10% methanol for analysis by a Waters Xevo tandem quadrupole mass spectrometer equipped with an electrospray ionization source and coupled to an Acquity UPLC system (Waters) as described (Ting et al., 2013). There is no standard for (DH)AAA, therefore, the data of (DH)AAA was given as relative peak intensity. For the AN-PW glycosides and conjugates, 100 mg of infiltrated leaf material was extracted in 300 ul methanol / 0.137% formic acid using 10 min sonication, after which the samples were centrifuged at 18,625 x g at 4 °C and filtered through a 0.45µm PTEE-membrane filter. The supernatant was analyzed on a LC-QTOF-MS (Waters). Since leaves showed signs of necrosis after 7 dpi, for samples collected from 7 dpi to 18 dpi the samples were freeze-dried and yield was determined by dry-weight rather than by fresh weight. For analysis 100 mg of dried samples was used following the same extraction procedure as for fresh-weight samples.

Apoplast fluid analysis using LC-MS

Apoplastic fluids were collected using the infiltration centrifugation technique with modifications (Joosten, 2012; Witzel et al., 2011). For the vacuum infiltration, the fresh leaves were cut into 5 cm strips and stack in a beaker filled with distilled water, and then applied vacuum with a vacuum pump. The vacuum was released slowly by opening the vent on the desiccator jar. For the apoplast fluid collection, leaf strips were placed on a grid resting on the conical bottom part of 50 ml plastic centrifuge tube (Greiner bio-one). Tubes were centrifuged at 400 g, 15 min at 4°C to collect the fluid in the bottom of the tube. After collection of the released apoplast fluid the leaf strips were weighted again, and subsequently shock frozen in liquid nitrogen for the metabolites analysis. The apoplast wash fluid was filtered through a 0.45 µm PTEE-membrane filter and the filtered fluid was dilute 5 times with 12.5% MeOH. For analysis of extracted compounds, the diluted apoplast fluid was injected into the UPLC-MRM-MS (Waters). The

concentration of the different products in the apoplast fluid was used to compare the different gene combinations in the transient expression assays.

Semi-quantification of LTP-EGFP fusion protein levels in agro-infiltrated leaves

We used LTP-GFP fluorescence as a semi-quantitative measure of LTP-EGFP protein accumulation. Leaves of *N. benthamiana* were infiltrated with *A. tumefaciens* carrying the *CaMV 35S::LTP-EGFP* expression constructs or with empty vector construct (control). Infiltrated leaves were harvested at 6 dpi and four regions of equal size at the same position were imaged per leaf under UV illumination in equal exposure time with a Leica MZIII fluorescence stereomicroscope and a Nikon Optiphot-2 coupled to a mercury lamp. Color images were split into individual R, G or B channels and the signal in the G channel was quantified in ImageJ 1.49n. Four biological replicate leaves (16 individual areas) were used to determine the average fluorescence signal in the G channel for each treatment.

Subcellular Localization of LTP-GFP and GFP-PDR

The LTP-GFP and GFP-PDR proteins were transiently expressed in the *N. benthamiana* leaves. Subcellular targeting analysis was carried out at different dpi by confocal laser scanning. In order to remove air from leaf sections, the leaf pieces were cut into 0.5 cm discs and mounted in perfluorodecalin (Sigma) (Littlejohn and Love, 2012). A Zeiss LSM 5 PASCAL lasser scanning microscope (Carl Zeiss) and a 63x (N.A. 1.4 oil) Plan Apochromat objective (Zeiss) was used for collect the images as described (Dong et al., 2013). The same samples were analyzed with laser excitation settings of 488-nm for GFP detection and 561-nm for RFP detection. Images were processed with Zeiss LSM image brower version 5.

Exclusion assay

4 weeks old *N. benthamiana* leaves were infiltrated with *A. tumefaciens* harboring *AaLTP3*, *AaPDR2*, *AaLTP3+AaPDR2* or *EV* as the control, each together with P19. In order to achieve the highest expressed protein levels, infiltrated leaves were kept on plant for 5 days and then harvested. For each treatment, four leaf disks of 1 cm diameter were excised from four individual leaves and were vacuum infiltrated with 3500 µl of 500 µM Na_2PO_4 buffer to make a con-

stant flow between the buffer and apoplastic region . Then 2500 μl of solution was taken out and a solution of AA was added into each well in a manner that each well got 0.625 mg of free compound. The 16-well plate was shaken gently for 2h on an orbital shaker and then the leaf disks were extracted with 300 μl of MeOH and 0.1% formic acid. Extracts were analyzed by LC-QTOF-MS.

Results

Isolation of *A. annua* glandular trichome *LTP* genes

Six different *LTP* sequences were identified in our *A. annua* trichome cDNA library (Bertea et al., 2006). Three of them (*AaLTP1*/2/3) were consistently detected, also in other published *A. annua* glandular trichome enriched cDNA libraries (Covello et al., 2007; Teoh et al., 2009), while three other LTP sequences (*AaLTP4*/5/6) were detected in an *A. annua* filamentous trichome enriched cDNA library (Soetaert et al., 2013) (Table 1). The expression of the six *LTP* genes was analyzed in different stages of *A. annua* leaf development and the apex of *A. annua* plants and compared to that of *ADS*, which encodes the first step in artemisinin biosynthesis. The expression profiles of *AaLTP4*/5/6 are different from the expression profile of *ADS* (Figure S1), which suggests they are not involved in AN-PW product transport, but in another process possibly in filamentous trichomes. *AaLTP1* and *AaLTP2* have a similar profile to that of *ADS* (Figure S1), while the expression profile of *AaLTP3* is different, showing a relative low expression in the shoot apex and much higher activity in older leaves. Since the pathway for artemisinin is active in glandular trichomes, *LTP1/2/3* were selected for further functional analysis in combination with AN-PW genes.

For the functional analysis of the *AaLTP* genes we also used two characterized *A. annua* ABC-G type transporter genes (*AaPDR1* and *AaPDR2*; Boutry, personal communication). Both *AaPDR1* and *AaPDR2* are present in multiple glandular trichome enriched cDNA libraries (Table 1). Expression profiling shows that the profile of *AaPDR1* is similar to that of *ADS* and *AaLTP1*/2, while the expression profile of *AaPDR2* is more similar to that of *AaLTP3* (Figure S1). The expression profile of *AaLTP3* and *AaPDR2* matches with the accumulation profile of DHAA over these developmental stages (Zhang et al., 2012). The three *LTP* genes with expression in glandular trichomes were cloned into binary expression vectors for functional analysis in combination with the AN-PW genes and the two *PDRs*.

Table 1. LTPs and PDRs expression in trichome databases. F: Filamentous trichomes; G: Glandular trichomes; ND: not detected; +: detected.

Data type	Tissues	LTP1	LTP2	LTP3	LTP4	LTP5	LTP6	PDR1	PDR2	Reference
RNAseq data	G vs. F	G	G	ND	F	F	F	G	ND	(Soetaert et al., 2013)
ESTs	G	+	+	+				+	ND	(Covello et al., 2007)
454 pyrosequencing transcriptome	G								ND	(Wang et al., 2009)
cDNA library	G + F	+	+	+	+	+	+	ND	+	(Bertea et al., 2006)
Trichome database	G + F	+	+	+				+	+	(Dai et al., 2010)
Tang's RNAseq	G vs. F	+	+	+				+	+	Personal communication
Brodelius' RNAseq	G vs. F							G	G	Personal communication

Subcellular localization of AaLTP1/2/3 upon transient expression in *N. benthamiana*

All three glandular trichome expressed *AaLTPs* contain a N-terminal signal peptide sequence, which is indicative for targeting to the ER and secretion to the apoplast. We tested the subcellular targeting of these AaLTPs by construction *LTP-GFP* fusion genes under control of the CaMV 35S promoter. Transient expression of these genes showed that at 4 dpi all LTP1-GFP, LTP2-GFP and LTP3-GFP were detected in internal membrane structures (presumably ER) (Figure 1A and S2). However, at 6-7 dpi the strongest GFP signal was observed outside the cells. Especially in the spongy mesophyll bridges of

fluorescent material between cells were observed, indicating local accumulation of the LTP-GFP fusion proteins in the apoplast (Figure 1B). Remarkably, for all three LTP-GFP fusion proteins, the extracellular accumulation seemed to be coordinated between cells in spongy mesophyll, resulting in organization of LTP-GFP proteins surrounding an intracellular cavity (Figure 1C).
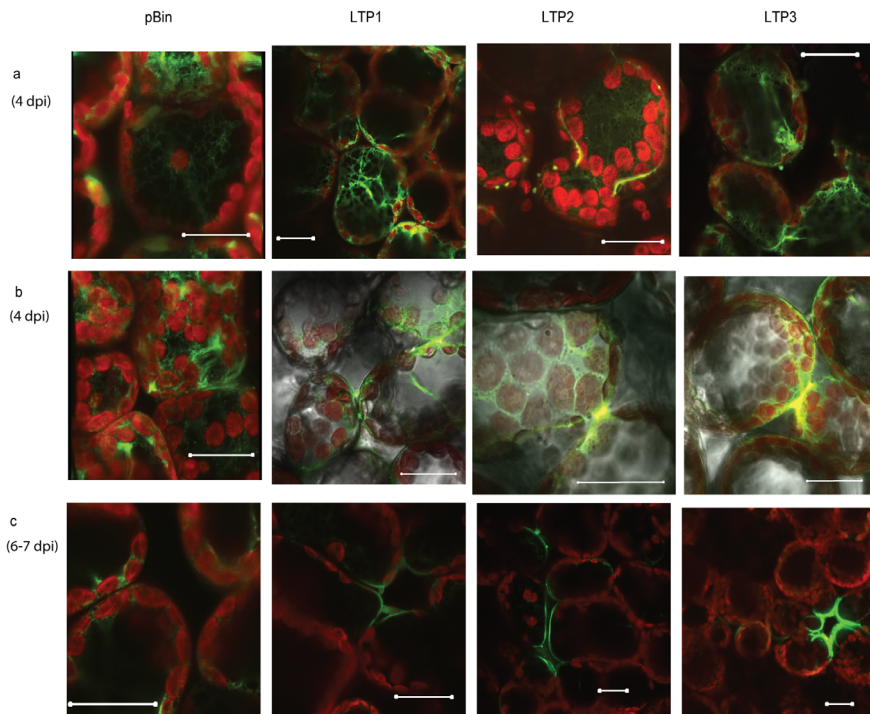
Figure 1 Subcellular localization of AaLTP1, AaLTP2 and AaLTP3. Confocal microscopy analysis of *N. benthamiana* leaves infiltrated with pBin:GFP, AaLTP1:GFP, AaLTP2:GFP and AaLTP3:GFP (a) intracellular localization of fusion proteins at 4 dpi; (b) intercellular bridges of fusion proteins between spongy mesophyll at 4 dpi; (c) fusion proteins lining cavities between mesophyll cells at 6-7 dpi. Artificial colors were given to GFP fluorescence (green) and auto fluorescence of chloroplasts (red). Bar: 20 µm.

AaLTPs+AaPDRs enhance free form AN-PW products in *N. benthamiana* leaves

We attempted to proof a putative role of AaLTP1 to 3 in the accumulation of AN-PW products by inhibiting their expression in *A. annua* using Virus Induced Gene Silencing (VIGS). However, our VIGS constructs did not dis-

play silencing activity in *A. annua* (data not shown). Therefore, we performed a functional analysis of the AaLTPs using transient co-expression with AN-PW and *AaPDR* genes in *N. benthamiana*. First, the AN-PW genes were co-infiltrated with either the three *AaLTP* genes or the two *AaPDR* genes. In all these experiments the relative dosage of agro-infiltrated AN-PW genes was kept constant, such that product accumulation was constant and the contribution of LTPs and PDRs could be accurately determined. Leaves were harvested six days post-agroinfiltration (dpi) and un-conjugated AN-PW intermediates were quantified by LC-MRM-MS as described (Ting et al., 2013). Figure 2 shows that co-expression of the AN-PW biosynthesis
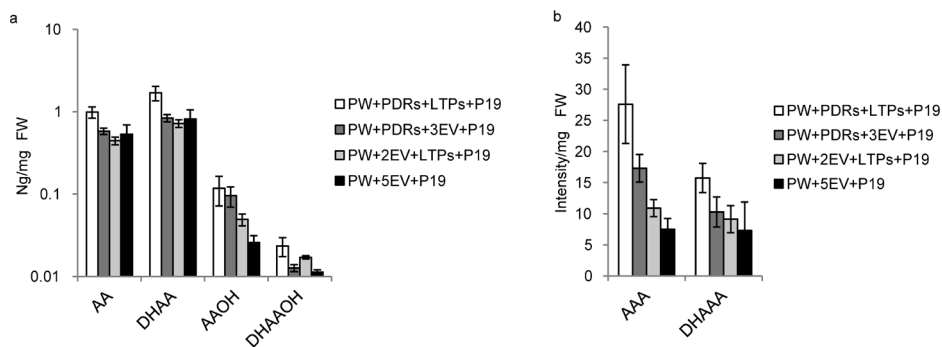
Figure 2 Artemisinin pathway intermediates and products in leaves of *N. benthamiana* infiltrated with AN-PW genes in combination with *LTPs* and *PDRs* as identified by LC-MRM-MS at 6dpi. AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; DHAAOH, dihydroartemisinic alcohol; DHAAA, dihydroartemisinic aldehyde; DHAA, dihydroartemisinic acid. 10 constructs were co-expressed in this experiment. Error bar is SE; p<0.05, n=4.

genes with the three *LTP* genes or the two *PDR* genes did not have a significant effect on the levels of (DH)AA, (DH)AAOH or (DH)AAA (Figure 2). However, the combination of AN-PW+*AaLTPs*+*AaPDRs* genes resulted in a significant increase in the accumulation of free AN-PW products compared with the pathway alone (Figure 2). LC-QTOF-MS analysis of the leaf extracts showed that the level of AN-PW glycosides and glutathione conjugates, which are more abundant than the free AN-PW products and intermediates, in the agroinfiltrated leaves, were not significantly affected by the *AaLTPs* and *AaDPRs* (Figure S3). The results suggest that AaLTPs, together with AaDPRs, can enhance sequestration of free form AN-PW intermediates away from conjugating enzyme activities. This raised the question whether these free form compounds reside inside or outside the cellsAN-PW.

AN-PW products accumulate in the apoplast of *N. benthamiana* leaves

To investigate whether the increased level of free form AN-PW intermediates and products mediated by co-expression of *AaLTPs+AaPDRs* with AN-PW is due to increased transport to the apoplast, we extracted and analyzed the apoplast. Hereto, leaves were harvested at 6 dpi and infiltrated with distilled water under vacuum. The infiltrated water was subsequently collected by mild centrifugation according to (Witzel et al., 2011). Analysis of this apoplast extract by UPLC-MRM-MS showed that AAOH, AA, DHAAOH and DHAA can be detected in the apoplast of *N. benthamiana* leaves expressing only the AN-PW genes (Figure 3). This suggests there is intrinsic transport activity for AN-PW intermediates in *N. benthamiana*. Rutin is part of the metabo-
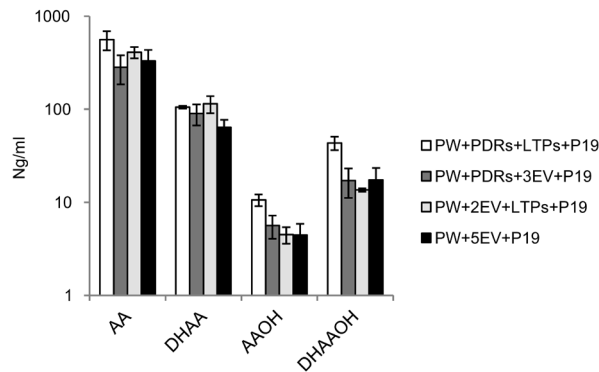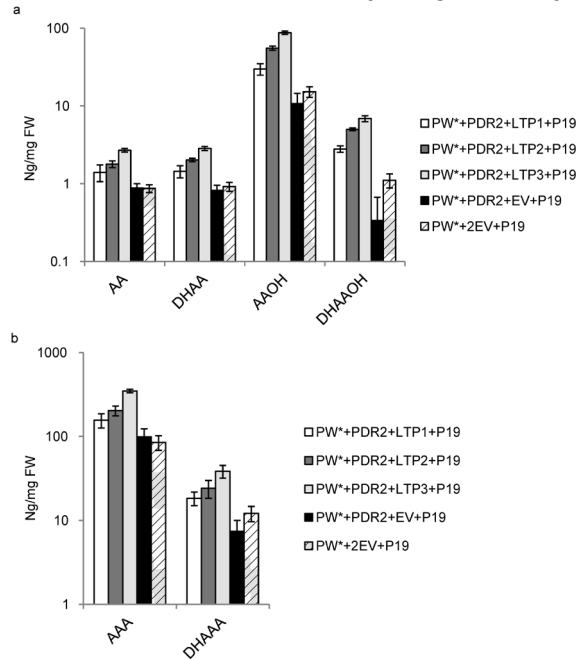

Figure 3 Artemisinin pathway intermediates and products in the apoplast of *N. benthamiana* leaves infiltrated with AN-PW genes with the *LTPs* and *PDRs* as identified by LC-MRM-MS at 6dpi. AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; DHAAOH, dihydroartemisinic alcohol; DHAAA, dihydroartemisinic aldehyde; DHAA, dihydroartemisinic acid. 10 constructs were co-expressed in this experiment. Error bar is SE; $p < 0.05$, n=4.

lome of *N. benthamiana* whole leaf extracts and it was reported to be only present in the cytoplasm and/or vacuole of plants (Markham et al., 2001; Marrs et al., 1995; Ökmen et al., 2013). As expected, rutin was not detected in the apoplast wash fluid (Figure S4). Analysis of apoplast samples by LC-QTOF-MS occasionally showed the presence of very low levels of DHAA-Hex3 (at most 5% of the total amount found in whole leaf extracts) (Figure S5). Combined, this indicates that only low level of cell damage occurred during the apoplast wash procedure and that one specific glycosylated AN-PW intermediate is transported to the apoplast at a low rate (Figure S5).

Analysis of the apoplast of leaves expressing either PW+*LTPs* or PW+*D-PRs* did not show a significant increase in AN-PW intermediates, con-

firming results from total leaf extracts (Figure 2). In contrast, expression of AN-PW+*AaPDRs*+*AaLTP* genes resulted in a significant increase in apoplastic levels of AAOH, AA, DHAAOH and DHAA (Figure 3), and DHAAA and AAA cannot detected in the apoplastHaving confirmed that the combined *LTPs* and *PDRs* can enhance apoplastic sequestration of AN-PW intermediates, we subsequently tested the specificity of the individual *LTPs* and *PDRs* in AN-PW product sequestration.

AaPDR2 is more effective in AN-PW product sequestration than AaPDR1

To determine the transport / sequestration activity of AaPDR1 and AaPDR2 in the transient expression assay in *N. benthamiana*, the AN-PW genes were co-expressed with all three *LTP* genes in combination with either *AaPDR1* or *AaPDR2*. The infiltrated leaves were harvested at 7 dpi and the apoplast was isolated and analyzed by UPLC-MRM-MS. Results show that for AA and DHAA AaPDR2 was more effective than AaPDR1 in transport to the apoplast (Figure 4). This difference in AaPDR activity could be relat-
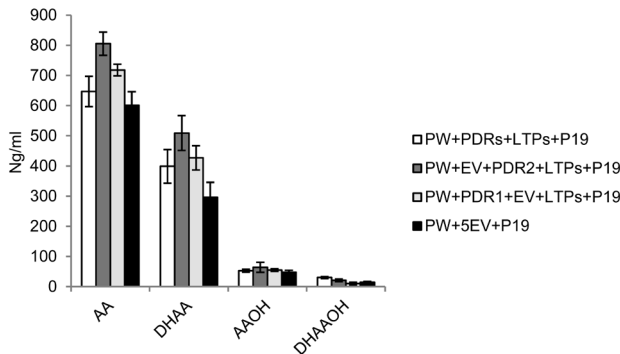
Figure 4 Artemisinin intermediates and precursors in the apoplast of *N. benthamiana* leaves infiltrated with AN-PW genes in combination with the *LTPs* and individual *PDRs,* as identified by UPLC-MRM-MS at 6dpi. AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; DHAAOH, dihydroartemisinic alcohol; DHAAA, dihydroartemisinic aldehyde; DHAA, dihydroartemisinic acid. 10 constructs were co-expressed in this experiment. Error bar is SE; p<0.05, n=4.

ed to difference in individual AaPDR structure or protein stability. Because AaPDR2 turned out to be most active, we subsequently tested the specificity of the individual *AaLTPs* in combination with the AN-PW genes and *AaPDR2*. In addition, co-expression of all LTPs and PDRs is less efficient than expressing specifically LTPs+PDR2 with the AN-PW genes as these

genes are co-expressed in the same cell in *N. benthamiana,* indicating the transport membrane dosage affects the sequestering of the metabolites as well. Transporter genes compete membrane sites, thus prevent the transport efficiency. The sequestering of (DH)AA for *PDRs + LTPs* + AN-PW is less efficient than *PDR2+LTPs*+PW at the same relative gene dosage.

AaLTP3 is most active in enhancing (DH)AA levels in *N. benthamiana*

To determine whether there is a difference in transport/sequestration activity of the individual *AaLTPs,* the AN-PW genes were co-expressed with *AaD-PR2* and the individual *AaLTP* genes in *N. benthamiana* leaves. Agro-infiltrated leaves were harvested at 7 dpi and free form AN-PW products were quantified in total leaf extracts. The results show that all three AaLTPs enhance the level of free (DH)AA compared with expression of AN-PW+*AaP-DR2* alone. However, AaLTP3 had the strongest effect on enhancing AN-PW product sequestration (Figure 5). A control experiment showed that the transcript levels of the individual *AaLTPs* and *AaPDR2* were very

Figure 5 Artemisinin pathway intermediates and products in leaves of *N. benthamiana* infiltrated with AN-PW genes with the *LTP1/2/3* and *PDR2,* as identified by LC-MRM-MS at 7dpi. AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; DHAAOH, dihydroartemisinic alcohol; DHAAA, dihydroartemisinic aldehyde; DHAA, dihydroartemisinic acid. 8 constructs were co-expressed in this experiment. Error bar is SE; p<0.01, n=4.

similar in these experiments (Figure S6), indicating that the difference in AaLTP sequestering activity must be either related to difference in individual AaLTP structure/function or difference in individual LTP protein stability.

Relative transport efficiency for artemisinin pathway intermediates

The full combination of *AN-PW+PDR2+LTP3* gave the highest AN-PW product levels in the apoplast, while *PW+PDR2*, or *PW+LTP3* only gave a small increase in AN-PW product levels in the apoplast compared with *PW+2EV+P19* (Figure 6). The ratio between the product level in the apoplast
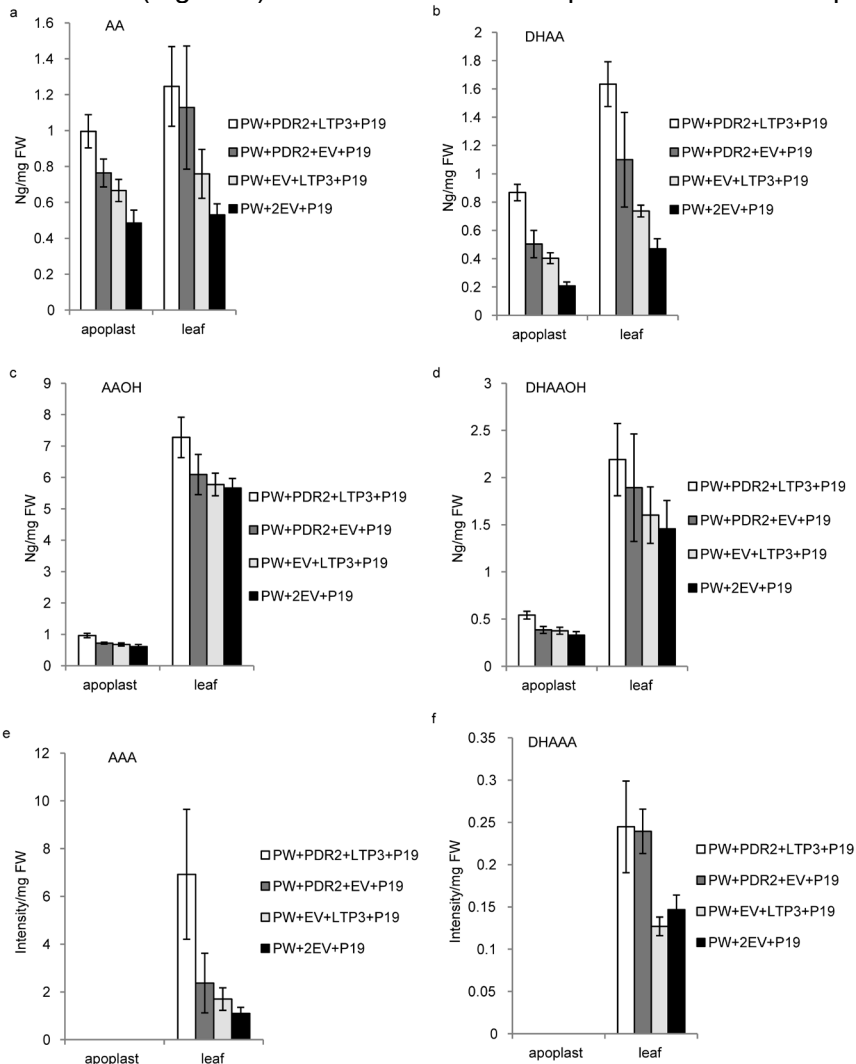
Figure 6 Distribution of artemisinin pathway intermediates and products in *N. benthamiana,* infiltrated with AN-PW genes with or without *LTP3* and *PDR2,* in the apoplast and in leaf extracts at 6dpi. 7 constructs were co-expressed in this experiment. The leaf weight before and after the apoplast extraction was used to calculate the remaining products in the apoplast and the apoplast production (ng/mg) was corrected by the apoplast concentration (ng/ml), apoplast fluid volume and remaining apoplast products. The products in leaf cells were corrected by the remaining leaf extraction and remaining apoplastic products. AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; DHAAOH, dihydroartemisinic alcohol; DHAAA, dihydroartemisinic aldehyde; DHAA, dihydroartemisinic acid. Error bar is SE; $p < 0.05$, n=5.

and in a total leaf extract was calculated for each of the AN-PW intermediates. This showed a higher specificity for transport of (DH)AA (up to 33%) to the apoplast than for (DH)AAOH (up to 10%). However, this specificity is the same as for the intrinsic transport activity in *N. benthamiana* as well as for leaves that have additional overexpression of *AaPDR2* (Figure 6). The fact that not all non-glycosylated AN-PW products are in the apoplast means some free compounds are still inside the cells, or extraction of the apoplast with water is not complete. Indeed, control experiments show that the relative transport efficiency of (DH)AA is very similar to the partitioning of these products between polar water phase and an inorganic chloroform phase (Figure S7), while the relative transport efficiency of (DH)AAOH is actually lower than the partitioning of (DH)AAOH between water and chloroform.

To determine whether the increased sequestration of free AN-PW products in the apoplast has a positive effect on total AN-PW activity we also measured the (DH)AA glycosides and glutathione conjugates using LC-QTOF-MS in leaves expressing the different combinations of AN-PW, *AaLTP3* and *AaPDR2*. These measurements show that the levels of glycosylated (DH)AAOH and glutathione AAA are lower than of the glycosylated (DH)AA products (Figure S8), consistent with a specific transport activity for the (DH)AA products. The total flux through the AN-PW is increased by *AaPDR2* and even stronger by *AaLTP3* or *AaLTP3+AaPDR2* (Figure S8).

High sequestration activity of LTP3 relates to high AaLTP3 protein stability

AaLTP3 was more effective in combination with AaPDR2 in sequestering (DH)AA than AaLTP1 or AaLTP2 at equal transcriptional activity (Figure 5). This could be due to higher specific activity of AaLTP3 towards (DH)AA or could be the result of differences in intrinsic protein stability. In order to investigate the intrinsic protein stability of the three *A. annua* LTPs we expressed AaLTP-GFP fusion proteins in *N. benthamiana* leaves and imaged

leaf fluorescence under an epi-fluorescence stereo-microscope at 5 dpi. At this stage the fluorescence of the AaLTP-GFP fusion protein was clearly above the background fluorescence of chlorophyll in the leaf samples and the AaLTP-GFP fluorescence was semi-quantified from the intensity in the green channel of RGB images taken at different positions of agro-infiltrated leaves. Results show that fluorescence of leaves expressing AaLTP3-GFP is significantly higher than that of leaves expressing AaLTP1-GFP or AaLTP2-GFP (Figure S9), while control experiments show that transcriptional activity of these genes is very similar. The higher sequestration activity of AaLTP3 therefore seems to directly relate to intrinsic stability of AaLTP3.

We also tested whether the interaction between AaLTP3 and AaPDR2 in sequestration of (DH)AA (Figure 7) is having an effect on AaLTP3 protein stability. We previously noticed that the synergistic interaction between the set of three AaLTPs +AaPDR2 for sequestration of AN-PW intermediates, coincided with an overall increase in GFP fluorescence signal when AaLTP-GFP fusion genes were used (Figure 7a). We therefore also tested whether AaLTP3 protein stability is affected by active AN-PW product transport. For this, *AaLTP3-GFP* was expressed with either *AaPDR2* alone or with the AN-PW genes or with the full set of AN-PW+*AaPDR2* (keeping the *AaLTP3-GFP* gene dosage the same between different treatments). AaLTP3-GFP fluorescence significantly increased in the presence of both *AaPDR2* and AN-PW gene expression, suggesting that AaLTP3-GFP protein levels are
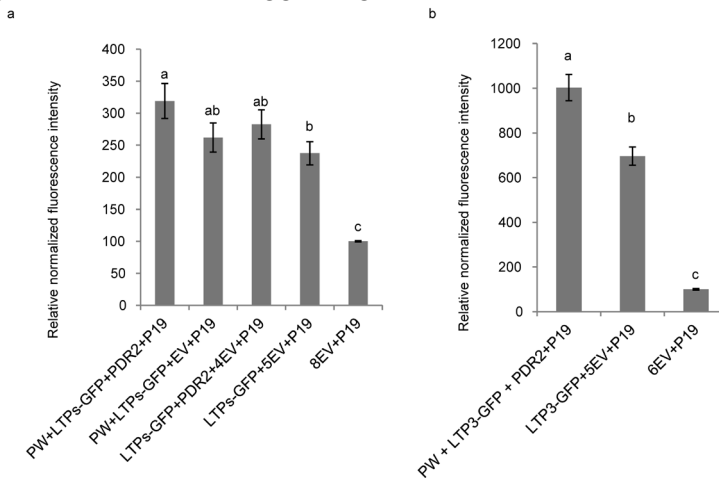
Figure 7 Relative fluorescence intensity of GFP signal in transient expression in *N. benthamiana*. A: Relative fluorescence of GFP-labelled LTPs; B: Relative fluorescence of LTP3-GFP. Relative gene dosage of LTPs-GFP was equal between all experiments by equilibrating with EV where needed. Results of two individual experiments combined (n≥54), **: P < 0.01, *: P < 0.05.

stabilized in the presence of active AN-PW product transport (Figure 7b).

Conversion of (DH)AA in *N. benthamiana* leaves is most effectively blocked by AaLTP3

So far we have shown that *AaLTP3+AaPDR2* can enhance transport of (DH) AA to the apoplast of *N. benthamina* leaves expressing the AN-PW genes. Another way by which transport and sequestration activity of proteins may be tested is by a product exclusion assay where the product is presented outside of the cells expressing the candidate proteins involved in transport. We therefore wanted to test the activity of AaLTP3 and AaPDR2 in excluding influx of (DH)AA infiltrated into the apoplast of *N. benthamiana* leaves. First we determined the fate of (DH)AA which was infiltrated into the apoplast of *N. benthamiana* leaves without transgene expression. Unexpectedly, this showed a very rapid conversion of AA and DHAA into (DH)AAA and (DH) AAOH, indicating (1) that the substrates are rapidly internalized into the leaf cells, (2) that endogenous enzymes very actively reduce the infiltrated substrates, leading to (DH)AAA and (DH)AAOH within minutes of (DH)AA infiltration) and full consumption of (DH)AA at 4 hr post apoplast infiltration (Figure 8). Next, we used *N. benthamiana* leaves at 6 dpi expressing either *EV*, *AaLTP3+EV*, *AaPDR2+EV* or *AaLTP3+AaPDR2* for infiltration of AA. We used the intracellular conversion to glycosides at two hours post apoplast infiltration as a measure of exclusion from the cells. Results show that leaves expressing EV were not able to exclude the substrate, resulting in high levels of AA glycosides and DHAA glycosides. Expression of only AaPDR2 was able to prevent ~27% of product conversion to glycosides, AaLDP3 was able to prevent ~74% of product conversion to glycosides and the combined action of AaLTP3+AaPDR2 prevented ~82% of product conversion. Combined, the results indicate that transport of compounds to the apoplast is not effective where there it is not combined with a retention system in the apoplast and that AaLTP3 can function in the retention of artemisinin products in the apoplast.

AaLTP3 + AaPDR2 enhance accumulation of artemisinin and arteanuin B

Previous attempts to detect the desired end products artemisinin (AN) or arteanniun B (AB) in *N. benthamiana* leaves expressing the AN-PW genes were unsuccessful when leaves were harvested at 7 dpi (Ting et al., 2013). However, we found that in leaves harvested at 9 dpi both AN and AB were
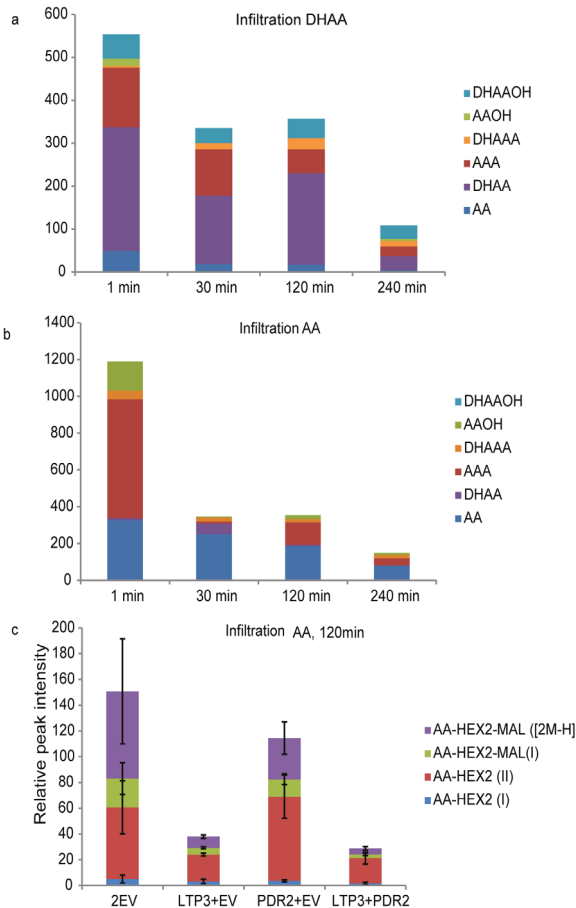
Figure 8. Bioconversion dynamics of DHAA or AA in *N. benthamiana* leaves. (a) Apoplast infiltration of DHAA (500 ng/mg) and harvest whole leaf discs at 1-240 min; (b) Apoplast infiltration of AA (500 ng/mg) and harvest at 0-4 h; Unit of AA, DHAA, AAOH, DHAAOH is ng/mg except for AAA and DHAAA (intensity) in a and b. (c) Apoplast infiltration of AA (2.67 mM) at 3 dpi of *N. benthamiana* leaves expressing AaPDR2+EV, AaLTP3+EV or AaPDR2+AaLTP3 at 3 dpi and harvest whole leaf discs 120 min after infiltration of AA into the apoplast.

detected and that levels of AN and AB increased until 13 dpi, after which AN and AB declined again (Figure 9). Both for AN and AB the accumulation was highest when the AN-PW genes were co-expressed with *AaLTP3+AaPDR2*. The accumulation of the different pathway intermediates (AAOH, AAA, AA, DHAAOH, DHAAA, DHAA) is shown in Figure S10. Results show a general trend of decrease in the alcohol (AAOH, DHAAOH) and acid products (AA, DHAA) but an increase in the aldehyde products (AAA, DHAAA) over time. The change in product profile coincides with an onset of necrosis in the agroinfiltrated leaves and raises the question wheth-

er conversion of (DH)AA to AN or AB is actually caused by leaf necrosis. The onset of necrosis after 7 dpi in leaves expressing the AN-PW genes (with or without *LTP3+PDR2*) prevents a reliable analysis of the apoplast content by the apoplast–wash method. Therefore, at present it is not known whether the accumulation of AN and AB is a consequence of continued (DH) AA transport to the apoplast between day 7-13 in agro-infiltrated leaves.
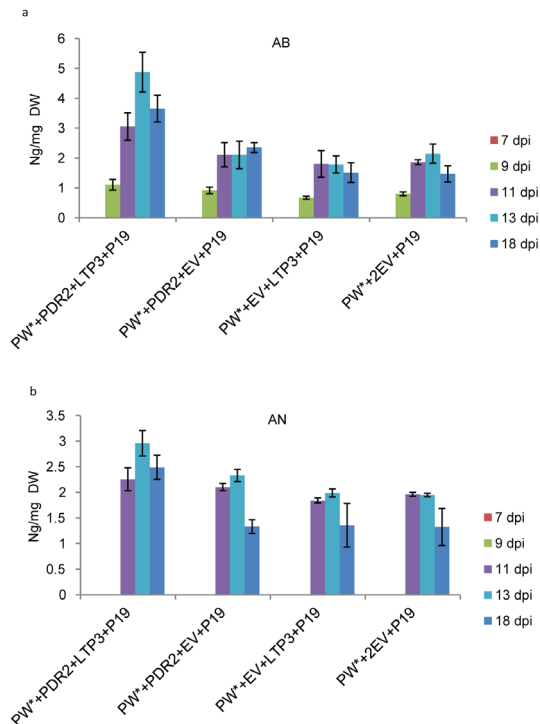
Figure 9 Production of AB and AN in leaf of *N. benthamiana* infiltrated AN-PW genes with the *LTP3* and *PDR2* as identified by LC-MRM-MS at different time points. 8 constructs were co-expressed in this experiment. AB, arteannuin B; AN, artemisinin. Error bar is SE; p<0.01, n=4.

Discussion

AaLTP3 affects the AN biosynthesis pathway activity in *N. benthamiana*

Here we have investigated the role for AaLTPs in AN-PW product accumulation using transient expression in *N. benthamiana*. We have collected mul-

tiple lines of evidence that AaLTP3 functions with AaPDR2 in the seques-
tration of AN-PW products: (1) *AaLTP3* is coexpressed with AN-PW genes
and PDR transporter genes in glandular trichomes (Table 1); (2) the *AaLTP3*
expression profile is similar to that of *AaPDR2* (Figure S1) and AaLTP3
activity on free AN-PW compound accumulation is highest in combination
with AaPDR2 (Figure 5); (3) AaLTP3 protein is extracellular (Figure 1),
while AaLTP3 enhances AN-PW product accumulation in the apoplast; (4)
AaLTP3 functions in retention of AA in the apoplast (Figure 8); (5) AaLTP3
protein stability is increased by AN-PW activity (Figure S9) suggesting
some kind of functional interaction between AaLTP3 and AN-PW activity.

The need of AaLTP3 function in *planta*

The experiments with apoplast infiltration of AA and DHAA have shown
that both of these compounds are rapidly taken up by the cells of *N. benth-
amiana* and that endogenous enzyme activity in *N. benthamiana* causes a
strong reverse flux of (DH)AA to (DH)AAA and (DH)AAOH. This is suppos-
edly caused by endogenous alcohol dehydrogenases (ADHs) or oxidore-
ductases that reduce (DH)AA. Such activity may compromise accumulation
of desired AN-PW end-products in *N. benthamiana.* A similar activity was
detected in *A. annua* by Ryden et al, who characterized an oxidoreduc-
tase Red1, expressed in *A. annua* flowers that catalyses the reduction of
DHAAA to DHAAOH (Rydén et al., 2010). This suggests that for heterol-
ogous expression systems and possibly also for *A. annua* the removal of
(DH)AA from the cytosol through transport to and retention in the apoplast
by AaLTP3 may be an important driver for the AN-PW flux towards (DH)
AA. Indeed, the specific removal of (DH)AA from the endogenous AN-
PW in *N. benthamiana* increased the overall flux through the AN-PW, as
both free (DH)AA and glycosylated (DH)AA product accumulation were in-
creased by co-expression of *PDR2* and *LTP3* (Figure 5, 6 and S8), while
there was much less effect on free and glycosylated (DH)AAOH (Figure S8).

Specificity of AaLTPs for AN-PW products

There are several lines of evidence that this effect of AaLTP3 is specific for
(DH)AA and is not caused by generation of a general (lipid) sink for sesquit-
erpenes in the apoplast: (1) in contrast to AaLTP3, AaLTP1 and AaLTP2 have
little or no effect on free (DH)AA accumulation in the apoplast. (2) AaLTP3 pro-
tein stability is increased by AN-PW activity. Finally, transient co-expression
of *LTP3* with the costunolide biosynthetic pathway from feverfew (a sesquiter-

pene PW very similar to the AN-PW; (Liu et al., 2011)) in *N. benthamiana* did not affect costunolide accumulation, suggesting specificity for (DH)AA and not any other lipid molecule (A. Beyraghdar Kashkooli, personal communication).

Putative model of AaLTP3 function

The retention activity of AaLTP3 for AA cannot be explained by a simple one to one binding of LTP3 to AA. If we assume an extreme high AaLTP3 production at 3 dpi of 1 mg/gr FW in *N. benthamiana*, the estimated Molar ratio of (AaLTP3):(infiltrated AA) is approximately 1: 635. Although at present we can only speculate on how AaLTP3 fulfills the retention activity in the apoplast, we favor the model in which AaLTP3 transports AN-PW products from the space between the plasma-membrane and cell wall to the outer side of the cell wall in the intercellular space. Once deposited at this site a reflux of (DH)AA back into the cell is then largely prevented by LTP3. The empty LTP may subsequently diffuse back to the plasma membrane to be re-loaded by another (DH)AA molecule.

Concluding remark

In conclusion, we here suggest that LTP3 plays a role in the transport of artemisinin pathway products. In combination with PDR2, LTP3 facilitates sequestration of DH(AA) into the apoplast and prevents reflux back into the cells. This newly discovered function of LTPs opens up new possibilities for the engineering of biosynthesis pathways of high value terpenes in heterologous expression systems.

Acknowledgements

References

Aziz, N., Paiva, N., May, G., and Dixon, R. (2005). Transcriptome analysis of alfalfa glandular trichomes. Planta *221*, 28-38.

Bertea, C.M., Voster, A., Verstappen, F.W.A., Maffei, M., Beekwilder, J., and Bouw-

meester, H.J. (2006). Isoprenoid biosynthesis in *Artemisia annua*: Cloning and heterologous expression of a germacrene A synthase from a glandular trichome cDNA library. Archives of Biochemistry and Biophysics *448*, 3-12.

Bouwmeester, H.J., Wallaart, T.E., Janssen, M.H.A., van Loo, B., Jansen, B.J.M., Posthumus, M.A., Schmidt, C.O., De Kraker, J.-W., König, W.A., and Franssen, M.C.R. (1999). Amorpha-4,11-diene synthase catalyses the first probable step in artemisinin biosynthesis. Phytochemistry *52*, 843-854.

Brown, G.D. (2010). The biosynthesis of artemisinin (Qinghaosu) and the phytochemistry of *Artemisia annua* L. (Qinghao). Molecules *15*, 7603-7698.

Chatzopoulou, F., Makris, A., Argiriou, A., Degenhardt, J., and Kanellis, A. (2010). EST analysis and annotation of transcripts derived from a trichome-specific cDNA library from *Salvia fruticosa*. Plant Cell Rep *29*, 523-534.

Cheng, H.-C., Cheng, P.-T., Peng, P., Lyu, P.-C., and Sun, Y.-J. (2004). Lipid binding in rice nonspecific lipid transfer protein-1 complexes from *Oryza sativa*. Protein Science : A Publication of the Protein Society *13*, 2304-2315.

Choi, Y.E., Lim, S., Kim, H.-J., Han, J.Y., Lee, M.-H., Yang, Y., Kim, J.-A., and Kim, Y.-S. (2012). Tobacco NtLTP1, a glandular-specific lipid transfer protein, is required for lipid secretion from glandular trichomes. The Plant Journal *70*, 480-491.

Covello, P.S., Teoh, K.H., Polichuk, D.R., Reed, D.W., and Nowak, G. (2007). Functional genomics and the biosynthesis of artemisinin. Phytochemistry *68*, 1864-1871.

Dong, L., Miettinen, K., Goedbloed, M., Verstappen, F.W.A., Voster, A., Jongsma, M.A., Memelink, J., Krol, S.v.d., and Bouwmeester, H.J. (2013). Characterization of two geraniol synthases from *Valeriana officinalis* and *Lippia dulcis*: similar activity but difference in subcellular localization. Metabolic Engineering *20*, 198-211.

Duke, M.V., Paul, R.N., Elsohly, H.N., Sturtz, G., and Duke, S.O. (1994). Localization of artemisinin and artemisitene in foliar tissues of glanded and glandless biotypes of *Artemisia annua* L. International Journal of Plant Sciences *155*, 365-372.

Duke, S.O., and Paul, R.N. (1993). Development and fine structure of the glandular trichomes of *Artemisia annua* L. International Journal of Plant Sciences *154*, 107-118.

Harada, E., Kim, J.-A., Meyer, A.J., Hell, R., Clemens, S., and Choi, Y.-E. (2010). Expression profiling of tobacco leaf trichomes identifies genes for biotic and abiotic stresses. Plant and Cell Physiology *51*, 1627-1637.

Huang, M.-D., Chen, T.-L.L., and Huang, A.H.C. (2013). Abundant type III lipid transfer proteins in *Arabidopsis Tapetum* are secreted to the locule and become a constituent of the pollen exine. Plant Physiology *163*, 1218-1229.

Joosten, M.A.J. (2012). Isolation of apoplastic fluid from leaf tissue by the vacuum infiltration-centrifugation technique. In Plant Fungal Pathogens, M.D. Bolton, and

B.P.H.J. Thomma, eds. (Humana Press), pp. 603-610.

Kader, J.C. (1997). Lipid-transfer proteins: a puzzling family of plant proteins. Trends in Plant Science *2*, 66-70.

Lange, B.M., Wildung, M.R., Stauber, E.J., Sanchez, C., Pouchnik, D., and Croteau, R. (2000). Probing essential oil biosynthesis and secretion by functional evaluation of expressed sequence tags from mint glandular trichomes. Proceedings of the National Academy of Sciences *97*, 2934-2939.

Littlejohn, G.R., and Love, J. (2012). A simple method for imaging *Arabidopsis* leaves using perfluorodecalin as an infiltrative imaging medium. Journal of Visualized Experiments : JoVE, 3394.

Liu, Q., Majdi, M., Cankar, K., Goedbloed, M., Charnikhova, T., Verstappen, F.W.A., de Vos, R.C.H., Beekwilder, J., van der Krol, S., and Bouwmeester, H.J. (2011). Reconstitution of the costunolide biosynthetic pathway in yeast and *Nicotiana benthamiana*. PLoS One *6*, e23255.

Lu, X., Zhang, L., Zhang, F., Jiang, W., Shen, Q., Zhang, L., Lv, Z., Wang, G., and Tang, K. (2013). AaORA, a trichome-specific AP2/ERF transcription factor of *Artemisia annua*, is a positive regulator in the artemisinin biosynthetic pathway and in disease resistance to *Botrytis cinerea*. New Phytologist *198*, 1191-1202.

Markham, K.R., Gould, K.S., and Ryan, K.G. (2001). Cytoplasmic accumulation of flavonoids in flower petals and its relevance to yellow flower colouration. Phytochemistry *58*, 403-413.

Marrs, K.A., Alfenito, M.R., Lloyd, A.M., and Walbot, V. (1995). A glutathione S-transferase involved in vacuolar transfer encoded by the maize gene *Bronze-2*. Nature *375*, 397-400.

Ökmen, B., Etalo, D.W., Joosten, M.H.A.J., Bouwmeester, H.J., de Vos, R.C.H., Collemare, J., and de Wit, P.J.G.M. (2013). Detoxification of α-tomatine by *Cladosporium fulvum* is required for full virulence on tomato. New Phytologist *198*, 1203-1214.

Olofsson, L., Engstrom, A., Lundgren, A., and Brodelius, P. (2011). Relative expression of genes of terpene metabolism in different tissues of *Artemisia annua* L. BMC Plant Biology *11*, 45.

Olsson, M.E., Olofsson, L.M., Lindahl, A.-L., Lundgren, A., Brodelius, M., and Brodelius, P.E. (2009). Localization of enzymes of artemisinin biosynthesis to the apical cells of glandular secretory trichomes of *Artemisia annua* L. Phytochemistry *70*, 1123-1128.

Rydén, A.-M., Ruyter-Spira, C., Quax, W.J., Osada, H., Muranaka, T., Kayser, O., and Bouwmeester, H. (2010). The molecular cloning of dihydroartemisinic aldehyde reductase and its implication in artemisinin biosynthesis in *Artemisia annua*. Planta Medica *76*, 1778-1783.

Schilmiller, A.L., Miner, D.P., Larson, M., McDowell, E., Gang, D.R., Wilkerson, C., and Last, R.L. (2010). Studies of a biochemical factory: tomato trichome deep expressed sequence tag sequencing and proteomics. Plant Physiology *153*, 1212-1223.

Soetaert, S., Van Neste, C., Vandewoestyne, M., Head, S., Goossens, A., Van Nieuwerburgh, F., and Deforce, D. (2013). Differential transcriptome analysis of glandular and filamentous trichomes in *Artemisia annua*. BMC Plant Biology *13*, 220.

Teoh, K.H., Polichuk, D.R., Reed, D.W., and Covello, P.S. (2009). Molecular cloning of an aldehyde dehydrogenase implicated in artemisinin biosynthesis in *Artemisia annua*. Botany *87*, 635-642.

Teoh, K.H., Polichuk, D.R., Reed, D.W., Nowak, G., and Covello, P.S. (2006). *Artemisia annua* L. (*Asteraceae*) trichome-specific cDNAs reveal CYP71AV1, a cytochrome P450 with a key role in the biosynthesis of the antimalarial sesquiterpene lactone artemisinin. FEBS Letters *580*, 1411-1416.

Ting, H.-M., Wang, B., Rydén, A.-M., Woittiez, L., van Herpen, T., Verstappen, F.W.A., Ruyter-Spira, C., Beekwilder, J., Bouwmeester, H.J., and van der Krol, A. (2013). The metabolite chemotype of *Nicotiana benthamiana* transiently expressing artemisinin biosynthetic pathway genes is a function of CYP71AV1 type and relative gene dosage. New Phytologist *199*, 352-366.

van Herpen, T.W., Cankar, K., Nogueira, M., Bosch, D., Bouwmeester, H.J., and Beekwilder, J. (2010). *Nicotiana benthamiana* as a production platform for artemisinin precursors. PLoS One *5*, e14222.

Voinnet, O., Rivas, S., Mestre, P., and Baulcombe, D. (2003). An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. The Plant Journal *33*, 949-956.

Wang, G., Tian, L., Aziz, N., Broun, P., Dai, X., He, J., King, A., Zhao, P.X., and Dixon, R.A. (2008). Terpene biosynthesis in glandular trichomes of hop. Plant Physiology *148*, 1254-1266.

Witzel, K., Shahzad, M., Matros, A., Mock, H.-P., and Muhling, K. (2011). Comparative evaluation of extraction methods for apoplastic proteins from maize leaves. Plant Methods *7*, 48.

Yeats, T.H., and Rose, J.K.C. (2008). The biochemistry and biology of extracellular plant lipid-transfer proteins (LTPs). Protein Science : A Publication of the Protein Society *17*, 191-198.

Zhang, L., Lu, X., Shen, Q., Chen, Y., Wang, T., Zhang, F., Wu, S., Jiang, W., Liu, P., Zhang, L., *et al.* (2012). Identification of putative *Artemisia annua* ABCG transporter unigenes related to artemisinin yield following expression analysis in different plant tissues and in response to methyl jasmonate and abscisic acid treatments. Plant Mol Biol Rep *30*, 838-847.

Zhang, Y., Teoh, K.H., Reed, D.W., Maes, L., Goossens, A., Olson, D.J.H., Ross, A.R.S., and Covello, P.S. (2008). The molecular cloning of artemisinic aldehyde Δ11(13) reductase and its role in glandular trichome-dependent biosynthesis of artemisinin in *Artemisia annua*. Journal of Biological Chemistry *283*, 21501-21508.

Supplementary infomation
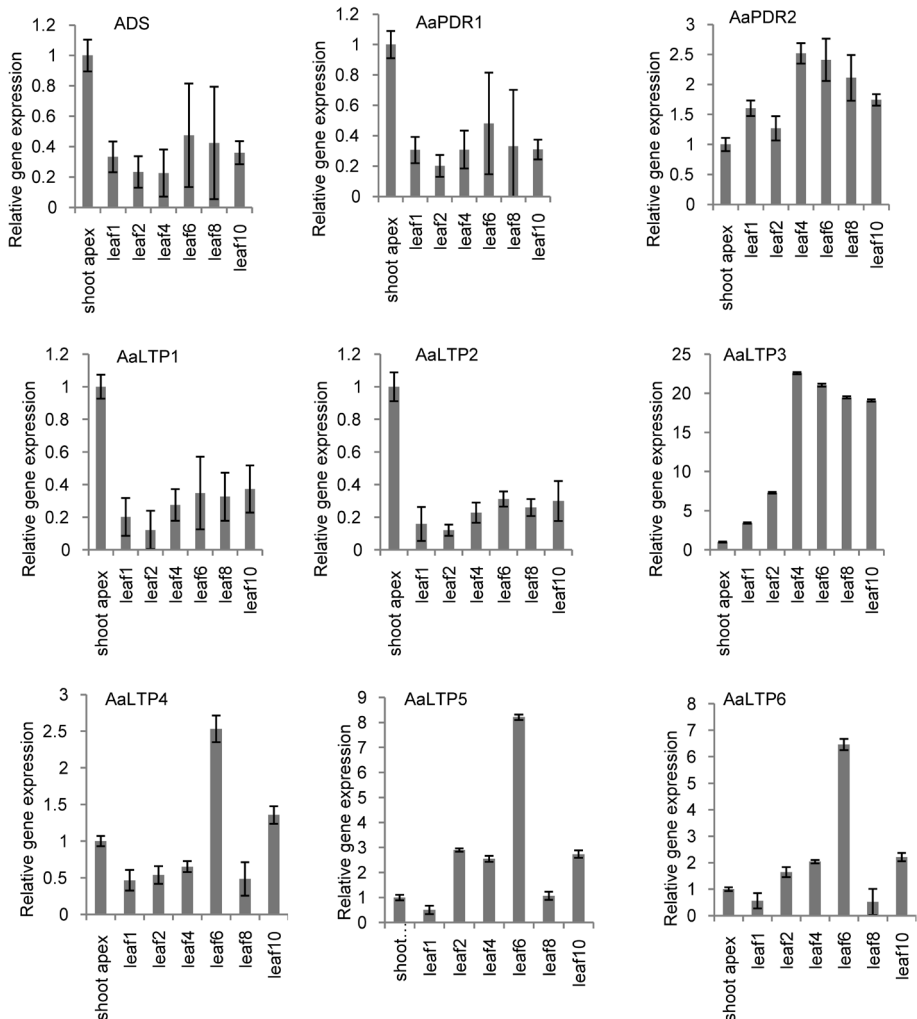
Figure. S1 Relative gene expression of *AaPDRs* and *AaLTPs* genes in shoot apex and different stages leaves of *A. annua*. Values indicate the mean fold relative to sample shoot apex, which was normalized to 1.

	pBin	LTP1	LTP2	LTP3

a
(4 dpi)

b
(4 dpi)

c
(6-7 dpi)
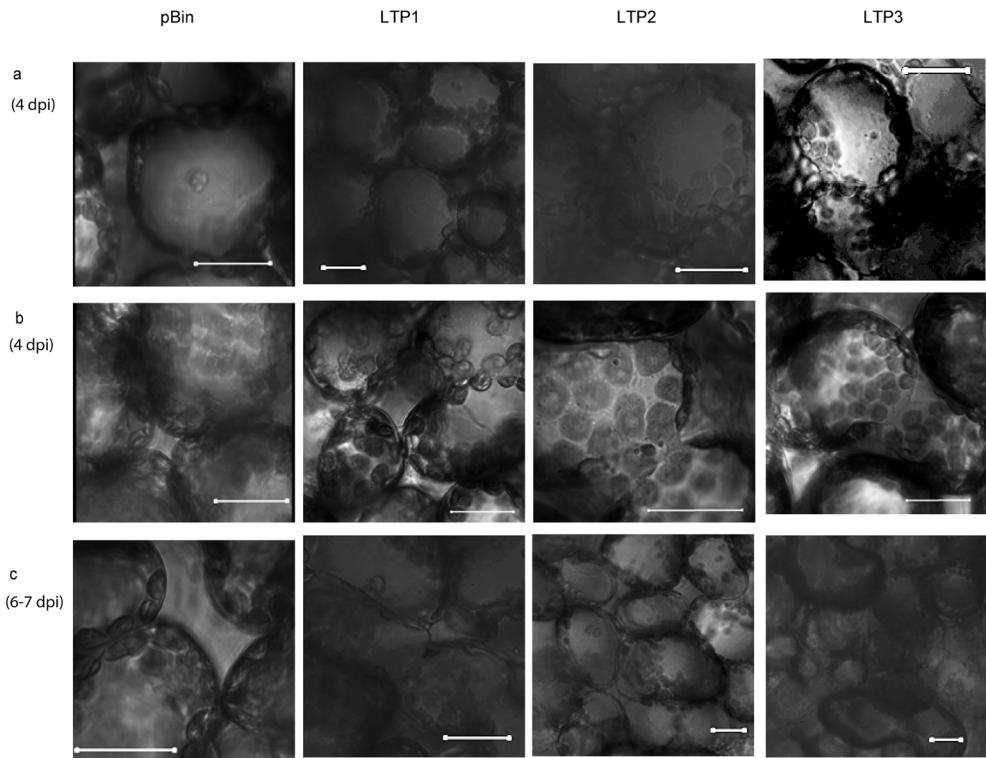
Figure S2 Bright field of AaLTP1/2/3 subcellular localization. Confocal microscopy analysis of *N. benthamiana* leaf infiltrated with *pBin:GFP*, *AaLTP1:GFP*, *AaLTP2:GFP* and *AaLTP3: GFP* separately in (A) intercellular at 4 dpi; (B) bridges between spongy mesophyll at 4 dpi; (C) cavity between the mesophyll cells at 6-7 dpi. Bar: 20 μm.
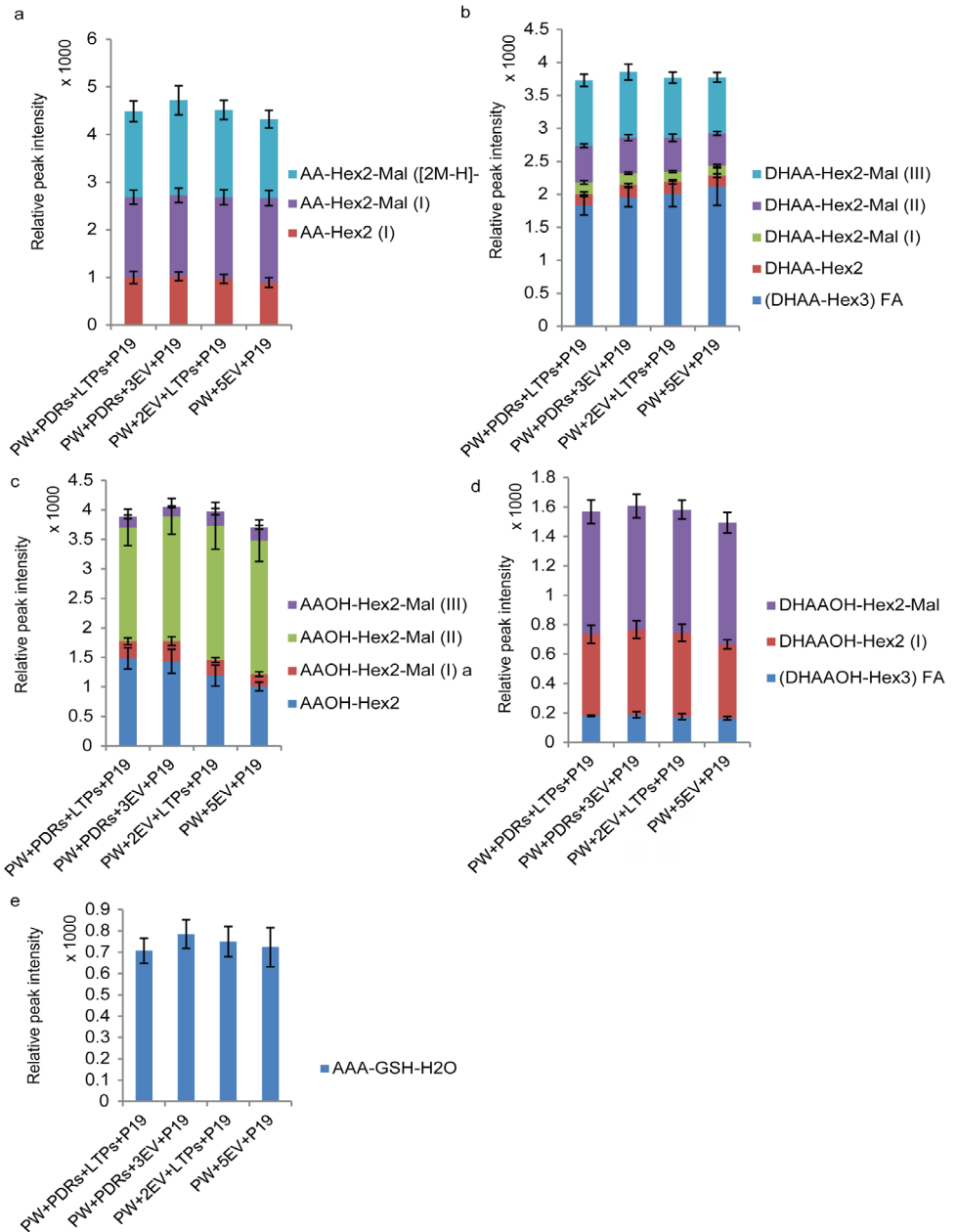
Figure S3 Artemisinin glycosides conjugations in leaves of *N. benthamiana* infiltrated AN-PW genes with the *LTPs* and *PDRs* as identified by LC-QTOF-MS at 6dpi. AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; DHAAOH, dihydroartemisinic alcohol; DHAAA, dihydroartemisinic aldehyde; DHAA, dihydroartemisinic acid. 10 constructs were co-expressed in this experiment. Error bar is SE; p<0.05, n=4.
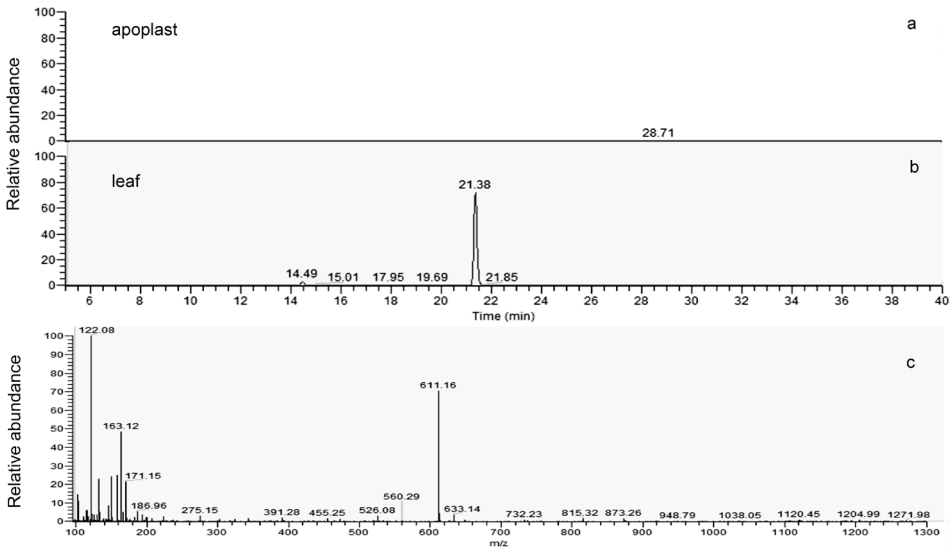
Figure S4 Detection of rutin in the apoplast samples under the LC-Orbitrap-MS. A: apoplast sample, B: remaining leaf sample after apoplast wash, C: mass spectrum of rutin.
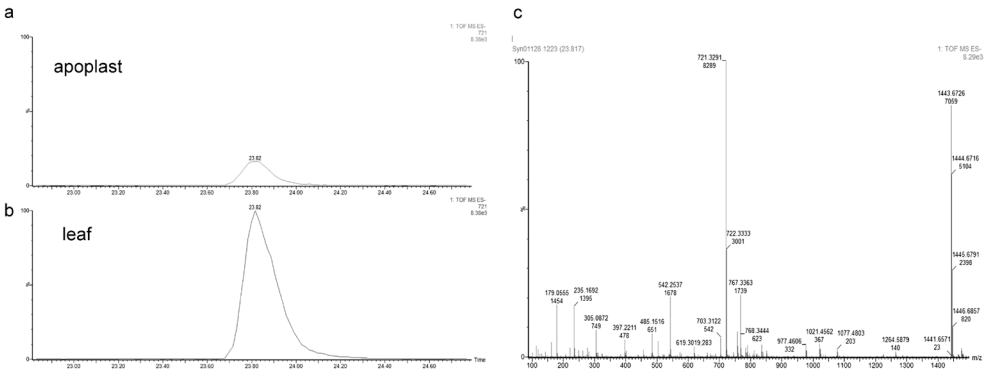
Figure S5 Detection of (DHAA-Hex3)FA in the apoplast samples under the LC-Orbitrap-MS. A: apoplast sample, B: remaining leaf sample after apoplast wash, C: mass spectrum of (DHAA-Hex3)FA. DHAA, dihydroartemisinic acid.
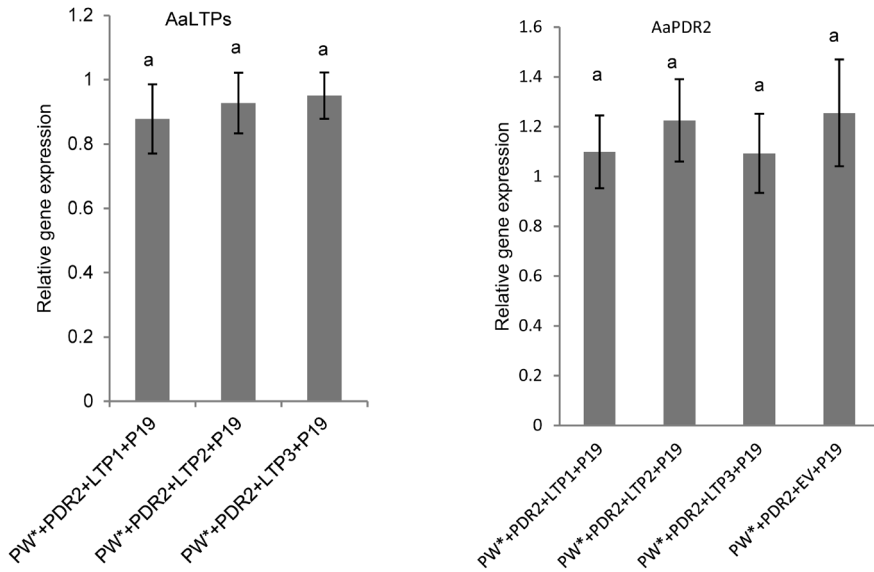
Figure S6 The relative gene expression of *AaPDR2* and *AaLTP1/2/3* in the *N. benthamiana* leaves. error bar is SE; p<0.01, n=4.
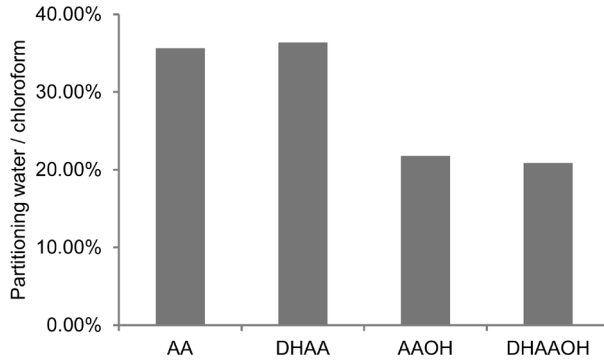
Figure S7 Partinioning of artemisinin-intermediates standards in water vs. chloroform. AAOH, artemisinic alcohol; AA, artemisinic acid; DHAAOH, dihydroartemisinic alcohol; DHAA, dihydroartemisinic acid.
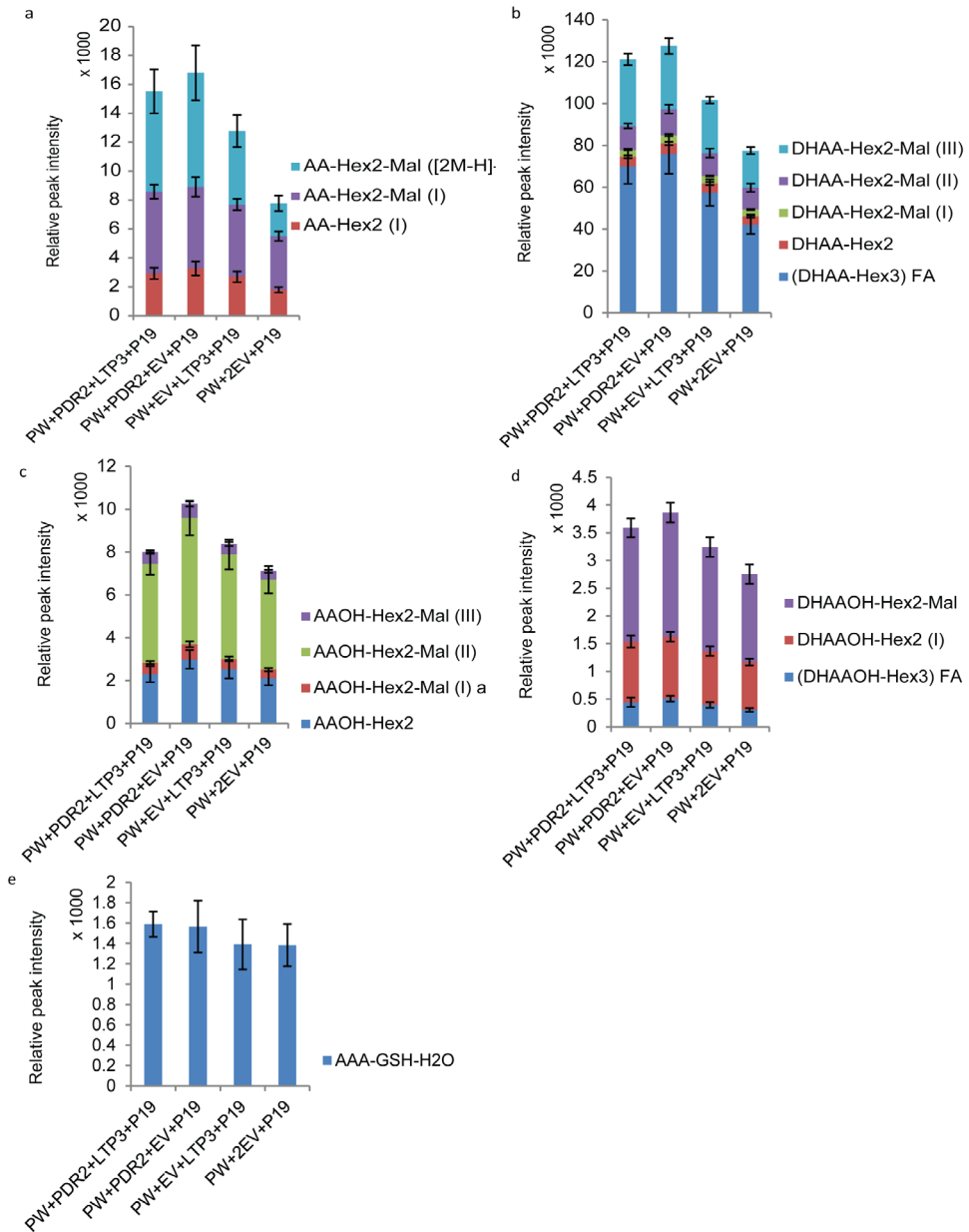
Figure S8 Artemisinin-related glycosides and glutathione conjugates production in leaf extraction of *N. benthamiana* infiltrated AN-PW genes with *LTP3* and *PDR2* as identified by LC-QTOF-MS at 6 dpi. AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; DHAAOH, dihydroartemisinic alcohol; DHAAA, dihydroartemisinic aldehyde; DHAA, dihydroartemisinic acid. 7 constructs were co-expressed in this experiment. Error bar is SE; $p<0.05$, n=5.
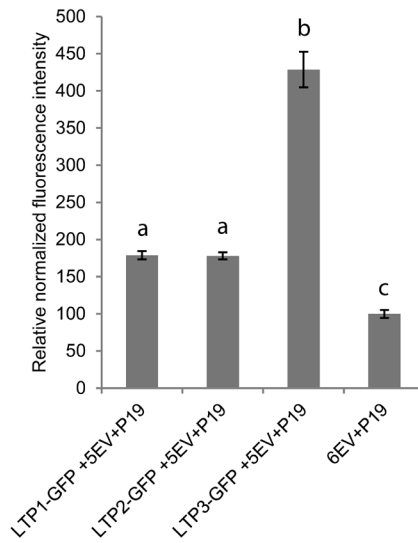
Figure S9 Intrinsic fluorescence of individual LTP-GFP, transiently expressed in *N. benthamiana*. LTP1: AaLTP1-GFP. LTP2: AaLTP2-GFP. LTP3:AaLTP3-GFP. EV: Empty Vector. **: P < 0.01
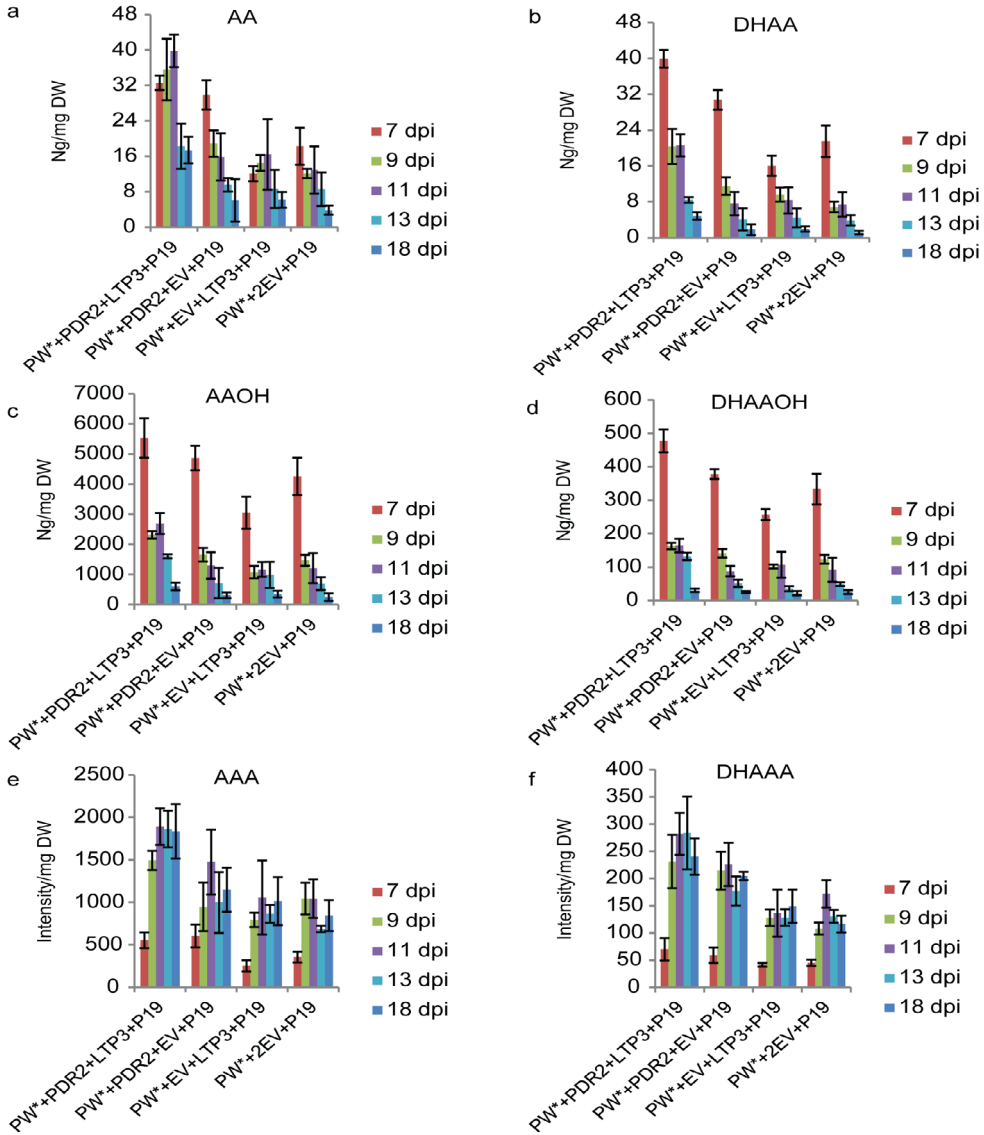
Figure S10 Production of AN-PW compounds in leaf of *N. benthamiana* infiltrated AN-PW genes with the *LTP3* and *PDR2* as identified by LC-MRM-MS at different time points. AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; DHAAOH, di-hydroartemisinic alcohol; DHAAA, dihydroartemisinic aldehyde; DHAA, dihydroartemisinic acid. 8 constructs were co-expressed in this experiment. Error bar is SE; p<0.01, n=4.

Table S1 Primers used in this paper

Primers	Sequence(5' > 3')
β-actin_Fw	CCAGGCTGTTCAGTCTCTGTAT
β-actin_Rev	CGCTCGGTAAGGATCTTCATCA
AaADS_Fwd	GGAGTATGCCCAAACCTTGA
AaADS_Rev	TCGTCTCCCATACGTGTGAA
AaLTP1_Fwd	GCAGTAGAAGGCGAGGTGAC
AaLTP1_Rev	AAGCAGCTTGTTTGTCAGCA
AaLTP2_Fwd	AAACTCCGCTGCTAAAACGA
AaLTP2_Rev	TCGGTGCTAAGGCTGATCTT
AaLTP3_Fwd	AGCACCCTATGCTGAGGCTA
AaLTP3_Rev	AGTTTGACGATCAGGGGTTG
AaPDR1_Fwd	AGAACGAGCTGCTGGAATGT
AaPDR1_Rev	GAGCAACGAGCAAAACATGA
AaPDR2_Fwd	GTTGGTTCTGGATTGGTGCT
AaPDR2_Rev	CCGAGTCTTGGTTGTTGGTT
AaLTP1-F	TTCCATGGTTGGAAAGGTTGTGTTGG
AaLTP1-R	GCGGCCGCTCATTGTATCGTTGAGCAAT
AaLTP2-F	TTCCATGGCAAGTATGACAATGAG
AaLTP2-R	GCGGCCGCTCACTTGACCTTGTTGCAGT
AaLTP3-F	TTCCATGGCAAGGATGGCAATGATTGT
AaLTP3-R	GCGGCCGCTCACTGCACCTTGGTGCAGT
EGFP_c-term-F	GAAGGAGCGGCCGCGAGCAAGGGCGAGGAGCTG
EGFP_c-term-R	CACAAAGAGCTCTTTATACAGCTCGTCCATGC
AaLTP1_gfp-F	AACACCGGATCCAATGGTTGGAAAGGTTGTGTTGGTC
AaLTP1_gfp -R	GACAAAGCGGCCGCTTGTATCGTTGAGCAATCAGTTGTGG
AaLTP2_gfp-F	AACACCGGATCCAATGGCAAGTATGACAATGAGGGTTTTATG
AaLTP2_gfp -R	GACAAAGCGGCCGCCTTGACCTTGTTGCAGTCGGTG
AaLTP3_gfp-F	AACACCGGATCCAATGGCAAGGATGGCAATGATTGTTTC
AaLTP3_gfp -R	GACAAAGCGGCCGCCTGCACCTTGGTGCAGT
AaLTP1_rfp-F	CACCATGGTTGGAAAGGTTGTGTTG
AaLTP1_rfp -R	TTGTATCGTTGAGCAATCAGTTG
AaLTP2_rfp-F	CACCATGGCAAGTATGACAATGAGG
AaLTP2_rfp -R	CTTGACCTTGTTGCAGTCG
AaLTP3_rfp-F	CACCATGGCAAGGATGGCAATG
AaLTP3_rfp -R	CTGCACCTTGGTGCAGTC

Chapter   *4*

Interaction between ectopically induced flavonoid and sesquiterpenoid biosynthetic pathways in *Nicotiana benthamiana*

Bo Wang[1], Nikolay Outchkourov [1,2], Jules Beekwilder[1,2], Harro J. Bouwmeester[1], Alexander R. van der Krol[1]

[1]Laboratory of Plant Physiology, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands.
[2]Plant Research International, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

Summary

The sesquiterpenoid biosynthetic pathway in the glandular trichomes of *Artemisia annua* is responsible for the production of the important anti-malarial compound artemisinin. Also the flavonoid biosynthesis pathway is expressed in *A. annua* glandular trichomes, however, the functional significance of this has not been studied so far. Here we investigated the interaction between these two pathways using transient expression assays in *Nicotiana benthamiana*. The endogenous *N. benthamiana* flavonoid biosynthetic pathway was stimulated by transient expression of the *Antirhinum majus* transcription factor *Rosea1* (*ROS*) and the artemisinin (AN) biosynthesis pathway (PW) by expression of AN-PW genes. Upon coexpression of AN-PW genes and *ROS* product formation through the AN-PW increased. Gene expression analysis showed that this can be attributed to stimulation by *ROS* of the expression of endogenous mevalonate kinase (*NbMVK*), which catalyzes a rate-limiting step in the mevalonate pathway which supplies the substrate for sesquiterpenoid biosynthesis. Finally, we showed that flavonoid biosynthesis competes with product formation by the AN-PW. Together these observations match with the result that highest accumulation of artemisinin was obtained by coexpression of AN-PW, *ROS*, and transporter genes with or without *CHS^RNAi*.

Key words

Flavonoids, sesquiterpenoids, artemisinin, *Artemisia annua, Nicotiana benthamiana*

Introduction

Flavonoids and terpenoids are two super families of secondary metabolites in plants and here we studied the interaction between these two major biosynthesis pathways. Flavonoids are synthesized in the phenylpropanoid biosynthesis pathway. The first committed step for flavonoid biosynthesis is catalyzed by the enzyme chalcone synthase (CHS) which combines 3 molecules of malonyl-CoA with one p-coumaroyl-CoA (formed in the shikimate biosynthesis pathway) to produce naringenin chalcone. This basic structure of naringenin chalcone can be further modified by reductases, isomerases, hydroxylases, and glycosyltransferases to make various subclasses of flavonoids, such as flavones, flavonols, anthocyanins, proanthocyanidins (Ferrer et al., 2008; Glas et al., 2012). Most of the structural genes in the flavonoid biosynthesis pathway are coordinately regulated by a ternary complex comprising three types of transcription factors (TFs); a R2R3-MYB TF, a basic helix–loop–helix (bHLH) TF and a WD-repeat (WDR) TF (Patra et al., 2013). Especially the MYB and bHLH TFs have been successfully used to ectopically induce the flavonoid biosynthesis pathway and enhance anthocyanin content in several different plant species (Outchkourov et al., 2014; Schwinn et al., 2014; Stommel et al., 2009; Tai et al., 2013). Here we used transient expression of the *Antirhinum majus* MYB TF *Rosea1* (*ROS*) to ectopically enhance production of flavonoids in *Nicotiana benthamiana*.

Terpenoids are composed of isoprene units, which are synthesized in two separate pathways: the mevalonic acid (MVA) pathway in the cytosol, and the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway in plastids (Vranová et al., 2013). These isoprene units are converted by terpene synthases into different types of terpenoids (monoterpenoids, sesquiterpenoids, diterpenoids, etc), which may again be further modified by hydroxylation, oxidation, reduction, glycosylation, acetylation or other modifying enzyme activities, which mostly occur in the cytosol and on the ER membrane. Here we used ectopic expression of the sesquiterpenoid biosynthesis pathway genes from *Artemisia annua* in *N. benthamiana* as representative of a sesquiterpene biosynthesis pathway, to study the interaction with the flavonoid biosynthesis pathway as activated by *ROS*.

Previous research in tomato has indicated a requirement of flavonoid biosynthesis for type VI trichome development, and sesquiterpene biosynthesis is located in type VI trichomes as well (Kang et al., 2014; Kang et al., 2010). Moreover, a common transcriptional control of the flavonoid and terpene biosynthesis pathway by a WRKY TF has been demonstrated in

Catharanthus roseus (Suttipanta et al., 2011) and by the Arabidopsis MYB transcription factor *PAP1* in *Rosa hybrida* (Zvi et al., 2012). In *A. annua*, the localization of flavonoid biosynthesis seems to overlap with that of sesquiterpene biosynthesis, i.e. in the glandular trichomes, as both the genes from the artemisinin biosynthesis pathway (AN-PW) and the flavonoid biosynthesis pathway were identified in a glandular trichome cDNA library (Wang et al., 2009). A common transcriptional regulation in *A. annua* has not been demonstrated, but both pathways seem to show similar activity profiles in *A. annua* as flavonoids and artemisinin (AN) reach their highest contents at the full blooming stage in the aerial parts of *A. annua* (Baraldi et al., 2008) and both accumulation of anthocyanin and AN is promoted by overexpression of the blue light receptor CRY1 (Hong et al., 2009).

It has been suggested that the co-localization of two major secondary metabolite pathways in the same glandular trichome cell may serve to optimize the function of trichome glands in a dynamic environment (Kang et al., 2014). However, at present it is not clear whether molecules of the sesquiterpene biosynthesis pathway affect flavonoid biosynthesis or vice-versa. An interaction between these two types of molecules is perhaps suggested by skin penetration studies, where terpenoids are routinely used to enhance the penetration of flavonoids (Saija et al., 1998), indicating the these molecules together may affect penetration of biological membranes.

To determine if there is an interaction between the flavonoid biosynthesis and sesquiterpene biosynthesis pathways, we expressed the AN-PW from *A. annua* (with transporters) and the flavonoid biosynthesis pathway (as ectopically activated by the *ROS* TF) alone or together by agroinfiltration of *N. benthamiana* leaves. Co-expression of the endogenous flavonoid biosynthesis pathway, as activated through ectopic expression of *ROS*, with the AN-PW genes resulted in 3-fold increase in the level of free and glycosylated artemisinic acid, while other AN-PW intermediates showed an up to 2-fold increase upon co-expression with *ROS*. This could not be attributed to differences in AN-PW gene transcription. Subsequent experiments showed that the increase in the AN-PW activity may be caused by the upregulation by *ROS* of mevalonate kinase (*NbMVK)*, which catalyzes the rate-limiting step in the mevalonate pathway which supplies substrate for AN biosynthesis. Rather than enhancing transport of sesquiterpenoids over membranes, flavonoid biosynthesis actually reduces AN-PW product levels. Indeed, experiments with CHS[RNAi] indicate a competition between the AN-PW and the *ROS*-activated flavonoid-PW. Finally, we demonstrate that the highest level of AN and arteannuin B in *N. benthamiana* are reached by ectopic expres-

sion of the *AN-PW, ROS,* and transporter genes with or without *CHS^{RNAi}*.

Materials and methods

Gene expression analysis

100mg of infiltrated fresh leaf materials were extracted and four biological replicates were prepared for RNA isolation and gene expression analysis. Real-time quantitative PCR was performed using the iQ5 RT-PCR (BioRad). The RT-PCR cycling was: 95 °C (3min); 40cycles at 95 °C (10s), 60 °C (30s) for PCR real-time acquisition; 95 °C (1min); 55 °C (1min); and 81cycles at 55 °C (10sec) for melt curve acquisition. Primers were list in table S1. Three technical replicates were run for each cDNA sample. ΔΔCT was calculated as follows: ΔCT=CT(Target)-CT(Actin), ΔΔCT was normalized using ΔCT. The fold change value was calculated by 2-ΔΔCT.

Transient expression in *N. benthamiana* leaves

Transient expression in leaves of *N. benthamiana* was done as described (Van Herpen et al., 2010). The agro-bacteria cells were collected at 3000rpm, 15min, 20°C. The dosage of each construct was resuspended and adjusted to a final OD600 of 0.5 in the agro-infiltration buffer. Individual expression constructs transformed *A. tumefaciens* strains were co-infiltrated into *N. benthamiana* leaves with 1ml syringes. The total dosage of *A. tumefaciens* between treatments was the same by dilution with empty vector in each set of experiments. Infiltrated leaves were harvested for chemical analysis at the 5 day post infiltration (dpi).

Metabolites analysis using LC- MS

Infiltrated leaves were harvested and ground to a fine powder in liquid nitrogen for metabolite analysis of infiltrated leaves in *N. benthamiana*. For the free compounds, 100mg of leaf powder were weight and extracted in 300ul methanol. After centrifuge and filtering, 10 times dilution of the supernatant with 10% methanol is performed on a Waters Xevo tandem quadrupole mass spectrometer equipped with an elect*ROS*pray ionization source and coupled to an Acquity UPLC system (Waters). For the glycosides and glutathione conjugations, 100mg of grinding infiltrated leaves were weight and extracted

in 300ul methanol with 0.137% formic acid. The samples were centrifuge at 12000rpm, 10min at 4 °C. The supernatant was performed on the LC-QTOF-MS. When necrosis was detected, the fresh samples were dried by the freezer dry after the grinding in the liquid nitrogen for the AN measurement.

Analysis apoplastic fluid using LC-MS

Apoplastic fluids were collected using the infiltration centrifugation technique with modifications (Joosten, 2012; Witzel et al., 2011). In order to get rid of the leaf necrosis, the leaves after the agro-infiltration were harvested at 6 or 7 dpi for the apoplast wash. For the vacuum infiltration, the fresh leaves were cut into 5 cm strips and stack in a beaker filled with demi water. A weight was placed on the leaves to prevent the leaves from floating. The vacuum was carried out 10 min in a desiccator using a vacuum pump. To prevent the cell broken, the vacuum was released slowly until the leaf become water-soaked and dark in colour. For the apoplast collection, the centrifuge tubes consist two part, the bottom part is a 50 ml plastic centrifuge tube (Greiner bio-one); the upper part has a perforated bottom and loosely fits on the bottom part. The vacuum infiltrated leaves was placed inside the upper centrifuge tube and centrifuged 400g, 15min at 4°C to collect the fluid in the bottom centrifuge tube. After the extraction of the apoplastic fraction, the residual leaf tissue was shock frozen in liquid nitrogen for the metabolites analysis. The apoplast fluids were centrifuged and filtered with 0.45µm PTEE-membrane filter. The filtered fluid was dilute into 4 times with 10% MeOH. The diluted apoplast fluid was injected into the LC-MS.

Electrolyte leakage assay

The electrolyte leakage assay was carried out as described (Bouwmeester et al., 2014). 8mm Leaf discs were punched from agro-infiltrated *N. benthamiana* leaves. Six leaf discs were floated on 4mL MQ water at room temperature under 250 rpm shaker for 2 hours. Four biology replicates were prepared. Conductivity was measured using a Mettler Toledo InLab 741 ISM conductivity probe (Mettler Toledo, Tiel, Netherlands). Then, samples with the leaf discs were autoclaved and measured for conductivity as well. Relative electrolyte leakage ratios were calculated by dividing conductivity values of samples before autoclaving with the conductivity values of samples after autoclaving.

Results

Co-expression of *ROS*+*DEL*+AN-PW genes leads to leaf necrosis

The artemisinin biosynthesis pathway (AN-PW) gene set was used for ecto-pic expression in *N. benthamiana* (N-terminal truncated 3-Hydroxy-3-Meth-ylglutaryl Coenzyme A Reductase (*tHMGR*)*,* farnesyl diphosphate synthase (*FPS*)*,* amorphadiene synthase (*ADS*), amorphadiene oxidase (*AMO*), ar-temisinic aldehyde double-bond reductase (*DBR2*), aldehyde dehydroge-nase 1 (*ALDH1*)) (Bouwmeester et al., 1999; Rydén et al., 2010; Teoh et al., 2009; Teoh et al., 2006; Ting et al., 2013; Zhang et al., 2008), while p19 was co-infiltrated to suppress possible silencing of transgenes (Voinnet et al., 2003). For activation of the flavonoid biosynthesis pathway in *N. benth-amiana* we tested transient expression of the *A. majus* transcription factors *ROS* (MYB) and *DEL* (bHLH) (Outchkourov et al., 2014). However, com-bined ectopic expression of the AN-PW genes with the *ROS* and *DEL* TFs resulted in necrosis within 3 days post agro-infiltration (dpi) in *N. benthami-ana* (Figure 1A). Previous results indicated that *ROS* is sufficient to induce
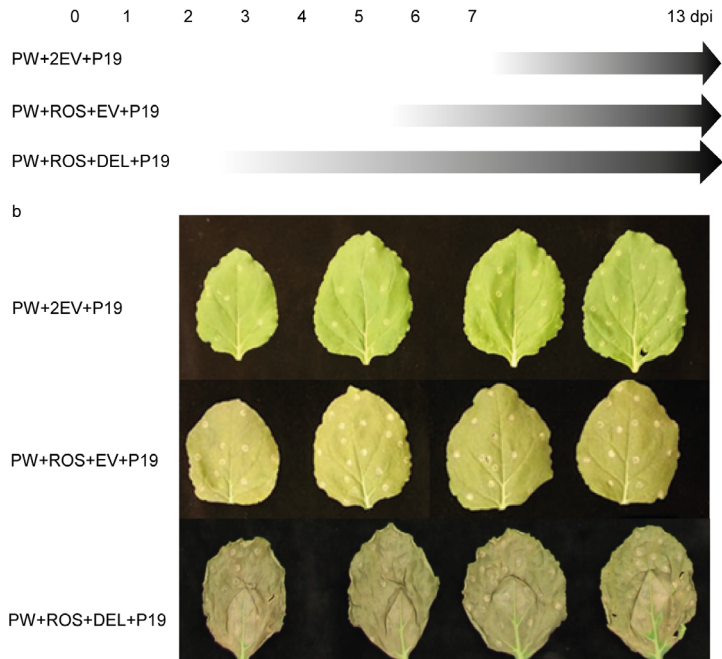
Figure 1. Necrosis in *N. benthamiana* expressing artemisinin pathway (PW), *Ro-sea1* (*ROS*) and *Della* (*DEL*). a, time course onset of necrosis symptom, the black ar-row indicates the necrosis; b, leaves infiltrated with different gene combination at 5 dpi.

low level accumulation of flavonoids, while DEL is only required to enhance functions of *ROS* upon transient expression in *N. benthamiana* (Outchkourov et al., 2014). Subsequently ectopic expression of only *ROS* with the AN-PW gene set was tested, in which case the onset of necrosis was delayed until about 6 dpi (Figure 1B). Leaves expressing AN-PW+*ROS*+*P19* in *N. benthamiana* showed accumulation of red anthocyanins at 3 dpi and only mild necrotic symptoms (Figure 1B). Therefore, in the following experiment the harvest time was set at 5 dpi for AN-PW product analysis and at 13 dpi for measurement of the final products artemisinin (AN) and arteannuin B (AB).

Co-expression of *ROS* + AN-PW genes enhances AN sesquiterpene biosynthesis

AN-PW genes were co-expressed with or without *ROS* in *N. benthamiana*, such that the relative dosage of the AN-PW genes was constant in the two treatments. Leaves were harvested at 5 dpi and the freeform AN intermediates were extracted in methanol and extracts were analyzed by the LC-trip-quad-MS as described (Ting et al., 2013). Figure 2a shows that both AA
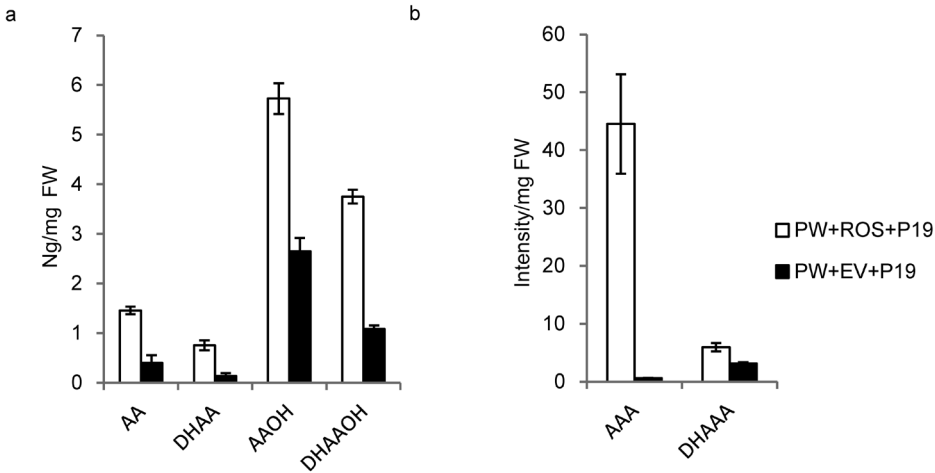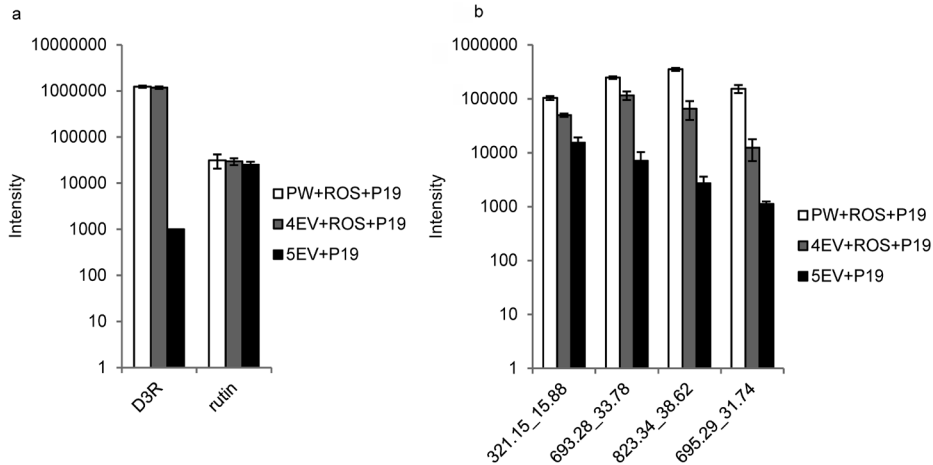

Figure 2. AN and its precursors in leaves of *N. benthamiana* infiltrated with artemisinin pathway (PW) and *Rosea1* (*ROS*) and analyzed by LC-triple-quad-MS at 5 dpi (a and b). AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; DHAAOH, dihydroartemisinic alcohol; DHAAA, dihydroartemisinic aldehyde; DHAA, dihydroartemisinic acid. Error bar is SE, n=4.

and DHAA production were significantly stimulated by the co-expression with *ROS*. Since we did not have proper standards for (DH)AAA, we used the (DH)AAA peak intensities for semi-quantification. AAA production was also strongly stimulated by co-expression with *ROS*, while *ROS* had only limited
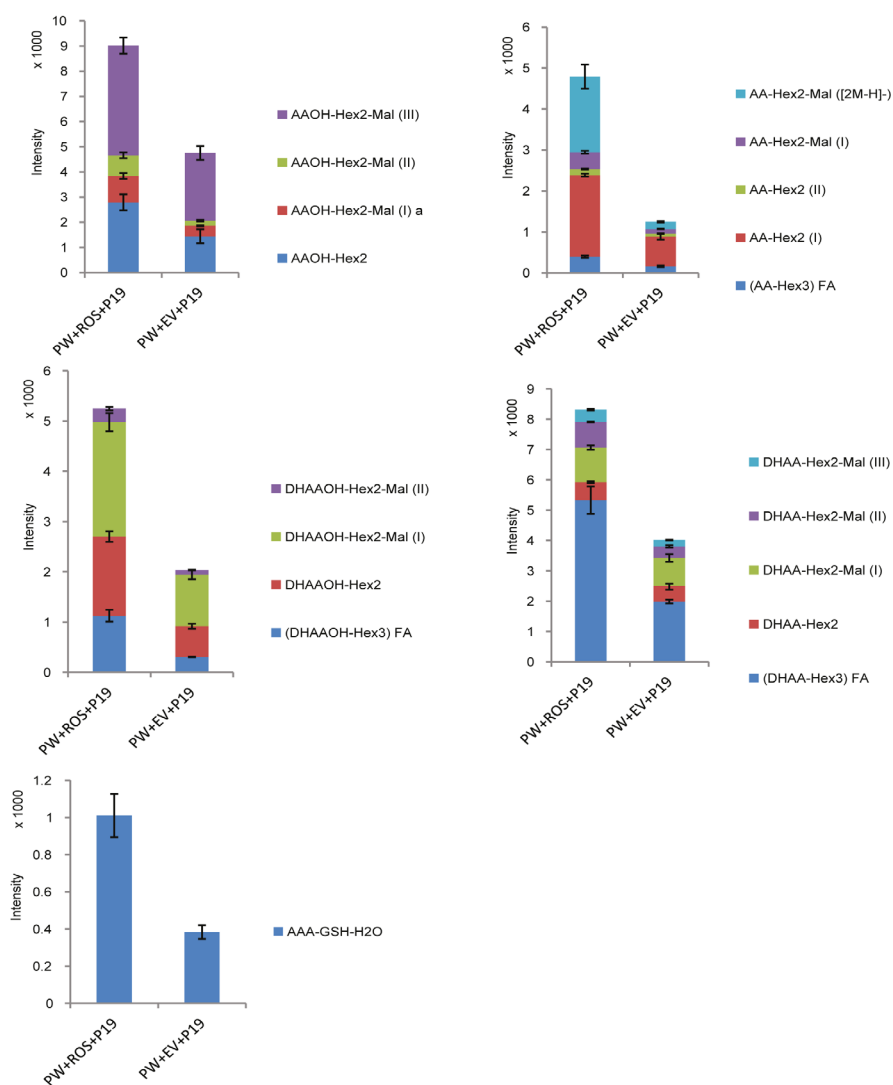
Figure 3. AN related glycosides and glutathione conjugates in leaves of *N. benthamiana* infiltrated with artemisinin pathway (PW) and *Rosea1* (*ROS)* and analyzed by LC-QTOF-MS at 5 dpi. AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; AB, arteannuin B; DHAAOH, dihydroartemisinic alcohol; DHAAA, dihydroartemisinic aldehyde; DHAA, dihydroartemisinic acid. Error bar is SE, n=4.

effect on accumulation of DHAAA (Figure 2b). In order to test the effect of *ROS* on the glycosides and glutathione conjugates of AN-PW intermediates, leaf tissue was extracted with methanol with 0.137% formic acid and extracts were analyzed by LC-QTOF-MS (Ting et al., 2013). Similar to the free AN-PW product levels, also the AN-PW related glycosides and glutathione conju-

gates showed a 2- to 3-fold increase upon co-expression with *ROS* (Figure 3). Combined, these results show that *ROS* enhances all AN-PW product levels.

AN-PW has little effect on major *ROS*-induced flavonoids

The effect of AN-PW products on flavonoid PW accumulation was quantified by comparing *ROS*+5EV+P19 with *ROS*+AN-PW+P19 and 2EV+P19, in which the *ROS* gene dosage was kept constant in two of the treatments. For testing the effect of AN-PW activity on flavonoids we analyzed the flavonoid rutin and a single anthocyanin, delphinidin-3-rutinoside (D3R). Previously it was shown that ectopic *ROS* expression in *N. benthamiana* results in the production of D3R and a number of unknown flavonoid products (Outchkourov et al., 2014). The concentration of rutin was not affected by *ROS*

Figure 4. Flavonoids in leaves of *N. benthamiana* infiltrated with artemisinin pathway (PW) and *Rosea1* (*ROS*) and analyzed by LC-triple-quad-MS at 5 dpi (a and b). a, major flavonoids rutin and delphinidin-3-rutinoside (D3R); b, minor unknown putative flavonoid compounds induced by ROS. Error bar is SE, n=4.

expression alone or *ROS*+AN-PW (Figure 4). However, the levels of D3R were greatly enhanced by the ectopic expression of *ROS*, while additional expression of AN-PW did not have an effect. In contrast, accumulations of the unknown compounds that are induced by *ROS* in *N. benthamiana* leaves were slightly enhanced by co-expression of *ROS*+AN-PW. Combined, the results show that *ROS* stimulates accumulation of AN-PW products, but that the AN-PW has no or little effect on accumulation of major flavonoids in *N. benthamiana* leaves, only minor compounds induced by *ROS* are slightly en-

hanced by AN-PW activity. Subsequently, we addressed the question wheth-
er the effect of *ROS* on AN-PW product accumulation is due to improved
transport of the AN-PW products to the apoplast of *N. benthamiana* leaves.

ROS activity mainly enhances sequestration of AN-PW free products inside the cells

To investigate whether the increased concentration of freeform AN-interme-
diates upon co-expression of *ROS* and AN-PW was due to storage in the
apoplast, we analyzed the apoplast fluid of agro-infiltrated leaves for free
(non-glycosylated) AN-PW product accumulation and compared this to the
level of free AN-PW compounds in total extracts of the leaves (see Chap-
ter3). Because at 5 days there were already signs of necrosis, which could
lead to contamination of apoplast fluid with cell content, the apoplast assay
was performed at 4 dpi. Apoplastic AA and DHAA only increased 1.5-fold
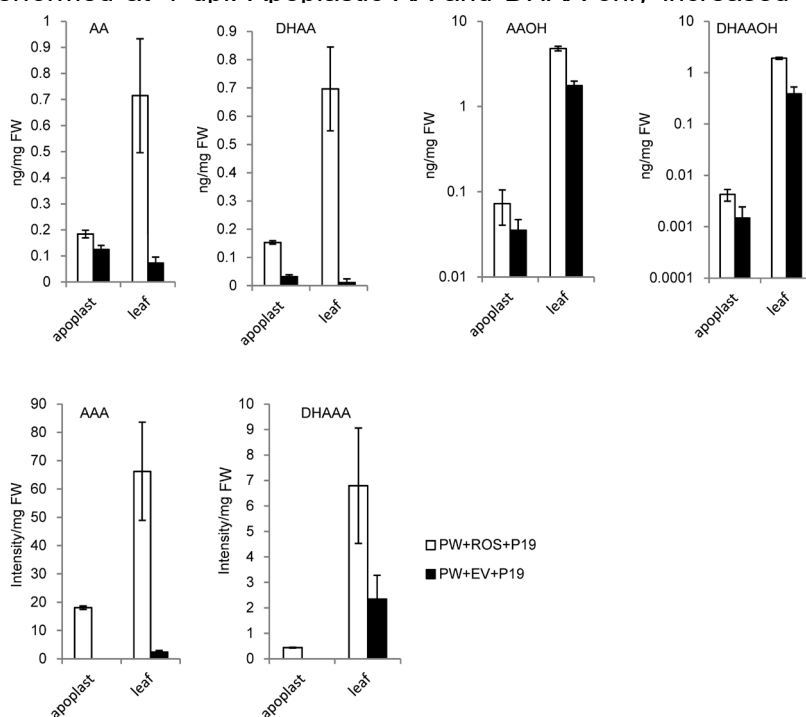
Figure 5. artemisinin pathway (PW) intermediate distribution in apoplast and leaf cells
of *N. benthamiana* infiltrated with PW and *Rosea1* (*ROS)* and analyzed by LC-tri-
ple-quad-MS at 5 dpi. AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, ar-
temisinic acid; AB, arteannuin B; DHAAOH, dihydroartemisinic alcohol; DHAAA, di-
hydroartemisinic aldehyde; DHAA, dihydroartemisinic acid. Error bar is SE, n=4.

and 4.8-fold respectively, while inside the leaf AA and DHAA increased 9.8-fold and 6.2-fold respectively, when AN-PW genes were co-expressed with *ROS* (Figure 5). The effect of *ROS* on the free form of AAOH and DHAAOH was less pronounced with only a small increase in the apoplast and total leaf extract concentration (Figure 5). Expression of only AN-PW genes in *N. benthamiana* leaves did not result in detectable levels of AAA or DHAAA in the apoplast, but upon coexpression of AN-PW with *ROS,* both AAA and DHAAA could be extracted from the apoplast. Rutin has been reported to reside in the cytosol and vacuole, and can be used to test for cell leakage (Markham et al., 2001; Marrs et al., 1995; Ökmen et al., 2013). Control measurements of apoplast fluid indicated that no rutin is present, indicating that (DH)AAA in the apoplast is not from leaking cells (Figure S1). Overall, the results indicate that ectopic expression of *ROS* may affect sequestration of free AN-PW products inside the cells, but that *ROS* has little effect on transport capacity of (DH)AA. However, *ROS* does seem to enhance transport of (DH)AAA to the apoplast.

ROS does not alter transcriptional activity of AN-PW genes in transient expression assays

Because *ROS* enhanced AN-PW product formation, this raised the question whether transcription of the different AN-PW transgenes (which are all under control of the CaMV 35S promoter) is somehow stimulated by *ROS*. Therefore the transcript levels of AN-PW genes were measured by qRT-PCR, but this showed that *ROS* does not affect their transcription (Figure 6). Alternatively, the overall increased AN-PW product level in the presence of *ROS* could be due to effects of *ROS* on endogenous *N. benthamiana* genes that encode enzymes involved in the biosynthesis of precursors for the sesquiterpene pathway. Previously, the transcriptional response to ectopic expression of *ROS* in tomato callus was analysed by RNAseq (Outchkourov et al., unpublished). We screened these tomato RNAseq data for all known genes related to either the MVA, MEP, terpenoid, flavonoid and shikimate pathways (Figure 12) (Dudareva et al., 2013; Falcone Ferreyra et al., 2012; Maeda and Dudareva, 2012; Petrussa et al., 2013; Saito et al., 2013; Vranová et al., 2013), to determine if any of these genes are affected by *ROS*. This analysis showed that in tomato callus the expression of a geranylgeranyl diphosphate synthase (*GGPPS*, Solyc07g064660) was downregulated 68-fold upon expression of *ROS,* while a mevalonate kinase (*MVK*, Solyc01g09884) was upregulated 2.2-fold. Downregulation of the GGPPS could potentially reduce competition for substrate flux towards sesquiterpene biosynthesis, while upregulation of MVK could potentially enhance the flux towards sesquiterpene biosynthesis

(Hinson et al., 1997). Therefore, if similar genes in *N. benthamiana* are affected by *ROS*, this could explain the overall higher AN-PW product level as result of AN-PW+*ROS*. The expression of the homologs of GGPPS and MVK in *N. benthamiana* was quantified by rt-PCR in the leaf samples expressing either AN-PW or AN-PW+*ROS*. Results show that *NbMVK* is indeed enhanced 3.0-fold by *ROS* (Figure 6), while *NbGGPPS* transcripts were not detected in leaves (not shown). Although HMGR was reported to be the rate-limiting step in the MVA pathway (Chappell et al., 1995; Stermer and Bostock, 1987; Tholl and Lee, 2011), *NbHMGR2* was not up-regulated by *ROS* activity in *N. benthamiana* (Figure 6). Thus, boosting of AN-PW products by *ROS* expression could be the direct result of boosting of *NbMVK* gene activity by *ROS*.
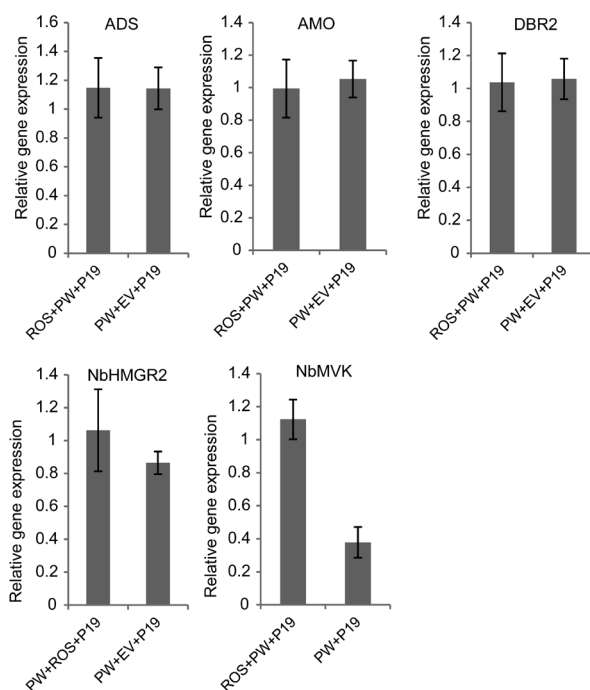
Figure 6. Relative gene expression of amorphadiene synthase *(ADS)*, amorphadiene oxidase *(AMO)*, artemisinic aldehyde double-bond reductase *(DBR2)* and mevalonate kinase (*NbMVK)* in leaves *of N. benthamiana* infiltrated with artemisinin pathway (PW) and *ROS*ea1 (*ROS)* and analyzed at 5 dpi. Gene expression analysis was carried out with 4 biological replicates and each of the biological replicates has three technical replicates. Error bar is SE, n=3.

Blocking flavonoid biosynthesis induced by *ROS* through CHS[RNAi] enhances flux through the AN biosynthesis pathway

Besides an effect of *ROS* on endogenous gene activity, which could affect flux through the AN-PW, it is also of interest to determine if flavonoids themselves affect AN-PW product accumulation. To test for this we expressed the AN-PW+*ROS* with and without a construct which blocks chalcone synthase (*CHS*), the first enzymatic step of the flavonoid pathway. *CHS* transcription was silenced by a *CHS[RNAi]* construct (Schijlen et al., 2007). Analysis of the *ROS*-induced flavonoid D3R and the endogenous flavonoid rutin, showed that the *CHS[RNAi]* construct was very effective in suppression of D3R and had only limited effect on rutin (Figure 7 and Figure S2). The silencing of flavonoid biosynthesis (in the presence of *ROS*) resulted in an increase in both freeform and glycosylated AN-PW related compounds (Figure 7 and Figure 8), indicating that activity of the flavonoid pathway somehow competes with the sesquiterpene pathway in cells. We note that the ratio of different AN-PW related glycosides is different for *N. benthamiana* leaves expressing AN-PW+*ROS* (Figure 3) and AN-PW+*ROS*+*CHS[RNAi]* (Figure 8), indicating that suppression of the flavonoid biosynthesis alters the AN-PW related leaf chemotype.
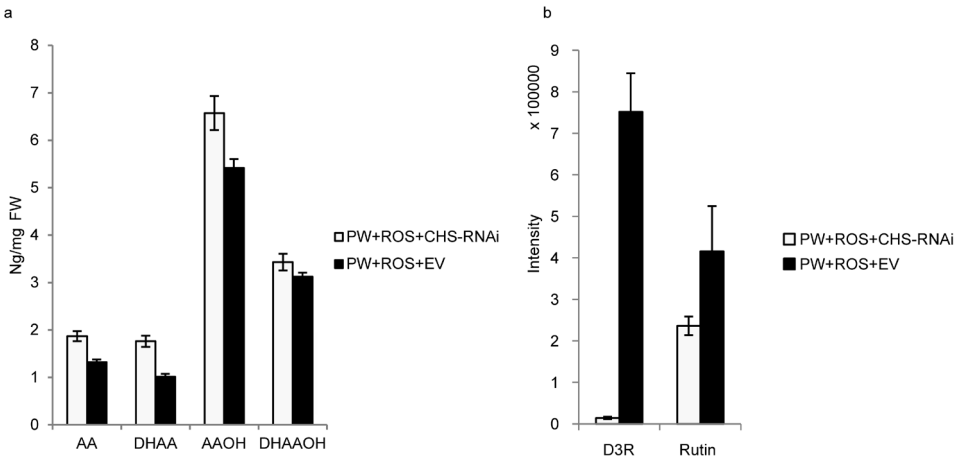
Figure 7. Targeted analysis of extracts of *N. benthamiana* leaves infiltrated with artemisinin pathway (PW) and *Rosea1 (ROS)* with and without *CHS[RNAi]* and analyzed by LC-triple-quad-MS at 5 dpi. a, AN intermediates; b, flavonoids. AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; AB, arteannuin B; DHAAOH, dihydroartemisinic alcohol; DHAAA, dihy droartemisinic aldehyde; DHAA, dihydroartemisinic acid. Error bar is SE, n=4.
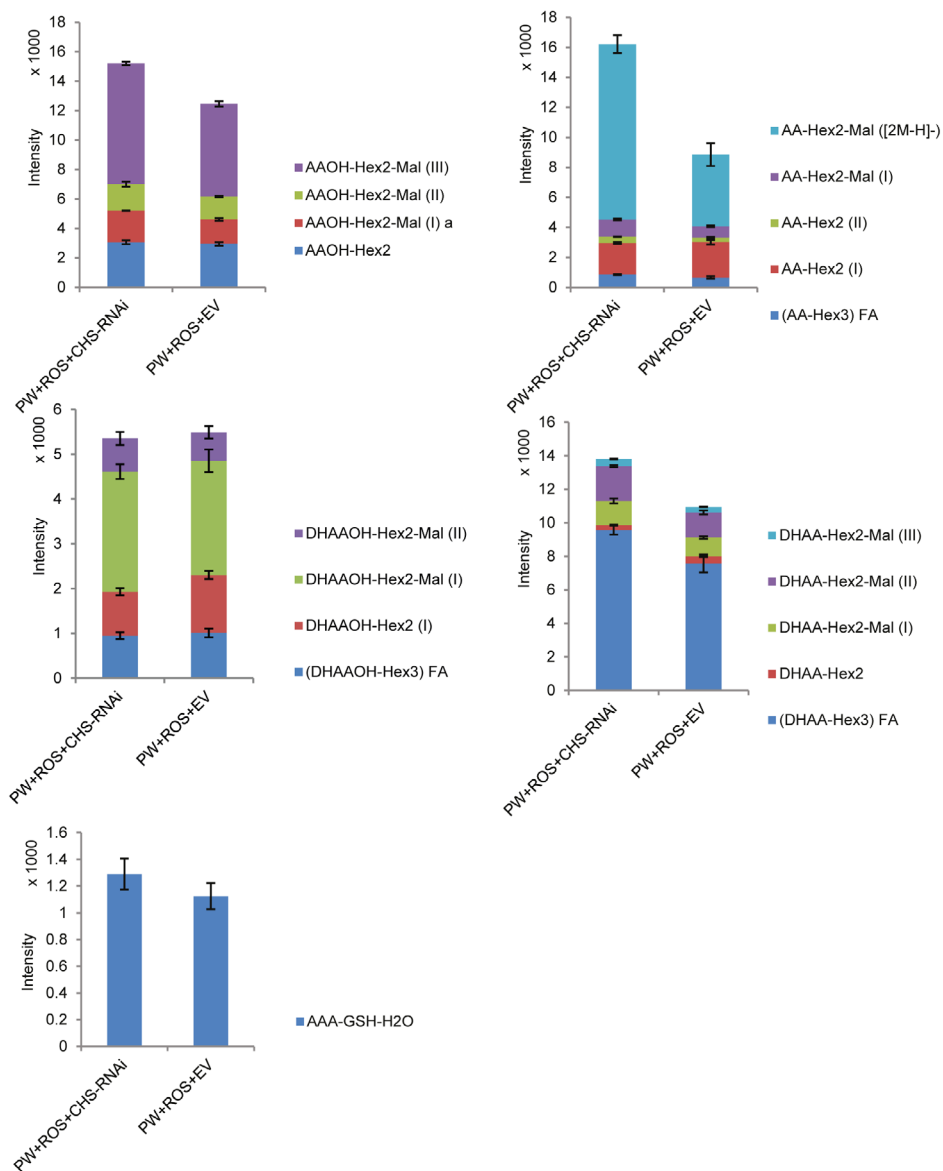
Figure 8. AN glycosides and glutathione conjugates in leaves of *N. benthamiana* infiltrated with artemisinin pathway (PW), *Rosea 1* (*ROS*) and CHS[RNAi] and analyzed by LC-triple-quad-MS at 5 dpi. AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; AB, arteannuin B; DHAAOH, dihydroartemisinic alcohol; DHAAA, dihydroartemisinic aldehyde; DHAA, dihydroartemisinic acid. Error bar is SE, n=4.

AaLTP3 and AaPDR2 have no influence on the effect of ROS on AN-PW product accumulation, but alter the glycoside chemotype of N. benthamiana leaves.

Because previously we have shown that AaPDR2 and AaLTP3 enhance sequestration of free AN-PW products in the apoplast (Chapter3), we subsequently tested the effect of *ROS* on AN-PW product accumulation in the presences of AaPDR2 and AaLTP3. To this end, the product profile in leaves agro-infiltrated with AN-PW+*ROS*+*AaPDR2*+*AaLTP3* was compared with that of AN-PW+*EV*+*AaPDR2*+*AaLTP3*. Note that, due to the additional two gene constructs (35S::AaPDR2 and 35S::AaLTP3) in these experiments the relative dosage of AN-PW genes was slightly lower than in those used in Figures 5 and 7. Results show again that *ROS* has more effect on sequestration of (DH)AA inside the leaf cells than on accumulation in the apoplast (Figure 9 and Figure S3), and, similar as in Figure 5,
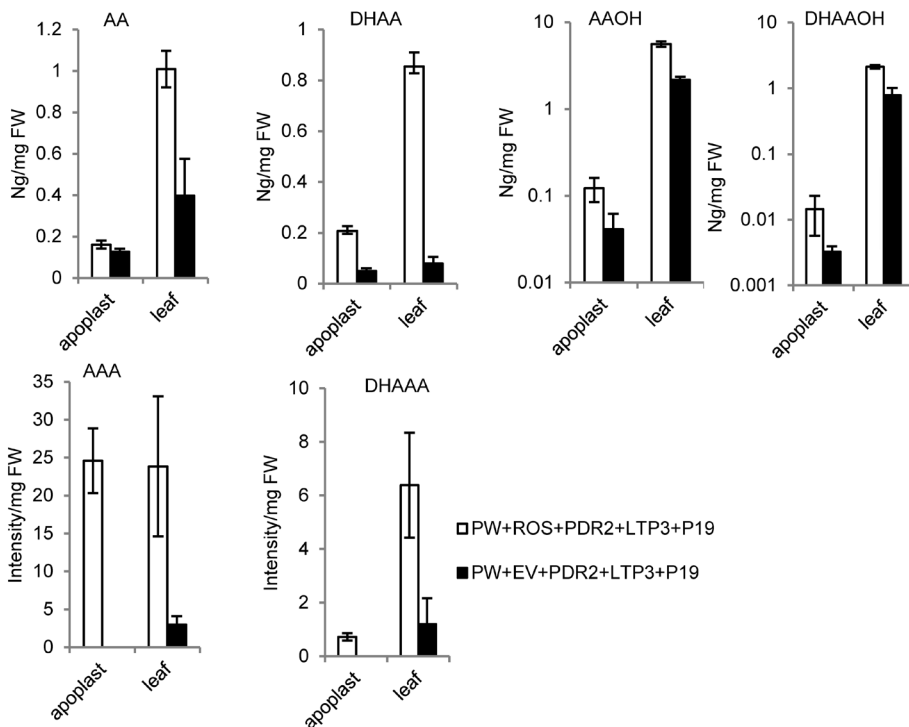
Figure 9. AN intermediate distribution in the apoplast and in leaf extracts of *N. benthamiana* infiltrated with artemisinin pathway (PW), *Rosea1* (*ROS*) and the *Artemisia annua* transporter genes (*PDR2*, *LTP3*) and analyzed by the LC-triple-quad-MS at 5 dpi. AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; AB, arteannuin B; DHAAOH, dihydroartemisinic alcohol; DHAAA, dihydroartemisinic aldehyde; DHAA, dihydroartemisinic acid. Error bar is SE, n=4.
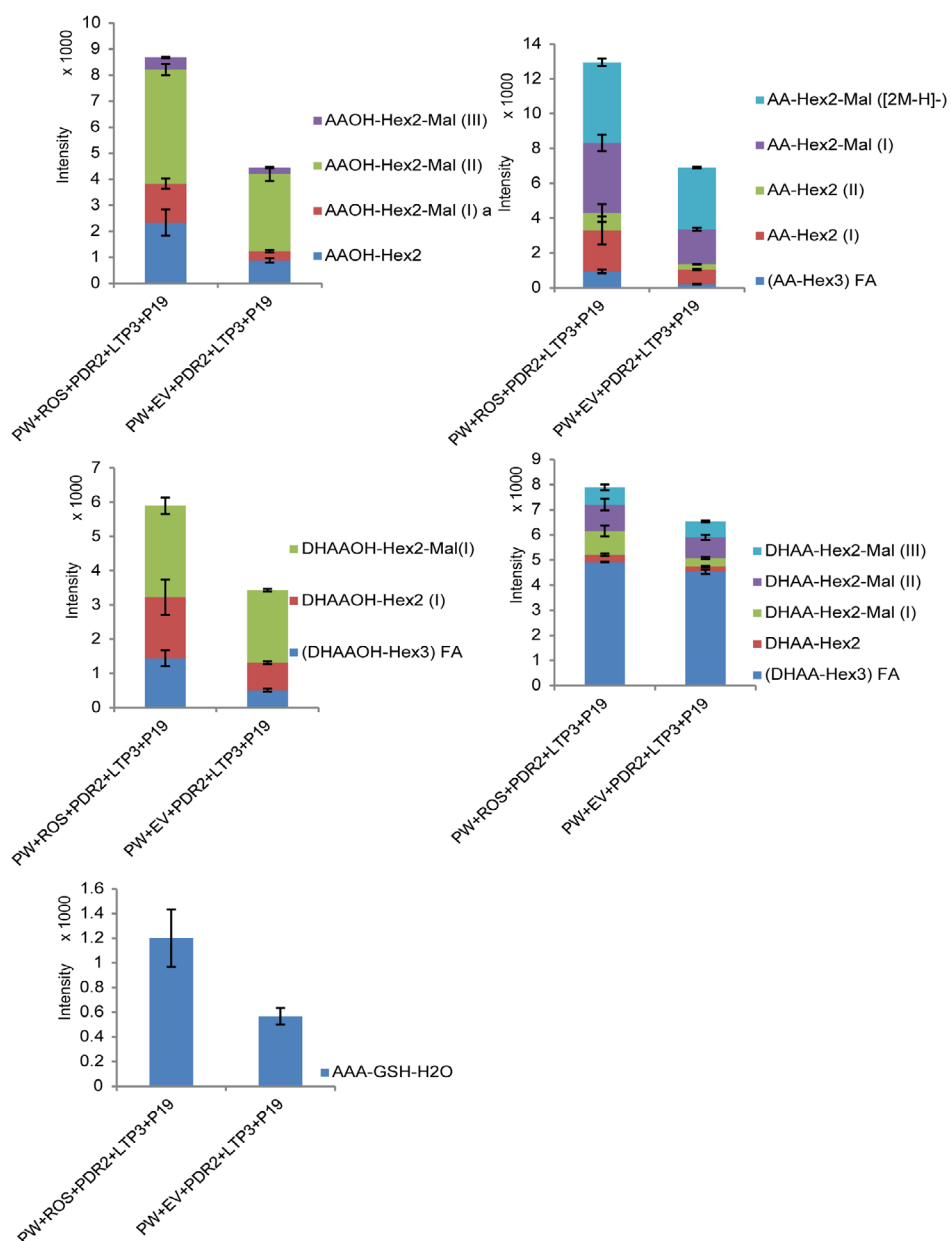
Figure 10. artemisinin pathway related glycosides and glutathione conjugates in leaves of *N. benthamiana* infiltrated with artemisinin pathway (PW), *Rosea1* (*ROS*) and transporter genes (*PDR2*, *LTP3*) analyzed by LC-triple-quad-MS at 5 dpi. AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; AB, arteannuin B; DHAAOH, dihydroartemisinic alcohol; DHAAA, dihydroartemisinic aldehyde; DHAA, dihydroartemisinic acid. Error bar is SE, n=4.

ROS stimulated accumulation of (DH)AAA in the apoplast. *ROS* also had a stimulating effect on glycosylated AN-PW product accumulation, as indicated by the increased accumulation of (DH)AA and (DH)AAOH glycosides (Figure 10). The effect of *ROS* on overall AN-PW product accumulation was not further influenced by coexpression of *AaPDR2* and *AaLTP3*. However, we did note that the ratio of different AN-PW related glycosides is different for *N. benthamiana* leaves expressing AN-PW+*ROS* (Figure 3) and AN-PW+*ROS*+*AaPDR2*+*AaLTP3* (Figure 10), indicating that the co-expression with *AaPDR2*+*AaLTP3* changes the AN-PW related leaf chemotype.
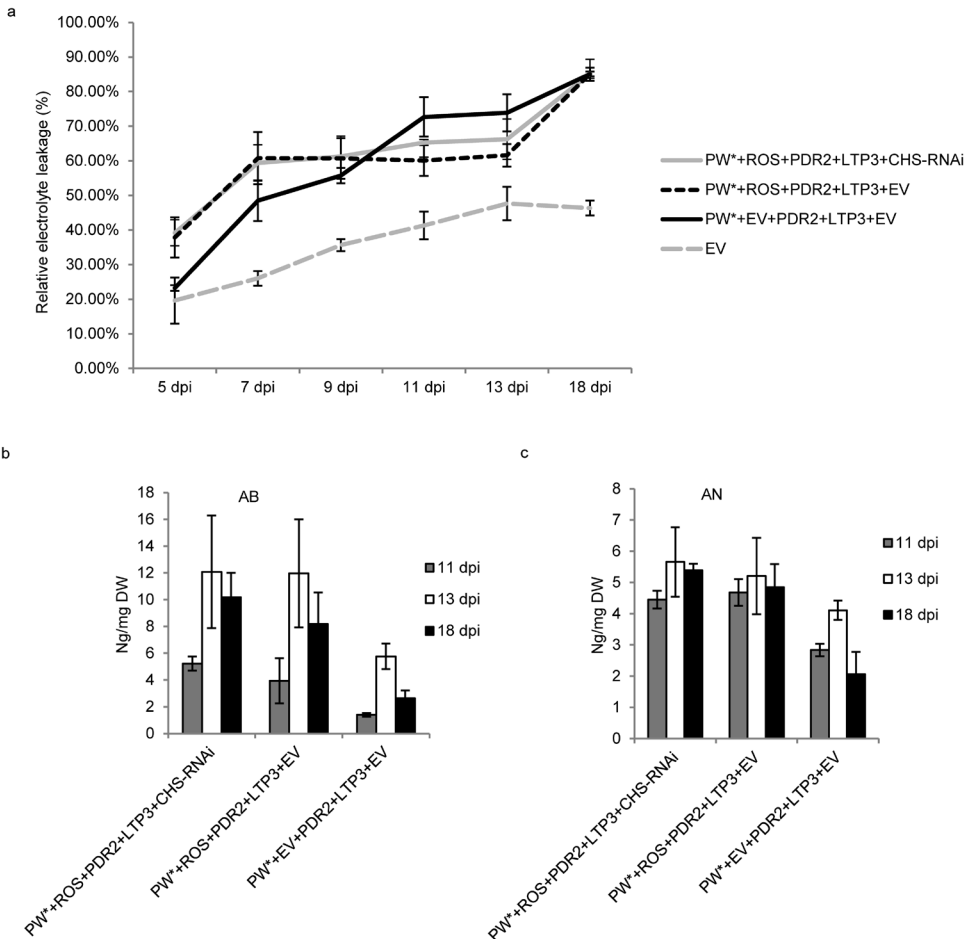
Figure 11. Final product accumulation and quantification of necrosis at different time points after infiltration with artemisinin pathway (PW*), *Rosea1* (*ROS*) and Artemisia transporter genes (*PDR2*, *LTP3*). a, Relative electrolyte leakage assay from 5 dpi to 18 dpi, b, arteannuin B(AB) detection from 11 dpi to 18 dpi; c, artemisinin (AN) detection from 11 dpi to 18 dpi. Error bar is SE, n=4.

ROS enhances AN and AB accumulation from AN-PW+*AaPDR2* + *AaLTP3* expression in *N. benthamiana* leaves

Previously we have shown that accumulation of the final products AN-PW products AN and AB only occurs after onset of necrosis in *N. benthamiana* leaves expressing the AN-PW genes (Chapter 3). Since leaves expressing AN-PW+*ROS* show visible signs of necrosis at day 6 we semi-quantified necrosis levels by ion-leakage assay (Bouwmeester et al., 2014). Results show that agro-infiltration of EV or AN-PW+*AaPDR2*+*AaLTP3* both resulted in ~20% ion leakage at 5 dpi, with a gradual increase to 40% at 18 dpi for EV, but to 80% for leaves expressing the AN-PW+*AaPDR2*+*AaLTP3* genes. When AN-PW+*AaPDR2*+*AaLTP3* were coexpressed with *ROS* the leakage at 5 dpi was already 40%, but was more or less steady at 60% from 7-13 dpi (Figure 11a). Subsequently we quantified AN and AB accumulation in *N. benthamiana* leaves agro-infiltrated with the different gene sets. The highest levels of AN and AB were reached at 13 dpi and product accumulation was highest for the expression of AN-PW+*ROS*+*AaPDR2*+*AaLTP3*+*CHS^RNAi* and AN-PW+*ROS*+*AaPDR2*+*AaLTP3*. AN or AB accumulation did not correlate with the level of necrosis, which was highest for leaves expressing *PW*+*AaPDR2*+*AaLTP3*+*EV* at 13 dpi (Figure 11 and Figure S4).

Discussion

ROS activity enhanced the AN metabolic flux

Here we have tested the interaction between the induced flavonoid biosynthesis pathway of *N. benthamiana* and an ectopically expressed AN sesquiterpene biosynthesis pathway. *ROS* activity resulted in a substantial and similar increase in both free form and glycosylated AN-PW products (Figure 2 and Figure 3). The effect of *ROS* on AN-PW+*AaPDR2*+*AaLTP3* (Figure 9) was similar to that of *ROS* on AN-PW alone (Figure 5 and Figure 11). Both results suggest that overall AN-PW activity is increased by *ROS*, while *ROS* does not affect transgene transcript levels (Figure 6). However, analysis of the expression of (precursor) pathway genes demonstrated a 2-fold upregulation of *NbMVK* (Figure 6) by *ROS*, which could result in an increased precursor flow into the AN-PW and which thus could account for the higher content of AN-PW products in the presence of *ROS* (Figure 2, 3 and 11). Previously, *HMGR* was reported as the rate-limiting step in the biosynthesis of isoprenoid compounds derived from the cytosolic MVA pathway (Chappell et al., 1995; Stermer and Bostock, 1987; Tholl and Lee, 2011). Indeed, co-infiltration of a

tHMGR with *ADS* resulted in strongly enhanced amorphadiene production in *N. benthamiana* (Van Herpen et al., 2010). However, *ROS* did not enhance transcriptional activity of endogenous *NbHMGR* (Figure 9). Since the AN-PW genes include *tHMGR* we propose that under transient expression conditions *HMGR* is no longer limiting for MVA pathway activity and that other enzymatic steps such as *MVK* become limiting. This could then explain why upregulation of endogenous *MVK* can further enhance the flux through the AN-PW (Hinson et al., 1997). Next to the effect of *ROS* on *MVK* expression, it is still possible that as result of *ROS* activity and higher flavonoid levels in the *N. benthamiana* leaves, proteins are less prone to oxidative damage. Indeed *ROS* expression does seem to lower and stabilize cell leakage between 9 dpi and 13 dpi (Figure 11a) and it could be that the increased flavonoid levels in these leaves also protect enzymes of the AN-PW, resulting in more prolonged activity, which than could contribute to increase AN-PW product levels.

Competition between the flavonoid and terpene pathway

In the test to determine if flavonoid compounds themselves affect AN-PW product accumulation, we found that inhibition of flavonoid biosynthesis, resulted in a 20% increase in AN-PW product levels (Figure 7 and Figure 8). However, final product accumulation of AB and AN at 13 dpi was not significantly affected by *CHS^{RNAi}* in the present of ROS (Figure 11). The effect of *ROS+CHS^{RNAi}* on AN-PW product intermediates at 5 dpi may be explained in different ways. For instance, phosphoenol pyruvate feeds both into the AN-PW and the flavonoid pathway (see Figure 12) and both pathways compete with each other for this compound. Thus the lower competition for phosphoenol pyruvate upon coexpression of *CHS^{RNAi}* could result in the increase in AN-PW product formation. Alternatively, it could be that upon *CHS^{RNAi}* expression, modifying P450 enzymes of the flavonoid biosynthesis pathway no longer consume NADPH nor utilize cytochrome P450 oxidoreductase (CPR) activity for P450 regeneration (Renault et al., 2014). This would leave more NADPH and CPR activity for the P450s of the AN-PW. We note that the apparent competition between the AN-PW and flavonoid biosynthesis PW could also be the result of exceptional high expression of the AN-PW genes in the context of the transient expression in *N. benthamiana* and therefore may be of less relevance to situations with more natural levels of PW gene expression.
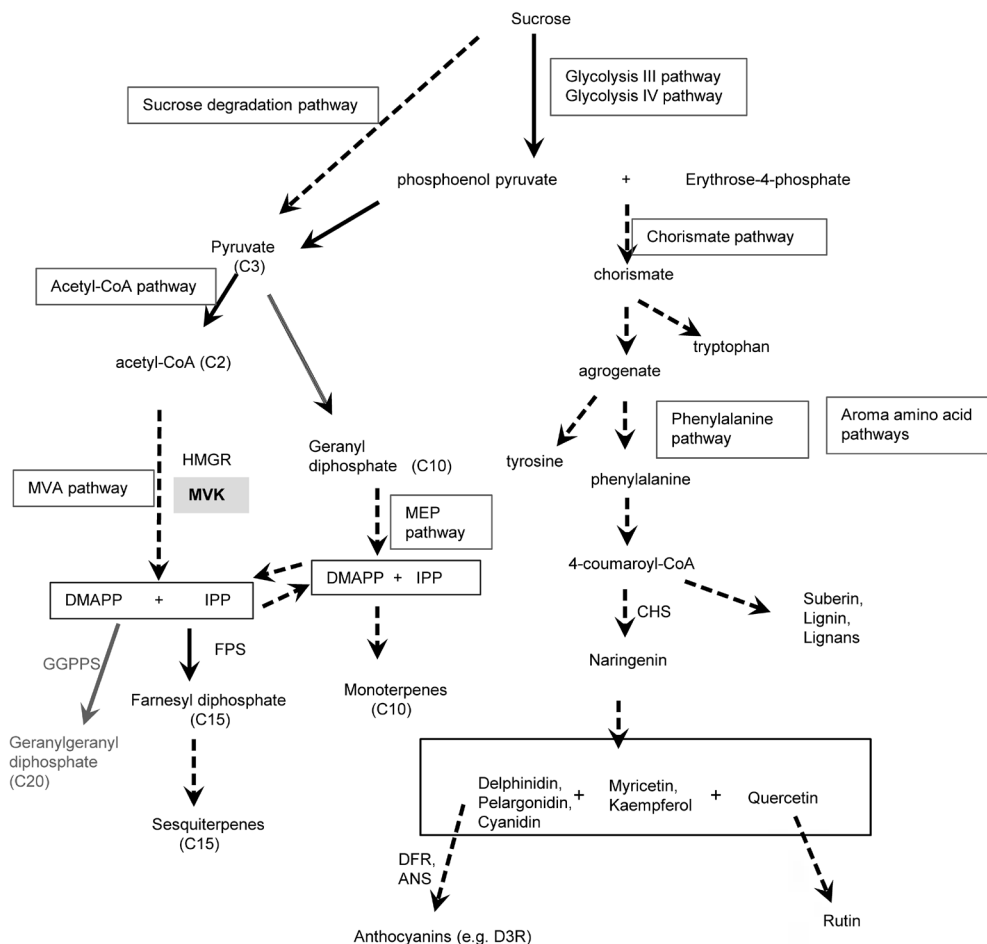
Figure 12. Schemtaic overview of the flavonoids and terpenoid biosynthesis pathways. Gene name abbreviations: HMGR, 3-hydroxy-3-methylglutaryl-CoA reductase; MVK, mevalonate kinase; GGPPS, geranylgeranyl diphosphate synthase; FPS, farnesyl diphosphate synthase; CHS, chalcone synthase; DFR, dihydroflavonol 4-reductase; ANS, anthocyanidin synthase.

AN-PW chemotype in *N. benthamiana* leaves affected by *ROS* and CHS[RNAi]

We previously defined the chemotype of the *N. benthamiana* leaves express-ing the AN-PW as the ratio between DHAA-hexose and AA-hexose, as quan-tified by peak intensities in MS chromatograms of leaf extract analysis (Ting et al., 2013). Table 1 lists this *N.benthamiana* leaf chemotype for the different treatments. Seasonal changes in growth conditions may affect flux through the PW and therefore in table 1 experiments executed at different times of

the year are indicated. The results show that ROS lowers the DHAA:AA leaf chemotype in two separate set of experiments (Figure 3 and Figure 10).

Table 1 ratio of DHAA-hex : AA-hex for the different gene combinations. Treatments performed at different times of the year are indicated by same number. AN-PW= V5+AMO+DBR2+ALDH1.

Treatment	DHAA-hex : AA-hex	Relative PW gene dosage	Harvest day
1AN-PW +EV+ p19 (this chapter)	3.21 (Figure 3)	4/6	(dpi)
1AN-PW+ROS +p19	1.74 (Figure 3)	4/6	5
2AN-PW+ROS+CHSRNAi	1.07 (Figure 8)	4/6	5
2AN-PW+ROS+EV	1.74 (Figure 8)	4/6	5
3AN-PW+EV+ AaPDR2+AaLT-P3+p19	0.95 (Figure 10)	4/8	5
3AN-PW+ROS +AaPDR2+AaLT-P3+p19	0.61 (Figure 10)	4/8	5

The ratio of DHAA:AA decreased from 3.21 to 1.74 by the *ROS* actvity, and even further by additional coexpressing with *CHS^{RNAi}* (to 1,07). Coexpression of the AN-PW with the *Artemisia* transporter genes combination *AaPDR2+AaLTP3* resulted in a DHAA:AA ration of 0,95, while coexpression of the AN-PW with transporters and *ROS* lowered the ratio to 0,61 (Table 1). In most of these treatments the relative gene dosage of the AN-PW was similar and also the inhibitor of transgene silencing (P19) does not seem to affect the leaf chemotype (e.g. compare AN-PW+*ROS*+P19 with AN-PW+*ROS*+EV). The effect of the transporter genes indicates that *AaPDR2+AaLTP3* have more effect on AA and therefore they are more efficient in stimulating the flux through the AA branch than through the DHAA branch of the AN-PW. Besides an effect of the treatments on the DHAA:AA chemotype, we also note difference in the ratio of the different types of AA-hex glycosides in *N. benthamiana* leaves expressing AN-PW+*EV* or AN-PW+*ROS* (Figure 3). At present it is not clear how ROS is affecting DHAA:AA or glycosylation chemotype. It could be that *ROS* affects expression of glycosyltransferases, which preferably act on AA. Alternatively, it could be due to a direct interaction between flavonoid and terpene molecules that somehow affects AA more than DHAA. However, this is not supported by the experiment with inhibition of flavonoid biosynthesis in the presence of *ROS*, which leads to even stronger effect on AA compared to DHAA. Although P19 did not affect the overall DHAA:AA chemotype, the ratio of AA-Hex$_2$-Mal to AA-

Hex$_2$ is higher in leaves expressing AN-PW+*ROS*+*CHS^{RNAi}* compared to that in leaves expressing AN-PW+*ROS*+*P19*. This indicates that either silencing or inhibition of flavonoid biosynthesis activity affects glycosylation. More careful investigation in the future is needed to determine whether and how the ratio of AA glycosides is affected by AN-PW agroinfiltration gene dosage, harvest dpi, *ROS* and or inhibition of transgene silencing by P19.

Concluding remarks

In conclusion, we here report that the TF *ROS* not only stimulates flavonoid biosynthesis genes but also stimulates gene expression for precursor input into the AN sesquiterpene pathway, presumably by upregulating *MVK*. In addition we have shown that the stimulated flavonoid biosynthesis activity competes with total sesquiterpene production at 5 dpi, but does not affect overall accumulation of AN/AB at 13 dpi in *planta*. The effect of activated flavonoid biosynthesis pathway activity on the type of AN-PW glycosylated products may be the consequence of interaction between flavonoid and sesquiterpenoid and needs further investigation

Acknowledgement

References

Baraldi, R., Isacchi, B., Predieri, S., Marconi, G., Vincieri, F.F., and Bilia, A.R. (2008). Distribution of artemisinin and bioactive flavonoids from *Artemisia annua* L. during plant growth. Biochemical Systematics and Ecology *36*, 340-348.

Bouwmeester, H.J., Wallaart, T.E., Janssen, M.H.A., Van Loo, B., Jansen, B.J.M., Posthumus, M.A., Schmidt, C.O., De Kraker, J.-W., König, W.A., and Franssen, M.C.R. (1999). Amorpha-4,11-diene synthase catalyses the first probable step in artemisinin biosynthesis. Phytochemistry *52*, 843-854.

Bouwmeester, K., Han, M., Blanco-Portales, R., Song, W., Weide, R., Guo, L.-Y., van der Vossen, E.A.G., and Govers, F. (2014). The *Arabidopsis* lectin receptor kinase LecRK-I.9 enhances resistance to *Phytophthora infestans* in *Solanaceous* plants. Plant Biotechnology Journal *12*, 10-16.

Chappell, J., Wolf, F., Proulx, J., Cuellar, R., and Saunders, C. (1995). Is the reaction

catalyzed by 3-hydroxy-3-methylglutaryl coenzyme A reductase a rate-limiting step for isoprenoid biosynthesis in plants? Plant Physiology *109*, 1337-1343.

Dudareva, N., Klempien, A., Muhlemann, J.K., and Kaplan, I. (2013). Biosynthesis, function and metabolic engineering of plant volatile organic compounds. New Phytologist *198*, 16-32.

Falcone Ferreyra, M.L., Rius, S.P., and Casati, P. (2012). Flavonoids: biosynthesis, biological functions, and biotechnological applications. Frontiers in Plant Science *3*, 222.

Ferrer, J.L., Austin, M.B., Stewart Jr, C., and Noel, J.P. (2008). Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. Plant Physiology and Biochemistry *46*, 356-370.

Glas, J.J., Schimmel, B.C.J., Alba, J.M., Escobar-Bravo, R., Schuurink, R.C., and Kant, M.R. (2012). Plant glandular trichomes as targets for breeding or engineering of resistance to herbivores. International Journal of Molecular Sciences *13*, 17077-17103.

Hinson, D.D., Chambliss, K.L., Toth, M.J., Tanaka, R.D., and Gibson, K.M. (1997). Post-translational regulation of mevalonate kinase by intermediates of the cholesterol and nonsterol isoprene biosynthetic pathways. Journal of Lipid Research *38*, 2216-2223.

Hong, G.-J., Hu, W.-L., Li, J.-X., Chen, X.-Y., and Wang, L.-J. (2009). Increased accumulation of artemisinin and anthocyanins in *Artemisia annua* expressing the *Arabidopsis* blue light receptor *CRY1*. Plant Mol Biol Rep *27*, 334-341.

Joosten, M.A.J. (2012). Isolation of apoplastic fluid from leaf tissue by the vacuum infiltration-centrifugation technique. In Plant Fungal Pathogens, M.D. Bolton, and B.P.H.J. Thomma, eds. (Humana Press), pp. 603-610.

Kang, J.-H., McRoberts, J., Shi, F., Moreno, J.E., Jones, A.D., and Howe, G.A. (2014). The flavonoid biosynthetic enzyme chalcone isomerase modulates terpenoid production in glandular trichomes of tomato. Plant Physiology *164*, 1161-1174.

Kang, J.-H., Shi, F., Jones, A.D., Marks, M.D., and Howe, G.A. (2010). Distortion of trichome morphology by the hairless mutation of tomato affects leaf surface chemistry. Journal of Experimental Botany *61*, 1053-1064.

Maeda, H., and Dudareva, N. (2012). The shikimate pathway and aromatic amino acid biosynthesis in plants. Annual Review of Plant Biology *63*, 73-105.

Markham, K.R., Gould, K.S., and Ryan, K.G. (2001). Cytoplasmic accumulation of flavonoids in flower petals and its relevance to yellow flower colouration. Phytochemistry *58*, 403-413.

Marrs, K.A., Alfenito, M.R., Lloyd, A.M., and Walbot, V. (1995). A glutathione S-transferase involved in vacuolar transfer encoded by the maize gene *Bronze-2*. Nature

375, 397-400.

Ökmen, B., Etalo, D.W., Joosten, M.H.A.J., Bouwmeester, H.J., de Vos, R.C.H., Collemare, J., and de Wit, P.J.G.M. (2013). Detoxification of α-tomatine by *Cladosporium fulvum* is required for full virulence on tomato. New Phytologist *198*, 1203-1214.

Outchkourov, N.S., Carollo, C.A., Gomez-Roldan, V., de Vos, R.C.H., Bosch, D., Hall, R.D., and Beekwilder, J. (2014). Control of anthocyanin and non-flavonoid compounds by anthocyanin-regulating MYB and bHLH transcription factors in *Nicotiana benthamiana* leaves. Frontiers in Plant Science *5*, 519.

Patra, B., Schluttenhofer, C., Wu, Y., Pattanaik, S., and Yuan, L. (2013). Transcriptional regulation of secondary metabolite biosynthesis in plants. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms *1829*, 1236-1247.

Petrussa, E., Braidot, E., Zancani, M., Peresson, C., Bertolini, A., Patui, S., and Vianello, A. (2013). Plant flavonoids—biosynthesis, transport and involvement in stress responses. International Journal of Molecular Sciences *14*, 14950.

Renault, H., Bassard, J.-E., Hamberger, B., and Werck-Reichhart, D. (2014). Cytochrome P450-mediated metabolic engineering: current progress and future challenges. Current Opinion in Plant Biology *19*, 27-34.

Rydén, A.-M., Ruyter-Spira, C., Quax, W.J., Osada, H., Muranaka, T., Kayser, O., and Bouwmeester, H. (2010). The molecular cloning of dihydroartemisinic aldehyde reductase and its implication in artemisinin biosynthesis in *Artemisia annua*. Planta Med *76*, 1778-1783.

Saija, A., Tomaino, A., Trombetta, D., Giacchi, M., De Pasquale, A., and Bonina, F. (1998). Influence of different penetration enhancers on *in vitro* skin permeation and *in vivo* photoprotective effect of flavonoids. International Journal of Pharmaceutics *175*, 85-94.

Saito, K., Yonekura-Sakakibara, K., Nakabayashi, R., Higashi, Y., Yamazaki, M., Tohge, T., and Fernie, A.R. (2013). The flavonoid biosynthetic pathway in *Arabidopsis*: structural and genetic diversity. Plant Physiology and Biochemistry *72*, 21-34.

Schijlen, E.G.W.M., de Vos, C.H.R., Martens, S., Jonker, H.H., Rosin, F.M., Molthoff, J.W., Tikunov, Y.M., Angenent, G.C., van Tunen, A.J., and Bovy, A.G. (2007). RNA interference silencing of chalcone synthase, the first step in the flavonoid biosynthesis pathway, leads to parthenocarpic tomato fruits. Plant Physiology *144*, 1520-1530.

Schwinn, K.E., Boase, M.R., Bradley, M., Lewis, D.H., Deroles, S.C., and Davies, K.M. (2014). MYB and bHLH transcription factor transgenes increase anthocyanin pigmentation in petunia and lisianthus plants, and the petunia phenotypes are strongly enhanced under field conditions. Frontiers in Plant Science *5*.

Stermer, B.A., and Bostock, R.M. (1987). Involvement of 3-hydroxy-3-methylglutaryl coenzyme A reductase in the regulation of resquiterpenoid phytoalexin synthesis in potato. Plant Physiology *84*, 404-408.

Stommel, J.R., Lightbourn, G.J., Winkel, B.S., and Griesbach, R.J. (2009). Transcription factor families regulate the anthocyanin biosynthetic pathway in *Capsicum annuum*. Journal of the American Society for Horticultural Science *134*, 244-251.

Suttipanta, N., Pattanaik, S., Kulshrestha, M., Patra, B., Singh, S.K., and Yuan, L. (2011). The Transcription Factor CrWRKY1 Positively Regulates the Terpenoid Indole Alkaloid Biosynthesis in Catharanthus roseus. Plant Physiology *157*, 2081-2093.

Tai, H.H., Goyer, C., and Murphy, A.M. (2013). Potato MYB and bHLH transcription factors associated with anthocyanin intensity and common scab resistance. Botany *91*, 722-730.

Teoh, K.H., Polichuk, D.R., Reed, D.W., and Covello, P.S. (2009). Molecular cloning of an aldehyde dehydrogenase implicated in artemisinin biosynthesis in *Artemisia annua*. Botany *87*, 635-642.

Teoh, K.H., Polichuk, D.R., Reed, D.W., Nowak, G., and Covello, P.S. (2006). *Artemisia annua* L. (*Asteraceae*) trichome-specific cDNAs reveal CYP71AV1, a cytochrome P450 with a key role in the biosynthesis of the antimalarial sesquiterpene lactone artemisinin. FEBS Letters *580*, 1411-1416.

Tholl, D., and Lee, S. (2011). Terpene Specialized Metabolism in *Arabidopsis thaliana*. The Arabidopsis Book / American Society of Plant Biologists *9*, e0143.

Ting, H.-M., Wang, B., Rydén, A.-M., Woittiez, L., van Herpen, T., Verstappen, F.W.A., Ruyter-Spira, C., Beekwilder, J., Bouwmeester, H.J., and van der Krol, A. (2013). The metabolite chemotype of *Nicotiana benthamiana* transiently expressing artemisinin biosynthetic pathway genes is a function of CYP71AV1 type and relative gene dosage. New Phytologist *199*, 352-366.

Van Herpen, T.W., Cankar, K., Nogueira, M., Bosch, D., Bouwmeester, H.J., and Beekwilder, J. (2010). *Nicotiana benthamiana* as a production platform for artemisinin precursors. PLoS ONE *5*, e14222.

Voinnet, O., Rivas, S., Mestre, P., and Baulcombe, D. (2003). An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. The Plant Journal *33*, 949-956.

Vranová, E., Coman, D., and Gruissem, W. (2013). Network analysis of the MVA and MEP pathways for isoprenoid synthesis. Annual Review of Plant Biology *64*, 665-700.

Wang, W., Wang, Y., Zhang, Q., Qi, Y., and Guo, D. (2009). Global characterization of *Artemisia annua* glandular trichome transcriptome using 454 pyrosequencing. BMC Genomics *10*, 465.

Witzel, K., Shahzad, M., Matros, A., Mock, H.-P., and Muhling, K. (2011). Comparative evaluation of extraction methods for apoplastic proteins from maize leaves. Plant Methods *7*, 48.

Zhang, Y., Teoh, K.H., Reed, D.W., Maes, L., Goossens, A., Olson, D.J.H., Ross, A.R.S., and Covello, P.S. (2008). The molecular cloning of artemisinic aldehyde Δ11(13) reductase and its role in glandular trichome-dependent biosynthesis of artemisinin in *Artemisia annua*. Journal of Biological Chemistry *283*, 21501-21508.

Zvi, M.M.B., Shklarman, E., Masci, T., Kalev, H., Debener, T., Shafir, S., Ovadis, M., and Vainstein, A. (2012). PAP1 transcription factor enhances production of phenylpropanoid and terpenoid scent compounds in rose flowers. New Phytologist *195*, 335-345.

Supporting information

Table S1 Primers

Primer name	Sequence from 5' to 3'
35S-prom-F	GCTCCTACAAATGCCATCA
35S-prom-R	GATAGTGGGATTGTGCGTCA
NbHMGR2-F	CCAGGGATAACAATGATGATTC
NbHMGR2-R	CAGCCAAGCGTAGTTGCG
Ubi3-F	GCCGACTACAACATCCAGAAGG
Ubi3-R	TGCAACACAGCGAGCTTAACC
NbMVK_3F	ATGGCGTGGAGATTTCATTC
NbMVK_3R	GGCACGGTTTTGATGTTTCT
NbGGPPS_1F	CTGGGCCTCTATCGAGTCTG
NbGGPPS_1R	CATGTGCTGGTGGGTGTAAA
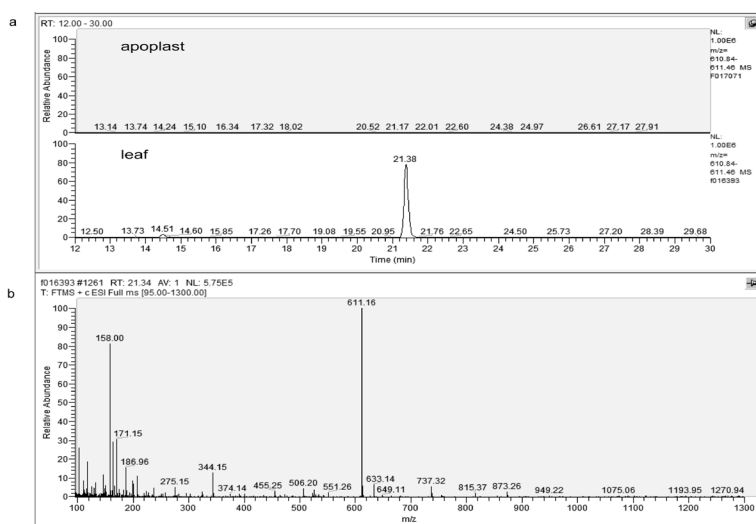
Figure S1 LC-orbitrap-MS chromatogram of [M+H]$^+$ = 611.160 in apoplastic fluid of *N. benthamiana* infiltrated with *AN-PW+ROS+P19*. a, apoplast sample and remaining leaf sample; b, Rutin MS spectrum of peak at 21.37 for *N. benthamiana* leaves infiltrated with *AN-PW+ROS*.
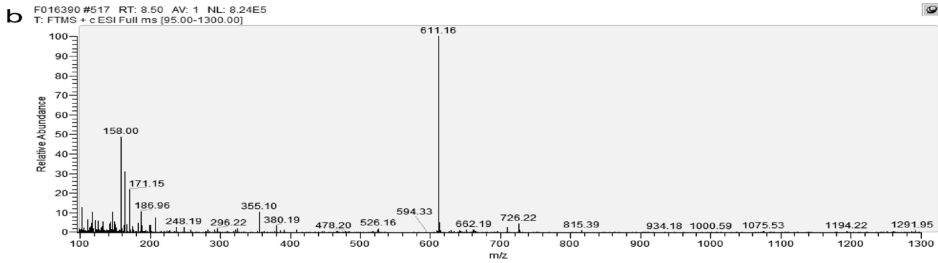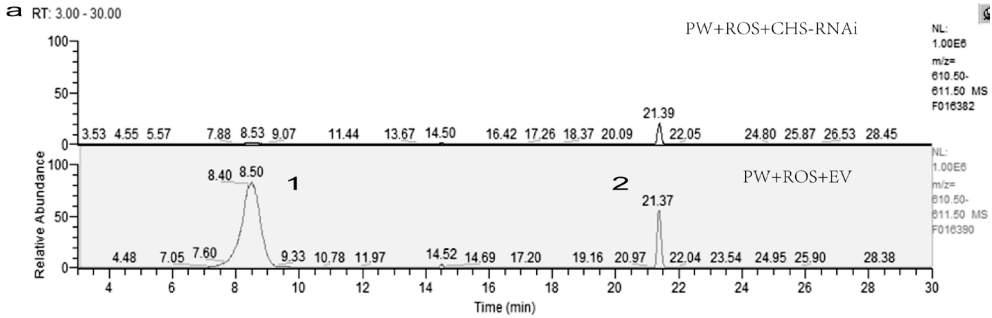

Figure S2 Characterization of flavonoid production in leaves of *N. benthamiana*. a: LC-orbitrap-MS chromatogram of [M+H]$^+$ = 611.160 at retention time=8.505min, b: MS spectrum of D3R of peak at 8.505 for *N. benthamiana* leaves infiltrated with AN-*PW+ROS+EV*.
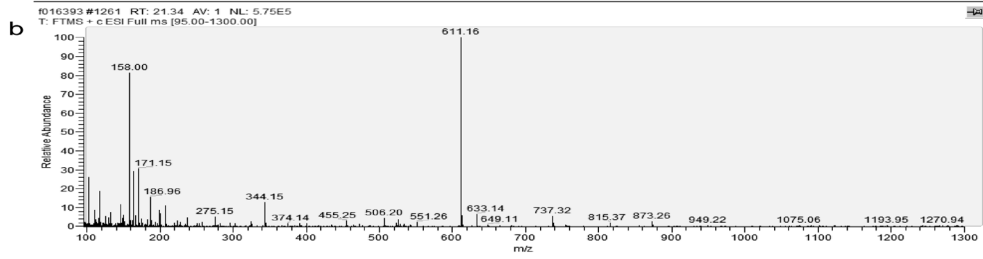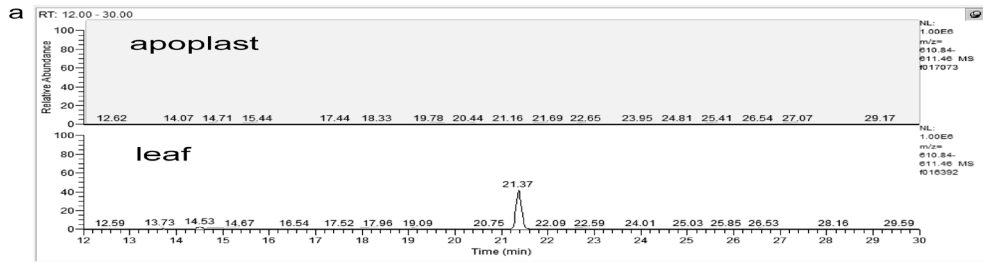
Figure S3 LC-orbitrap-MS chromatogram of AN-PW+*ROS*+*PDR2*+*LTP3*+*P19* at [M+H]$^+$ = 611.160 in apoplastic fluids of *N. benthamiana*. A, apoplast sample and remaining leaf sample; b, Rutin MS spectrum of peak at 21.37 for *N. benthamiana* leaves infiltrated with *AN-PW+ROS*.
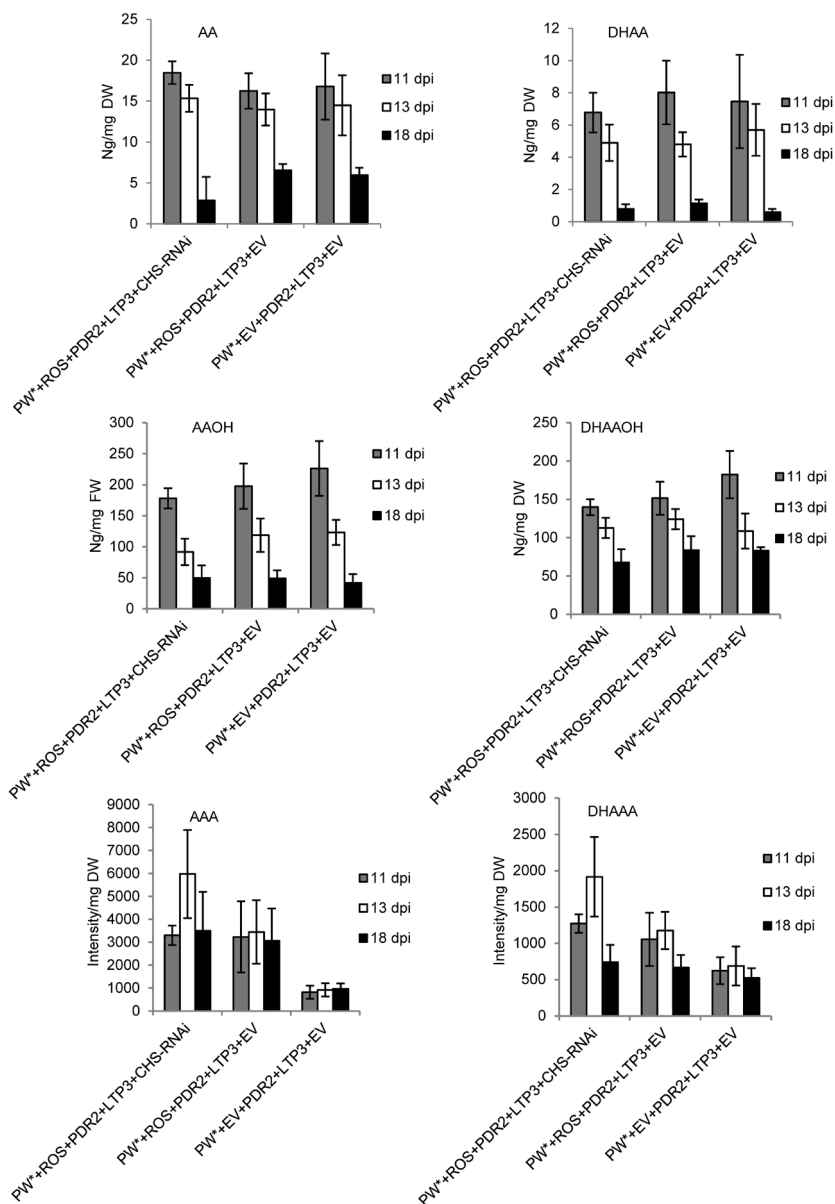

Figure S4 AN-intermediates accumulation from 11 dpi to 18 dpi in *N. ben-thaimama* infiltrated with artemisinin pathway (PW*), *Rosea1* (*ROS*) and Artemisia transporter genes (*PDR2*, *LTP3*); Error bar is SE, n=4.

Chapter 5

Higher terpene production from ectopic expressed terpene synthase genomic clones than from terpene synthase cDNAs in transgenic *Arabidopsis*

Bo Wang, Melida Rojas Chango, Thierry Delatte, Harro J. Bouwmeester, and Alexander R. van der Krol

Laboratory of Plant Physiology, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands.

Abstract

Intron splicing and splicing complexes remaining on processed mRNA may stimulate gene expression in eukaryotes in multiple ways. However, in metabolic engineering of plants, mostly ectopic expression of cDNAs under control of a strong promoter is used. Here we tested whether ectopic expression of a genomic clone consistently results in higher expression than the corresponding cDNA when under control of the same promoter. For this purpose the genomic and cDNA clone of the sesquiterpene synthases Caryophyllene Synthase (g*CST* and *cCST)* and Amorphadiene Synthase (*gADS* and *cADS*) were isolated from *Arabidopsis thaliana* and *Artemisia annua* respectively and expressed from the CaMV 35S promoter in *Arabidopsis*. For each expression construct 6-10 independent stable transformed plants were generated and overexpression (OE) in leaves was quantified. Average *CST* expression in *gCST*-OE lines is 27-fold higher than in *cCST*-OE lines, while average *ADS* expression in *gADS*-OE lines is 11-fold higher than in *cADS*-OE lines. The headspace volatiles of rosette plants were quantified by GC-MS, showing that average β-caryophyllene emission of the *gCST* overexpression (OE) lines is 58-fold higher than that of *cCST*-OE lines, while average amorphadiene emission of the *gADS*-OE lines is 5.4-fold higher than that of *cADS*-OE lines. Therefore, for both *CST* and *ADS* results indicate that genomic clones with introns boost transgene expression compared to the cDNA. The four expression constructs were also tested in transient expression assays using agro-infiltration of *N. benthamiana* leaves. However, in this context the stimulating effect of intron containing transgenes is largely lost.

Keywords

Introduction

Introns in eukaryotic pre-mRNAs are removed by splicing complex to generate mature mRNA. During intron splicing, exon junction complexes (EJC) bind to the pre-mRNA and next to a role in joining the 5'and 3' intron splice junctions sites, these complexes may have a role in maturated mRNA export and in decay of faulty nonsense-RNAs (Le Hir et al. 2001). Both in plants, insects, mice and humans some introns in genes are known to enhance expression and this effect has been named Intron-Mediated Enhancement (IME) (Mascarenhas et al. 1990). Introns and intron-splicing may stimulate the transcription in multiple ways (Le Hir et al. 2003; Niu and Yang 2011). For instance, (1) introns may contain transcriptional enhancer elements, which help stimulate expression from the upstream promoter. (2) the intron-splicing complex may interact with the RNA polymerase II and enhance reloading of the transcriptional machinery onto the promoter (Moabbi et al. 2012)). (3) the intron splicing complex may interact with the 7-methyl guanosine (m7G) 5'-end capping and 3'-end polyadenylation, enhancing the splicing efficiency and polyadenylation on either side of the terminal exon (Gunderson et al. 1998; Komarnitsky et al. 2000; Lewis and Izaurflde 1997; Proudfoot et al. 2002; Vagner et al. 2000). In some instances, (4) intron splicing affects exon RNA editing before splicing and may thus affect the actual protein product (Le Hir et al. 2003; Reenan 2001; Sowden and Smith 2001). (5) Residual protein complexes on the intron splice site on mature mRNA may promote export of mRNA from the nucleus (Legrain and Rosbash 1989; Luo and Reed 1999). (6) In addition, residual protein complexes on the intron splice site on mature mRNA may increase translational efficiency of mRNAs by affecting mRNA stability (Le Hir et al. 2003). (7) It has been demonstrated that exon splice-junctions near the 5'-end of some transcripts stimulate translation relative to an intronless mRNA, whereas splice-junctions in the 3'-UTR significantly reduce translation in Xenopus oocytes (Braddock et al. 1994; Matsumoto et al. 1998). Models have been developed to predict the enhancement of introns on gene expression, and the effect of an intron sequence and position on gene expression can thus be calculated and expressed as an IME score (Parra et al. 2011; Rose et al. 2008).

In our previous research on engineering of sesquiterpene biosynthesis in plants we produced *Arabidopsis* plants with overexpression of the *Arabidopsis* caryophyllene synthase gene (*CST*), using a genomic clone (*gCST*). While a previously reported *Arabidopsis* line expressing *CST* cDNA (*cCST*) emits ~ 4.0 ng·g^{-1} FW·h^{-1} (Huang et al. 2012), our line with overexpression of the *gCST* emits ~ 80 ng·g^{-1} FW·h^{-1} (Delatte, unpublished). In both cases

the transgene was driven by the same CaMV 35S promoter and both the *cCST* and *gCST* encode for identical proteins. Because in metabolic engineering of terpene production in heterologous plant hosts we aim for high production, we set out to determine whether there is a consistent higher expression from intron containing genes, as additional strategy to boost ectopic terpene production in plants. Here we show a proof of principle for the *Arabidopsis CST* and another sesquiterpene biosynthesis gene, encoding the first committed step in artemisinin biosynthesis in *Artemisia annua*, Amorphadiene synthase (*ADS*) (Bertea et al. 2006; Bouwmeester et al. 1999; Ting et al. 2013; Van Herpen et al. 2010). Indeed, in stable transformants the genomic fragment consistently results in higher expression and terpene production, but this boosting effect of an intron-containing gene is more or less lost in the context of transient expression in *N. benthamiana*.

Materials and methods

Quantitative reverse transcription PCR (qRT-PCR)

RNA was extracted from rosette leaves using RNeasy Plant Mini kit (Qiagen, Germany) with RNAase-Free DNase digestion (Qiagen, Germany). The quality and quantity of the RNA were assessed using the agarose gel electrophoresis and Nanodrop (Nanodrop Technologies, USA). cDNA was transcribed from 1000 ng of RNA using iScript cDNA Synthesis kit (BioRad, USA). Gene specific primer pairs for the qPCR analysis were *actin* (5'- CTA AGC TCT CAA GAT CAA AGG CTT A -3'/ 5'- ACT AAA ACG CAA AAC GAA AGC GGT T -3'), *ADS*-qF/ *ADS*-qR (5'- GGA GTA TGC CCA AAC CTT GA -3'/ 5'- TCG TCT CCC ATA CGT GTG AA -3') and *CST*-qF/ *CST*-qR (5'- TCA AGG ACG AAA ATG GGA AG -3'/ 5'- ATT GCG AGA TGA GGG CTA GA -3'). The primer pairs of house-keeping gene are *NbUbi3_* q/*NbUbi3*_qR (5'- GCC GAC TAC AAC ATC CAG AAG G-3'/ 5'-TGC AAC ACA GCG AGC TTAACC) and *Atactin2*_qF / *Atactin2*_qR (5'-CTAAGCTCTCAAGATCAAAGGCTTA-3'/ 5'-ACTAAAACG-CAAAACGAAAGCGGTT -3') were used for *Nicotiana benthamiana* and *Arabidopsis* samples, respectively. The qRT-PCR was performed on the iCycler iQ5 system (BioRad) as following program: 95 °C for 10 min, 40 cycles of 95 °C for 10 s, 55°C for 30 s, 95 °C for 1 min and followed by melting curve analysis from 65 to 95 °C. Relative gene expression was calculated using $\Delta\Delta$CT.

Vector construction

35S:gCST and *35S: cCST*. Genomic DNA was extracted from *Arabidopsis* (Col-0) leaves by the cetyltrimethylammonium bromide (CTAB) method. *gCST* and *cCST* were amplified separately using primer pair *CST*-F (5'- TTT TCC ATG GGG AGT GAA GTC AAC CGTC -3') and *CST*-R (5'- TTG GTT GCG GCC GCT CAA ATG GGT ATA G -3'). The NcoI and NotI restriction sites were introduced and cloned into the Impact Vector pIV1A2.1, which contains the CaMV35S promoter and Rbcs terminator (www.impactvector.com). The resulting pIV1A2.1/gCST and pIV1A2.1/cCST were cloned into pBinPlus vector using the LR reaction. The pBinPlus containing *35S: gCST* and *35S: cCST* were transformed to *A. tumefaciens* AGL-0 using the electroporation, separately.

35S:gADS and *35S: cADS*. Genomic DNA was extracted from A. *annua* flowers by the CTAB method. *gADS* and *cADS* were amplified separately using primer pair *ADS*-F (5'- CTA CTC TAG AAT GTC ACT TAC AGA AGA -3') and *ADS*-R (5'-GAC AAA CCA TGG TCA TAT ACT CAT AGG -3'). The XbaI and NcoI restriction sites were introduced and cloned into the Impact Vector pIV1A2.1, which contains the CaMV35S promoter and Rbcs terminator (www.impactvector.com). The resulting pIV1A2.1/gADS and pIV1A2.1/cADS were cloned into pBinPlus vector using the LR reaction. The pBinPlus containing *35S:gADS* and *35S:cADS* were transformed into *A. tumefaciens* AGL-0 using the electroporation.

Transformation of *Arabidopsis*

The expression constructs *35S:gCST*, *35S:cCST*, *35S:gADS* and *35S:cADS* were transformed into *Arabidopsis* Col-0 using the floral dipping method as described (Harrison et al. 2006). Transformed seeds were selected on MS medium containing kanamycin and 6-10 independent transformed lines for each expression construct were selected after 5 days on selective medium and transfer to soil for further analysis. Transformation was confirmed by qPCR on cDNA isolated from rosette leaves harvested from 3 weeks old primary transformants using primer pairs specific to each of the four expression constructs.

Plant growth conditions

Transformed seeds were sown on the MS medium with the kanamycin and stored in the dark at cold room (4°C). After 2 days, seeds were transferred in the room temperature until germination, and then sown in soil

in growth chamber (12 h light at 22 °C in the day and 12 h dark at 18 °C in the night). 3 week old plants were prepared for the further analysis.

Transient expression in *N. benthamiana*

The agro-infiltration was carried out in the 4th leaf of 5 week old *N. benthamiana* plants as described (Van Herpen et al., 2010). Briefly, *A. tumefaciens* strains transformed with individual expression constructs were infiltrated into *N. benthamiana* leaves using a 1 ml syringe without needle. Dosage compensation of infiltrated *A. tumefaciens* was done by diluting with an *A. tumefaciens* strain carrying an empty vector (EV) construct. Agrobacterium with a P19 expression construct was added to suppress post-transcriptional gene silencing (Voinnet et al. 2003). Infiltrated leaves were harvested at the 5 days post infiltration (dpi).

Headspace trapping and analysis of volatiles by GC-MS

Headspace analysis of stable transformed *Arabidopsis* was done using 3 weeks old primary transformants grown on soil. The volatiles were collected use dynamic headspace trapping as described (Houshyani et al. 2013). After *Arabidopsis* were enclosed in 1 L glass jar, the vacuum pump was used to draw air through the glass jar and the incoming air was purified through Tenax TA at 200 ml/min inlet. The outlet was trapped at 300 ml/min to collect the volatiles. The *gCST*-OE lines were trapped 4 hours from 10.00-13.00 at 22 °C. Wildtype plants and rest of the OE lines were trapped 24 hours (12 h light at 22 °C in the day and 12 h dark at 18 °C in the night). The headspace samples were analyzed by Thermodesorption GC-MS with a thermal desorber (Unity, Markes International Limited) and a Trace GC Ultra (Thermo Electron Corporation) coupled with DSQ mass spectrometer (Thermo Electron Corporation) as described (Houshyani et al. 2013). The GC-MS results were analyzed by the Xcalibur software (Thermo, Waltham). Quantification of terpene levels was accomplished by determining the peak area of the characteristic m/z (133) for β-caryophyllene, the characteristic m/z (93) for α-humulene, the characteristic m/z (119) for α-copene and the characteristic m/z (189) for amorphadiene. As the comparison, 5 open and 5 more closed *Arabidopsis* flowers were trapped for 24 h by twister in 10 mL vial in a glass vial and measured on GC-MS.

Quantification of untargeted metabolites by UPLC-MS analysis

After headspace analysis rosette leaves 3 and 4 were harvested from

3 week old primary transformants and immediately frozen in liquid Nitrogen. From this, 100 mg of frozen ground samples were extracted using 300 ul 99.9% MeOH with 0.133% formic acid. Analysis of non-volatile compounds in transgenic *Arabidopsis* rosette extracts was carried out by LC-QTOF-MS as described (Ting et al. 2013). The raw mass data was processed using metalign (www.metalign.nl). Multiple mass signals from the same compound were grouped using MSClust software (biotools.wurnet.nl). Significance of average intensity differences of aligned mass-signals between samples were assessed using student t-test ($p < 0.05$).

Results

Isolation of *gCST* and *gADS* genomic sequences

In order to investigate whether the introns affect the sesquiterpene production, the genomic DNA and cDNA of *Arabidopsis CST* (*gCST* and *cGST*) and *Artemisia annua ADS* (*gADS* and *cADS*) were isolated by PCR on (*Arabidopsis* or *A. annua*) genomic DNA or (*Arabidopsis* or *A. annua*) leaf cDNA, respectively. The fragments (*gCST*, *cCST*, *gADS*, *cADS*) were cloned downstream of the CaMV 35S promoter in a binary expression vector and *Arabidopsis* plants were transformed with the expression constructs *35S:gCST*, *35S:cCST*, *35S:gADS* and *35S:cADS* by the conventional floral dipping method (Harrison et al. 2006). Sequence analysis of the genomic clones showed that both the *Arabidopsis* g*CST* and *A. annua* g*ADS* gene have six introns (Fig. 1A and B). The *Arabidopsis* TAIR database indicates two splice variants for CST mRNA, with a small amino acid sequence difference in exon 5 (Fig. 1A). However, the alternative *CST* splice product was not detected in our leaf cDNA, also not using primer pairs specific for the alternative splice product (Fig. 1A). The intron sequences of *gCST* and *gADS* were evaluated using IMEter v2.0 (http://korflab.ucdavis.edu/cgi-bin/web-imeter2.pl) (Parra et al. 2011). The analysis indicates that intron 4 of *gCST* and intron 2 of *gADS* have the highest IME score of the six introns in each gene (Table 1). An IME score of 20 predicts a 5-fold increase in expression (Parra et al. 2011; Rose et al. 2008). Since the highest IME score is lower, but similar for both *gCST* and *gADS*, the prediction is that both may show a similar, but lower than 5-fold stimulation of gene expression, when compared to expression of their respective cDNAs.

Table 1. IMEter score of individual introns for *CST* and *ADS* genes

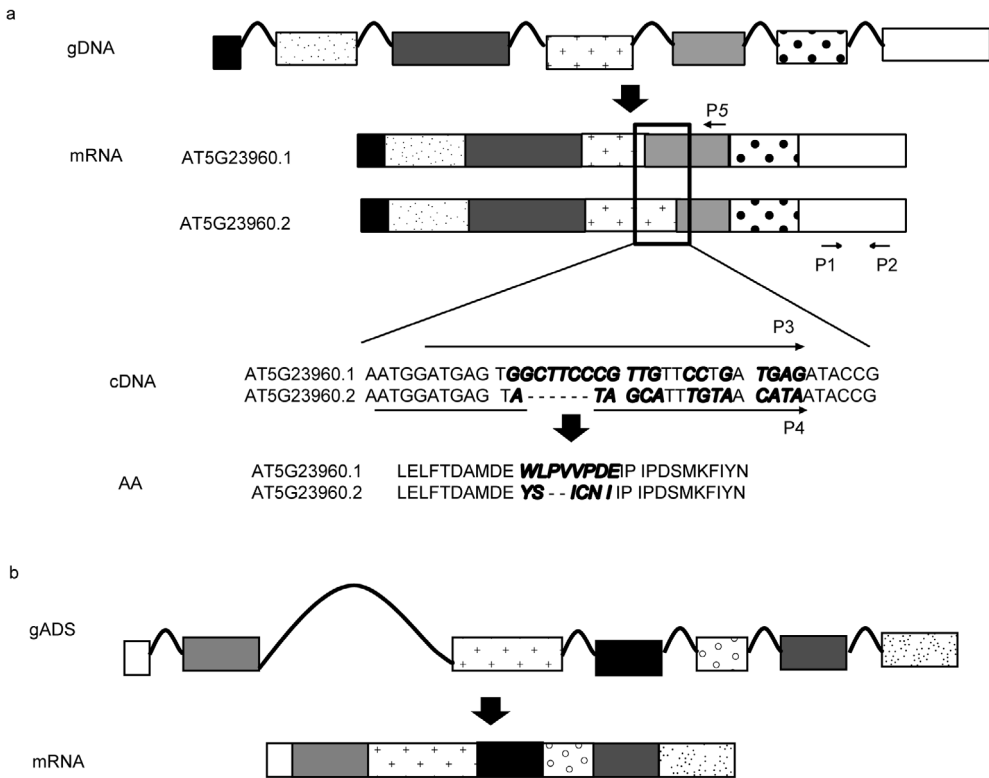| Intron	AT5G23960.1	AT5G23960.2	ADS
1	1.70	1.70	0.97
2	3.02	3.02	13.12
3	0.00	0.00	1.79
4	10.27	11.93	0.00
5	0.00	0.00	0.00
6	0.00	0.00	0.00

Fig.1. Sequence comparison of gDNA and mRNA of caryophyllene synthase (*CST)* and amorphadiene synthase *(ADS)* genes. a, *CST*; b, *ADS*.

Introns increased the transcription expression in *Arabidopsis*

For each of the four expression (*35S:gCST, 35S:cCST, 35S:gADS and*

35S:cADS) constructs 6-10 independent *Arabidopsis* transformants were selected, based on seedling resistance to kanamycin. After 5 days on selective medium, transformed seedlings were transferred to soil and plants were grown for 3 weeks at which RNA was isolated from two rosette leaves for quantification of transgene expression level by qPCR. On average the transcription level of *CST* was increased 74-fold for lines expressing *cCST*, compared to expression of the endogenous *CST* in untransformed *Arabidopsis*. In contrast, average transcription level of *CST* was 1978-fold higher for lines expressing *gCST*, compared to endogenous *CST* activity (Fig.
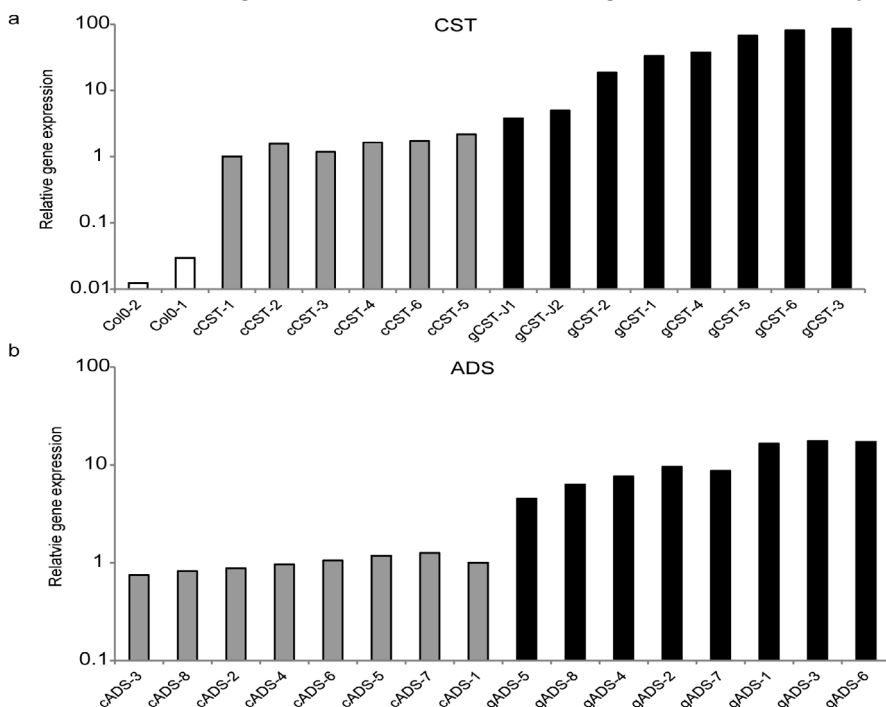
Fig.2. Transcription levels of caryophyllene synthase (*CST*) and amorphadiene synthase (*ADS*) genes in stable transformed *Arabidopsis*. a, gene expression of *CST* in transformed *Arabidopsis*, all values normalized to *cCST*-1; b, gene expression of *ADS* in transformed *Arabidopsis*, all values normalized to *cADS*-1.

2A). Therefore, on average, expression in *gCST* transformants is 26.7-fold higher than in *cCST* transformants. Similarly, *ADS* transcript levels in individual *gADS*-OE lines were consistently higher than the *ADS* transcript levels in *cADS*-OE lines (Fig. 2B), with average expression level of *ADS* in *gADS*-OE lines 11-fold higher than the average expression in cADS-OE lines. Therefore, for both *CST* and *ADS*, expression was higher from the intron-containing genomic DNA than from the corresponding cDNA. For

both the *CST* and *ADS* the increase in expression of the genomic clone compared to the cDNA clone was higher than predicted by the IME score.

Introns of *CST* increased the sesquiterpene emission in *Arabidopsis*

To test how the increased expression in transgenic *gCST*-OE and *gADS*-OE lines (compared to *cCST*-OE and *gADS*-OE lines, respectively) is affecting sesquiterpene emission, volatiles were trapped in the headspace of 3-week old primary transformants and analyzed by GC-MS. β-caryophyllene emission in eight selected *gCST*-OE lines was consistently higher than in the six selected *cCST*-OE lines (Fig. 3A). The average β-caryophyllene emission in the *gCST*-OE lines was 429.7-fold higher than in WT plants, while average β-caryophyllene emission in the *cCST*-OE lines was 7.4-fold higher than in WT plants. Therefore, average β-caryophyllene emission in *gCST*-OE lines was 58-fold higher than in *cCST*-OE lines, which is substantially more than the 26.7-fold difference in *CST* transcript levels between these two groups of lines (Fig. 3A and Fig. 4A-C). In individual lines, the amount of β-caryophyllene emitted matched well with the *CST* mRNA level. For instance, *gCST*-3 OE showed highest β-caryophyllene emission (Fig. 3A) and also the highest *CST* mRNA level (Fig. 2A). Besides β-caryophyllene, CST enzyme activity also produces some minor products, i.e. α-humulene and α-copaene (Chen et al., 2003). Figure 4A shows representative mass-traces of the headspace analysis of WT, *gCST*-OE and *cCST*-OE lines and the mass-spectra of the three products from CST enzyme activity. Although a peak is visible in the headspace of WT and *cCST*-OE at the position of the α-copaene peak in *gCST*-OE headspace (peak 4, 5, 3 respectively, Fig. 4A), the mass spectrum of peak 4 in WT is not from α-copaene. The predicted molecular composition of compound 4 in WT is $C_{15}H_{24}$, thus compound 4 is most likely another sesquiterpene from endogenous enzyme activity (Fig. 4B). In *cCST*-OE lines the mass spectrum of peak 4 position shows a combination of the mass spectra of the unknown compound and the mass spectrum of α-copaene (Fig. 4B). Curiously, in the *gCST*-OE lines the mass spectrum of the unknown compound is no longer detected and only the mass spectrum of α-copaene is detected. This may indicate that high ectopic expression in *gCST-OE* plants results in such high CST enzyme activity that it outcompetes the other endogenous sesquiterpene synthase which produces unknown compound 4. Therefore, the semi-quantification of the α-copaene peak intensity may be overestimated in cCST-OE lines (Fig. 3C). For the *gCST*-OE lines the increase in β-caryophyllene production was higher than the increase in *CST* transcript level compared with that in *cCST*-OE lines, suggesting that addition post-tran-
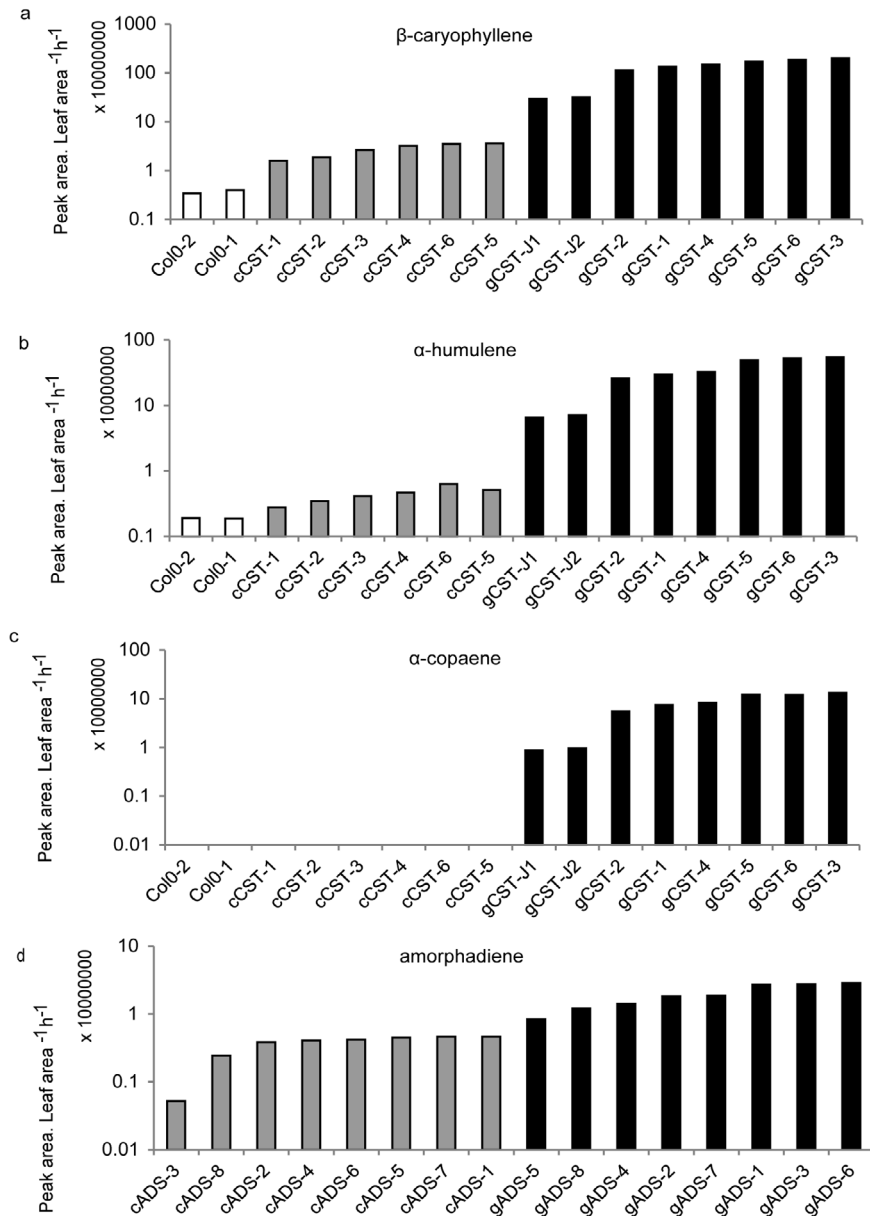
Fig. 3. Volatiles emission of caryophyllene synthase overexpression (*CST*-OE) and amorph-adiene synthase overexpression (*ADS*-OE) lines in *Arabidopsis*. a, β-caryophyllene emission; b, α-humulene emission; c, α-copaene emission; d, amorphadiene. The headspace of WT plants (Col-1, Col-2), *cCST*-OE lines, *cADS*-OE lines and *gADS*-OE lines was trapped for 24 hours, while the headspace of *gADS*-OE lines was trapped for 4 hours, using Tenax TA.
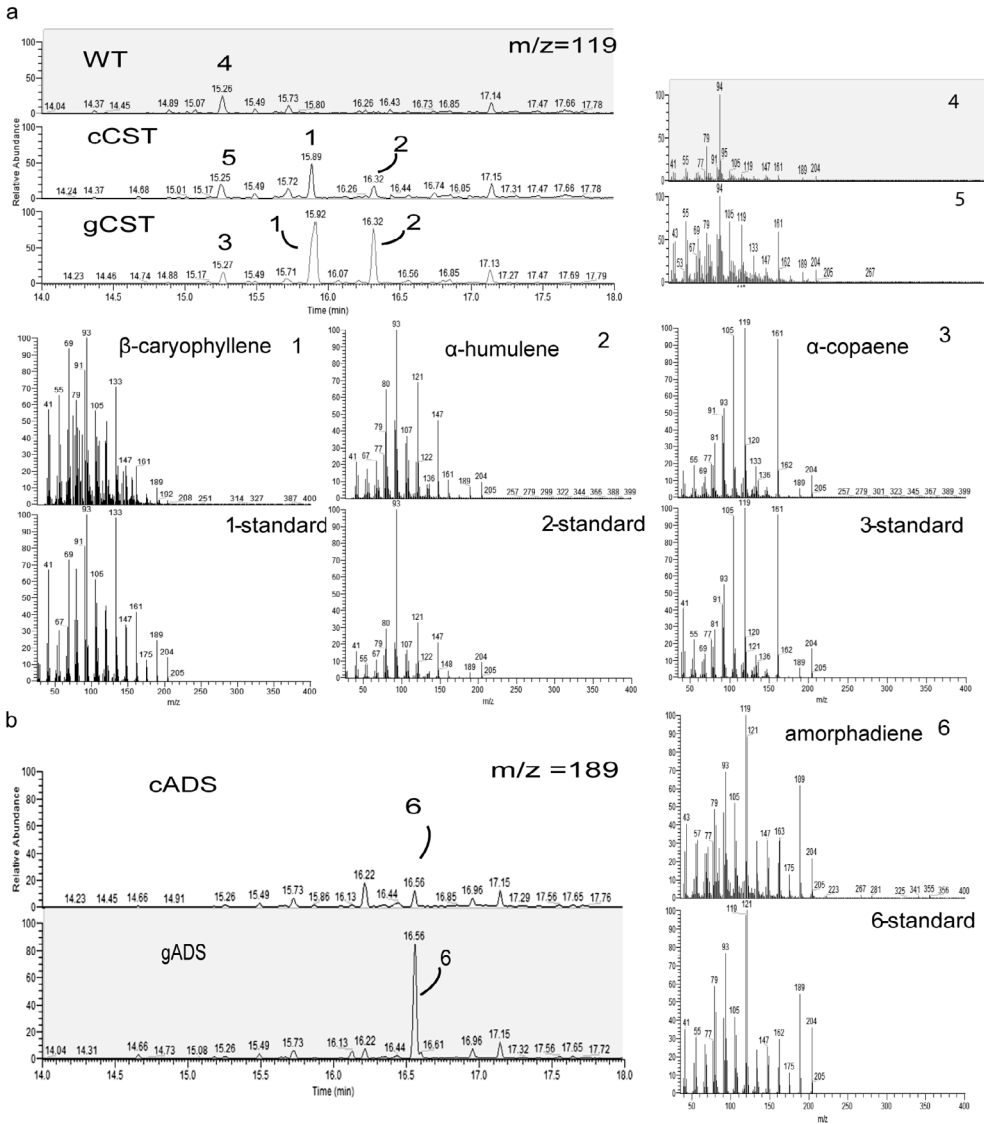
Fig. 4. GC-MS mass-traces for cDNA of caryophyllene synthase overexpression (*cCST*-OE) lines, genomic caryophyllene synthase overexpression (*gCST*-OE) lines at m/z=119, for cDNA of amorphadiene synthase overexpression (*cADS*-OE) lines and genomic amorphadiene synthase overexpression (*gADS*-OE) lines at m/z=189 in *Arabidopsis*. a, chromatogram of WT, for *cCST*-OE lines, *gCST*-OE line at m/z=189; Mass spectrum for compounds: 1, β-caryophyllene, 2, α-humulene and 3, α-copaene, 4, unknown, 5, unknown; b, chromatogram of cADS-OE and gADS-OE lines at m/z=189; Mass spectrum for compound: 6, amorphadiene. Note: mass-trace at m/z=119 was shown to clear show the α-copaene peak.

scriptional mechanisms may contribute to the higher emission rates from plants expressing genomic clones compared with plants expressing cDNAs.

Analysis of the headspace volatiles of the *gADS*-OE and *cADS*-OE lines shows that amorphadiene emission was also consistently higher in *gADS*-OE lines (Fig. 3D and Fig. 4F-G). The average amorphadiene emission in the *gADS*-OE lines was 5.4-fold higher than in *cADS*-OE lines (Fig. 3D), while the average *ADS* transcript level in *gADS*-OE lines was 11-fold higher than in *cADS*-OE lines (Fig. 2B). The *A. annua ADS* in *Arabidopsis* therefore seems to show less benefit from a post-transcriptional intron effect on overall gene expression, than the post-transcriptional intron effect on the *Arabidopsis CST* gene.

We note that (based on peak area in mass-traces), amorphadiene emission is only 1.2% of the average β-caryophyllene emission in transgenic *Arabidopsis* (Fig. 3A-D), while standards of β-caryophyllene and amorphadiene indicate that both compounds are detected with similar efficiency. We therefore considered that amorphadiene may be converted by endogenous *Arabidopsis* Cytochrome P450 or other enzymes to non-volatile amorphadiene derived products, just as in *A. annua*. In *A. annua,* amorphadiene is modified by an Amorphadiene Oxidase P450 enzyme (ADO) to non-volatile artemisinic acid (Teoh et al. 2006). Moreover, in a heterologous expression system oxidized amorphadiene could be further subject to glycosylation, as happens during transient expression of *ADS* and *ADO* in *N. benthamiana* (Ting et al. 2013; Van Herpen et al. 2010). Therefore, leaf extracts from both *gADS*-OE and *cADS*-OE *Arabidopsis* lines were analyzed by LC-QTOF-MS. However, in leaf extracts, no ADS-derived compounds were detected (Fig. S1).

Introns show less effect in the transient assay in *N. benthamiana*

We have shown that in stable transformants the expression of sesquiterpene synthases may be boosted by making use of genomic clones, which contain introns, rather than making use of a cDNA. The transient expression in *N. benthamiana* by agro-infiltration of terpene pathway genes is an efficient and fast way to reconstruct whole metabolic pathways *in planta* (Ting et al. 2013; Van Herpen et al. 2010). However, we have also experienced that when multiple individual constructs are co-expressed, the yield decreases due to gene dosage dilution effects. Because transient expression for pathway elucidation could benefit from intrinsic higher expression in agro-infiltrated leaves, we tested whether intron-mediated-enhancement also works in the context of transient assays in *N. benthamiana*. For this, *N. benthamiana* leaves were agro-infiltrated with either the *gCST*, *cCST*,
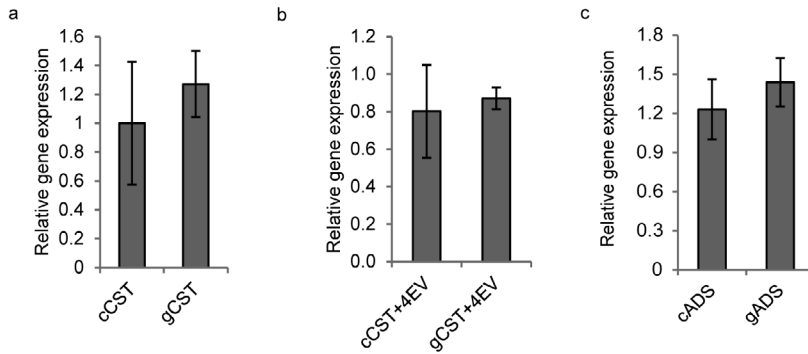
a

b

c

Fig. 5. Transcription levels of caryophyllene synthase gene (*CST*) and amorphadiene synthase gene (*ADS*) in leaves of *N. benthamiana* infiltrated with *cCST, gCST, cADS* and *gADS*. a, gene expression of *CST* with the non-diluted gene dosage, n=4; b, gene expression of *CST* with 5 x diluted gene dosage, n=9; c, gene expression of *ADS*, n=3.

gADS or *cADS* expression constructs and leaves harvested at 5 dpi were used to measure volatile emission. Results show that transient expression of *gCST* in *N. benthamiana* results in similar β-caryophyllene emission as for *cCST* expression (Fig. 6A). Also, there is no significant difference in *CST* transcript levels between the two treatments (Fig. 5A). When the dosage of agro-infiltrated *cCST* or *gCST* is reduced, by diluting agro-infiltrated bacteria with *Agrobacterium* carrying an empty vector (EV), overall emission levels are reduced, but caryophyllene emission in leaves with gCST expression are slightly higher than in leaves with cCST expression (Fig. 5B).
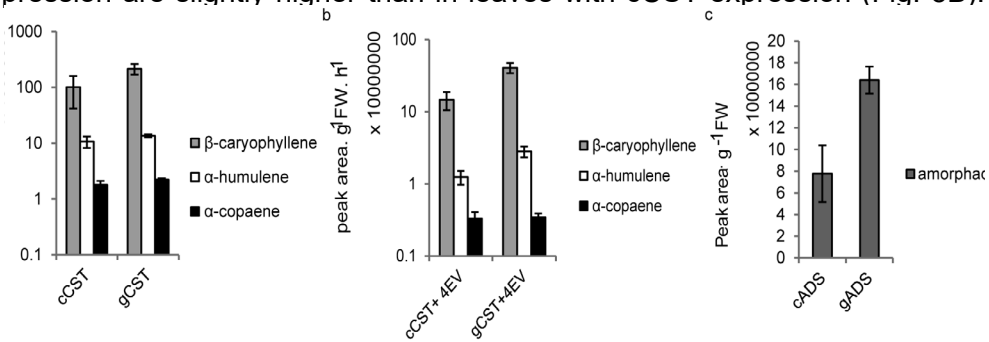

b

c

Fig. 6. Volatile analysis on leaves of *N. benthamiana* infiltrated with cDNA of caryophyllene synthase *(cCST),* genomic caryophyllene synthase *(gCST),* cDNA of amorphadiene synthase *(cADS)* and genomic amorphadiene synthase *(gADS)* genes as identified by GC-MS at 5dpi. a, β-caryophyllene, α-humulene and α-copaene emission with the non-diluted gene dosage, n=4; b, β-caryophyllene, α-humulene and α-copaene emission with 5 x diluted gene dosage, n=9; c, amorphadiene emission, n=3. The headspace of infiltrated leaves was trapped for 4 hours using Tenax TA.

However, there is still no significant difference between CST mRNA levels in *gCST* and *cCST* agro-infiltrated leaves (Fig. 5B). In contrast, amorphadiene emission was increased 2-fold in *N. benthamiana* leaves expressing *gADS* compared with leaves expressing *cADS* (Fig. 6C), although *ADS* transcript levels were not significantly different between the two treatments (Fig. 5C).

Discussion

Introns enhance the expression of sesquiterpene biosynthesis genes

Here we have shown that engineering of ectopic terpene production in plants may benefit from the use of genomic clones containing introns, compared to ectopic expression of cDNAs. The intron-enhancement effect was demonstrated for two sesquiterpene synthase genes, *CST* and *ADS* (Fig. 3). In the *gCST*-OE lines the enhancement in product formation was higher than the increase in *CST* expression, suggesting additional post-transcriptional effects besides an effect on promoter activity as has been described before for IME (Parra et al. 2011; Rose et al. 2008). However, although *ADS* expression was also enhanced by the presence of introns, the increase in amorphadiene production was lower than the increase in *ADS* transcription. This indicates that for this gene the post-transcriptional effects of IME do not seem to work. When we assume that for both the *Arabidopsis gCST* and *A. annua gADS* the *Arabidopsis* pre-mRNA maturation protein complexes and exon-junction protein complexes on mRNA are identical, this suggest an additional role for mRNA exon sequences or a role for spatial distribution of exon-junction protein complexes on mRNA in post-transcriptional IME effects. Apparently this post-transcriptional IME effect is less effective for the heterologous *ADS* mRNA sequence than for the endogenous *CST* mRNA sequence.

Lower ectopic amorphadiene production than ectopic β-caryophyllene production in *Arabidopsis*

When quantified by peak area, the average amorphadiene emission amounts to only 1.2% of the average β-caryophyllene emission in transformed *Arabidopsis* (Fig. 3A-D). This difference in emission is not explained by conversion of amorphadiene to non-volatile products, as no ADS-related compounds were detected by LC-QTOF-MS in a leaf extract (Fig. S1). The K_m of CST recombinant enzyme for farnesyl diphosphate (FPP) (2.1 µM (Chen et al. 2003; Tholl et al. 2005) is higher than that of ADS for FPP

(0.6-0.9 µM: (Bouwmeester et al. 1999; Mercke et al. 2000). This difference in K_m CST and K_m ADS does not explain the quantitative difference in sesquiterpene headspace production between the *gADS*-OE and *gCST*-OE lines. An alternative explanation for the low ADS activity in *Arabidopsis* may be that endogenous enzymes (e.g. CST) form a metabolon complex with FPPS, thereby limiting availability of FPP for ectopically produced ADS.

Introns have no clear effect in transient expression

Although introns enhance gene expression and sesquiterpene emission in stable transformed transgenic *Arabidopsis*, in transient assays in *N. benthamiana* leaves the effect of introns is much less obvious. There is no clear difference in transcript levels for ectopic expression of genomic and cDNA fragments. However, for *ADS* in the transient assay in *N. benthamiana* there is a 2-fold increase in amorphadiene production, indicating that there still is a small post-transcriptional effect from IME, which is not observed in stable transformed *Arabidopsis* (Fig. 3D). Also, when gene dosage was reduced, the effect of IME for in *N. benthamiana* became more apparent for *gCST* and *cCST* (Fig. 6B). Therefore, the artificial high expression levels under transient expression conditions seem to mask effects of IME manifested at normal gene dosage in stable transformed plants.

Variable product ratios from CST activity in stable transformants

The ratio of the three products (β-caryophyllene:α-humulene:α-copaene) was different in WT and *gCST*-OE and *cCST*-OE lines (Fig. 7A-B). The α-humulene proportion in the headspace of rosette plants was significantly increased in *gCST*-OE lines compared with *cCST*-lines (Fig. 7A). In untransformed WT plants CST activity was highest in flowers (Huang et al. 2012) and headspace analysis from inflorescences showed that the β-caryophyllene:α-humulene:α-copaene ratio in the headspace of WT flowers is similar to the β-caryophyllene:α-humulene:α-copaene ratio in the headspace of *gCST*-OE rosette leaves (Fig. 7A-B). The relative high proportion of α-humulene in the rosette leaf headspace of untransformed WT *Arabidopsis* plants may be due to overestimation of the α-humulene peak areas which was close to background levels (Fig. 4A).

In addition, product ratios may be influenced by saturation of the trapping agent. Indeed different trapping methods (Tenax vs. twister) resulted in different product ratios (Fig. 7B). Saturation of the trapping agent may be pre-
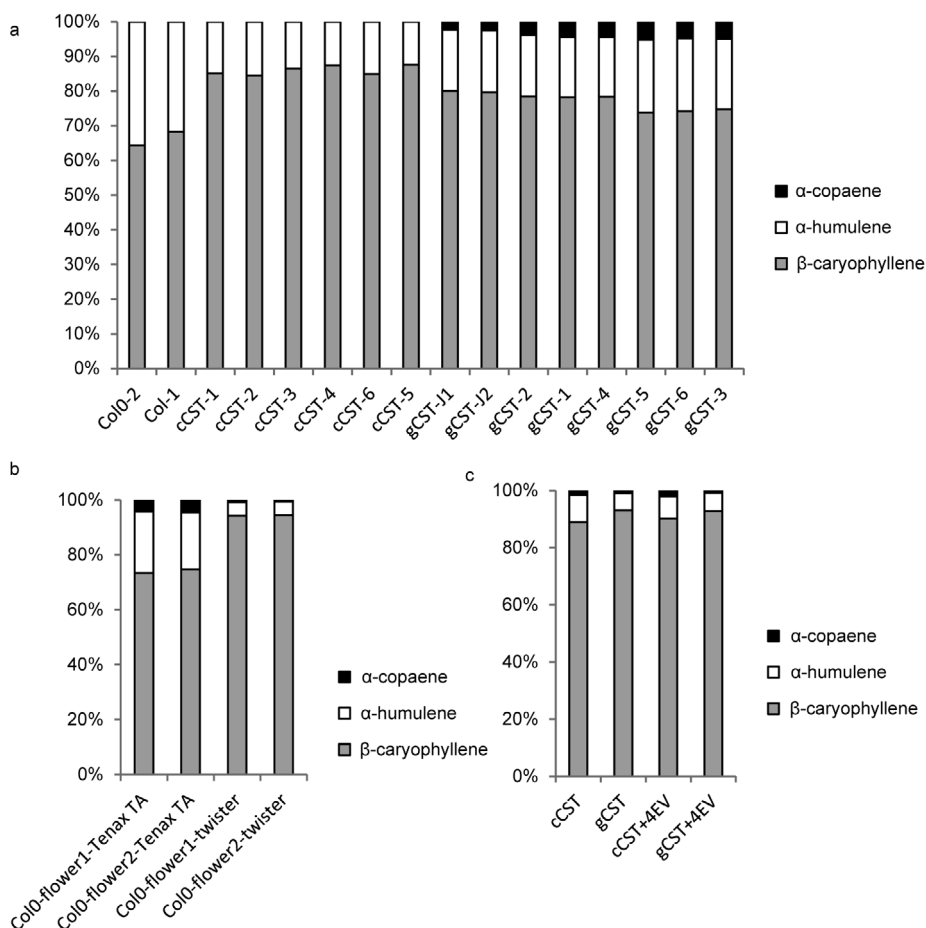
Fig. 7. Ratio of β-caryophyllene, α-humulene and α-copaene. a, WT, cDNA of caryophyllene synthase (*cCST*-OE*)*, genomic caryophyllene synthase (*gCST*-OE*)* overexpression lines; b, wild-type flower tissues using different trapping agents (Tenax TA and twister) with 24 hour trapping; c, transient expression in leaves of *N. benthamiana* infiltrated with *cCST, gCST*. The headspace of WT plants (Col-1, Col-2), *cCST*-OE lines, *cADS*-OE lines and *gADS*-OE lines was trapped for 24 hours, while the headspace of *gADS*-OE lines was trapped for 4 hours using Tenax TA.

vented by varying the headspace trapping interval time or by varying the gene dosage (e.g. in the transient assays). The peak intensities for caryophyllene were decreased 5- to 8-fold when the *cCST* and *gCST* gene dosages were diluted 5-fold upon transient expression in *N. benthamiana* (Fig. 6 A-B), however, expression of *cCST* or *gCST* at different gene dosage does not affect product ratio (Fig. 7C), suggesting that variation in product ratio is not related to volatile trapping saturation. Besides putative technical reasons

for variation in headspace product ratio from CST activity, it could also be that variation in products is related to different forms of CST derived from alternative splicing of the genomic *CST* gene. Two gene isoforms for *CST*, *AT5G23960.1* and *AT5G23960.2*, have been reported in the TAIR database. These isoforms proteins show a small difference near the conserved catalytic side and therefore the two forms could differ in product specificity. However, no *CST* splice variant *AT5G23960.2* was detected by us either in leaf RNA from untransformed WT plants or in leaf RNA from *gCST*-OE plants.

Conclusions

Here we have demonstrated that the use of intron containing genes is an efficient way to up-regulate heterologous terpenoid production in stably transformed transgenic plants, while the IME effect in transient expression is virtually absent.
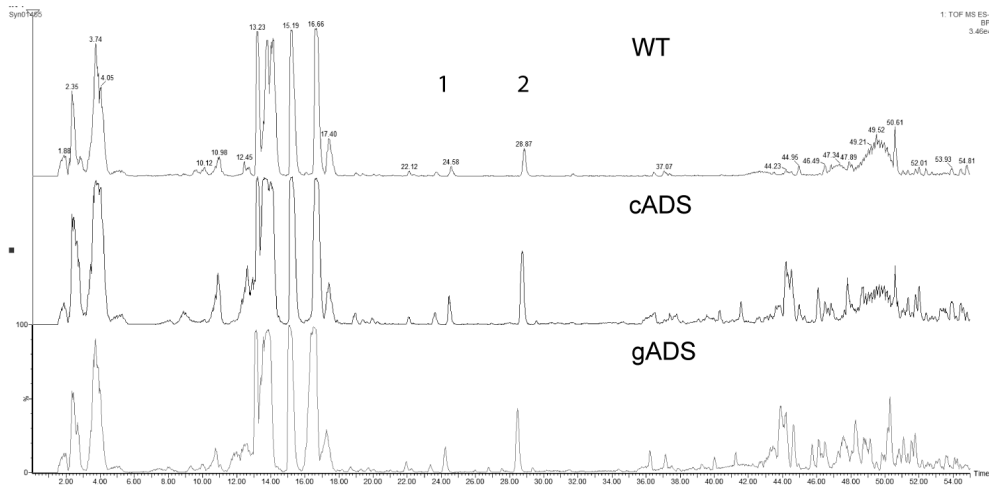
Supplementary data

Fig. S1. LC-QTOF-MS chromatograms for WT and cDNA of amorphadiene synthase (*cADS*-OE), genomic amorphadiene synthase (*gADS*-OE) lines in *Arabidopsis*. No ADS-related peaks were detected, Peak 1 and 2 are not significantly different between the samples (n=3).

Supplementary Fig.S1. LC-QTOF-MS chromatograms for amorphadiene synthase lines in *Arabidopsis*.

Acknowledgments

Reference

Bertea CM, Voster A, Verstappen FWA, Maffei M, Beekwilder J, Bouwmeester HJ (2006) Isoprenoid biosynthesis in *Artemisia annua*: cloning and heterologous expression of a germacrene A synthase from a glandular trichome cDNA library Arch Biochem Biophys 448:3-12 doi:http://dx.doi.org/10.1016/j.abb.2006.02.026

Bouwmeester HJ et al. (1999) Amorpha-4,11-diene synthase catalyses the first probable step in artemisinin biosynthesis Phytochemistry 52:843-854 doi:http://dx.doi.org/10.1016/S0031-9422(99)00206-X

Braddock M, Muckenthaler M, White MR, Thorburn AM, Sommerville J, Kingsman AJ, Kingsman SM (1994) Intron-less RNA injected into the nucleus of *Xenopus oocytes* accesses a regulated translation control pathway Nucleic Acids Res 22:5255-5264

Chen F, Tholl D, D'Auria JC, Farooq A, Pichersky E, Gershenzon J (2003) Biosynthesis and emission of terpenoid volatiles from *Arabidopsis* flowers Plant Cell 15:481-494 doi:10.1105/tpc.007989

Gunderson SI, Polycarpou-Schwarz M, Mattaj IW (1998) U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase Mol Cell 1:255-264 doi:http://dx.doi.org/10.1016/S1097-2765(00)80026-X

Harrison SJ, Mott EK, Parsley K, Aspinall S, Gray JC, Cottage A (2006) A rapid and robust method of identifying transformed *Arabidopsis thaliana* seedlings following floral dip transformation Plant Methods 2:19-19 doi:10.1186/1746-4811-2-19

Houshyani B, Assareh M, Busquets A, Ferrer A, Bouwmeester HJ, Kappers IF (2013) Three-step pathway engineering results in more incidence rate and higher emission of nerolidol and improved attraction of *Diadegma semiclausum* Metab Eng 15:88-97 doi:http://dx.doi.org/10.1016/j.ymben.2012.10.002

Huang M, Sanchez-Moreiras AM, Abel C, Sohrabi R, Lee S, Gershenzon J, Tholl D (2012) The major volatile organic compound emitted from *Arabidopsis thaliana* flowers, the sesquiterpene (E)-β-caryophyllene, is a defense against a bacterial pathogen New Phytol 193:997-1008 doi:10.1111/j.1469-8137.2011.04001.x

Komarnitsky P, Cho E-J, Buratowski S (2000) Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription Gene

Dev 14:2452-2460

Le Hir H, Gatfield D, Izaurralde E, Moore MJ (2001) The exon–exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay EMBO J 20:4987-4997 doi:10.1093/emboj/20.17.4987

Le Hir H, Nott A, Moore MJ (2003) How introns influence and enhance eukaryotic gene expression Trends Biochem Sci 28:215-220 doi:http://dx.doi.org/10.1016/S0968-0004(03)00052-5

Legrain P, Rosbash M (1989) Some cis- and trans-acting mutants for splicing target pre-mRNA to the cytoplasm Cell 57:573-583 doi:10.1016/0092-8674(89)90127-x

Lewis JD, Izaurflde E (1997) The role of the cap structure in RNA processing and nuclear export Eur J Biochem 247:461-469 doi:10.1111/j.1432-1033.1997.00461.x

Luo M-j, Reed R (1999) Splicing is required for rapid and efficient mRNA export in metazoans P Natl Acad Sci USA 96:14937-14942 doi:10.1073/pnas.96.26.14937

Mascarenhas D, Mettler I, Pierce D, Lowe H (1990) Intron-mediated enhancement of heterologous gene expression in maize Plant Mol Biol 15:913-920 doi:10.1007/bf00039430

Matsumoto K, Wassarman KM, Wolffe AP (1998) Nuclear history of a pre-mRNA determines the translational activity of cytoplasmic mRNA EMBO J 17:2107-2121 doi:10.1093/emboj/17.7.2107

Mercke P, Bengtsson M, Bouwmeester HJ, Posthumus MA, Brodelius PE (2000) Molecular cloning, expression, and characterization of amorpha-4,11-diene synthase, a key enzyme of artemisinin biosynthesis in *Artemisia annua* L Arch Biochem Biophys 381:173-180 doi:http://dx.doi.org/10.1006/abbi.2000.1962

Moabbi AM, Agarwal N, El Kaderi B, Ansari A (2012) Role for gene looping in intron-mediated enhancement of transcription P Natl Aaca Sci USA 109:8505-8510 doi:10.1073/pnas.1112400109

Niu D-K, Yang Y-F (2011) Why eukaryotic cells use introns to enhance gene expression: splicing reduces transcription-associated mutagenesis by inhibiting topoisomerase I cutting activity Biol Direct 6:24-24 doi:10.1186/1745-6150-6-24

Parra G, Bradnam K, Rose AB, Korf I (2011) Comparative and functional analysis of intron-mediated enhancement signals reveals conserved features among plants Nucleic Acids Res 39:5328-5337 doi:10.1093/nar/gkr043

Proudfoot NJ, Furger A, Dye MJ (2002) Integrating mRNA processing with transcription Cell 108:501-512 doi:http://dx.doi.org/10.1016/S0092-8674(02)00617-7

Reenan RA (2001) The RNA world meets behavior: A→I pre-mRNA editing in animals Trends Genet 17:53-56 doi:http://dx.doi.org/10.1016/S0168-9525(00)02169-7

Rose AB, Elfersi T, Parra G, Korf I (2008) Promoter-proximal introns in *Arabidopsis thaliana* are enriched in dispersed signals that elevate gene expression Plant Cell

20:543-551 doi:10.1105/tpc.107.057190

Sowden MP, Smith HC (2001) Commitment of apolipoprotein B RNA to the splicing pathway regulates cytidine-to-uridine editing-site utilization Biochem J 359:697-705

Teoh KH, Polichuk DR, Reed DW, Nowak G, Covello PS (2006) *Artemisia annua* L. (*Asteraceae*) trichome-specific cDNAs reveal CYP71AV1, a cytochrome P450 with a key role in the biosynthesis of the antimalarial sesquiterpene lactone artemisinin FEBS Lett 580:1411-1416 doi:http://dx.doi.org/10.1016/j.febslet.2006.01.065

Tholl D, Chen F, Petri J, Gershenzon J, Pichersky E (2005) Two sesquiterpene synthases are responsible for the complex mixture of sesquiterpenes emitted from *Arabidopsis* flowers Plant J 42:757-771 doi:10.1111/j.1365-313X.2005.02417.x

Ting H-M et al. (2013) The metabolite chemotype of *Nicotiana benthamiana* transiently expressing artemisinin biosynthetic pathway genes is a function of CYP71AV1 type and relative gene dosage New Phytol 199:352-366 doi:10.1111/nph.12274

Vagner S, Vagner C, Mattaj IW (2000) The carboxyl terminus of vertebrate poly(A) polymerase interacts with U2AF 65 to couple 3'-end processing and splicing Gene Dev 14:403-413

Van Herpen TW, Cankar K, Nogueira M, Bosch D, Bouwmeester HJ, Beekwilder J (2010) *Nicotiana benthamiana* as a production platform for artemisinin precursors PLoS One 5:e14222 doi:10.1371/journal.pone.0014222

Voinnet O, Rivas S, Mestre P, Baulcombe D (2003) An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus Plant J 33:949-956 doi:10.1046/j.1365-313X.2003.01676.x

Chapter 6

General Discussion

Artemisinin (AN) is an important sesquiterpenoid molecule for the treatment of malaria produced in the plant *Artemisia annua*. The biosynthesis pathway of AN is branched, yielding both AN and arteannuin B (AB) and different *A. annua* chemotypes are recognized based on either high AN and low AB (high artemisinin plants, HAP) or low AN and high AB content (low artemisinin plants, LAP). However, due to the low natural concentration of AN, even in *A. annua* of the HAP chemotype, yields of AN do not meet the pharmaceutical demand. In order to improve the production of AN, it is important to understand AN biosynthesis and transport and storage of AN and its precursors in order to optimally apply available technologies for metabolic engineering. In this thesis, options for AN production in a heterologous plant host (*Nicotiana benthamiana*), using transient expression of AN pathway (AN-PW) genes, were explored and used to validate various strategies which may enhance AN product accumulation. In Chapter 2 I transiently expressed AN-PW genes in *N. benthamiana* to identify the mechanistic basis for the presence of an AB and AN branch in the pathway and showed that there are gene variants that are responsible for these different 'chemotypes'. In Chapter 3 I used transient expression of AN-PW genes to characterize putative AN-PW product related transporter and carrier genes, leading to the identification of LTPs with specific sequestering activity. In Chapter 4 I used transient expression of AN-PW genes to study the interaction between the flavonoid and sesquiterpenoid biosynthesis pathways. Finally, In Chapter 5 I used transient expression of AN-PW genes to study the function of introns for enhancement of gene expression efficiency.

Here, in Chapter 6, I discuss the results in my thesis in a broader perspective. I will discuss major limitations we have encountered and potential solutions to metabolic engineering that still may be explored. Moreover, some of our findings go beyond engineering of the AN-PW and have implications for metabolic engineering in general (e.g. control of glycosylation, compartmentalization and sequestering of secondary metabolites, choice of host plant for heterologous production of secondary metabolites). Also, the current yields of AN in wildtype *A. annua* and in different other, heterologous, hosts through metabolic engineering will be compared. Finally, I will discuss how some results may also be of relevance to transport of plant hormones and therefore link my work to a more general understanding of plant physiology and development.

Evaluation of transient expression for metabolic pathway analysis

Throughout this thesis we have made use of the transient expression of genes in *N. benthamiana*. Most of the secondary metabolites are synthesized by multiple enzymes, and the reconstruction of an entire secondary metabolite pathway therefore requires the coordinated expression of multiple transgenes in a single cell. The transient expression system of *N. benthamiana* allows for quick assessment of PW genes as not all genes need to be in the same expression construct, but can be cloned in individual expression vectors and subsequently *N. benthamiana* leaves can be agro-infiltrated with all constructs at once by mixing the different *A. bacterium* strains containing these expression vectors. The fact that PW end products are produced after co-agro-infiltration in *N. benthamiana* leaves indicates that cells are transformed by multiple *A. bacterium* strains, which results in expression of multiple transgenes and full reconstitution of the PW within a single cell. Examples the use of multiple *A. bacterium* with different expression vectors are the reconstitution of the AN pathway in *N. benthamiana* through coexpression of 4 PW genes (this thesis), the parthenolide biosynthesis pathway in *N. benthamiana* though co-expression of 6 PW genes (Liu et al., 2014) and reconstitution of the secologain biosynthesis PW in *N. benthamiana* though co-expression of even 12 PW genes (Miettinen et al., 2014). However, in the latter case the total flux through the PW was very much reduced and needed to be boosted by infiltration of PW intermediates to obtain detectible levels of the end product (Miettinen et al., 2014). It could be argued that the more different *A. bacteria* are co-infiltrated into the *N. benthamiana* leaf, the higher the chances are that some of the cells will not be expressing all transgenes, resulting in a mixture of full and partial PW activity.

In this thesis, the combination of (1) *AaLTP3+AaPDR2+AN-PW* showed the highest free product formation, while the combination of (2) *AaLTP1+AaLTP2+AaLTP3+2xEV+AN-PW* and (3) *AaPDR1+AaPDR2+3xEV+AN-PW* did not increase the (DH)AA (Chapter 3). The difference is that in combination (1) the PW genes are co-infiltrated with just 2 additional agrobacterium strains, while in combinations (2) and (3) the PW genes are co-infiltrated with 5 additional agrobacterium strains. The final relative gene dosage of the AN-PW genes may therefore affect total end-product accumulation and that is why in a single experiment, the relative dosage of the AN-PW genes was always kept constant.

Since in all experiments we find both early (AAOH, AAA, DHAAOH, DHAAA) and late intermediates of the AN pathway (AA, DHAA), this raises the question whether early intermediate accumulation is caused by incomplete PW reconstitution in cells. Also, because in Chapter 2 we did experiments in *N. benthamiana* expressing only AN-PW genes, while in Chapter 4 we did experiments expressing AN-PW+*AaPDR2*+*AaLTP3*+*ROS*, the relative dosage of AN-PW genes between these experiment is different. Therefore, if comparing the product levels obtained in different experiments we need to take into account the different relative AN-PW gene dosage that has been used.

To obtain insight into the effect of relative gene dosage we tested the AN-PW gene set, consisting of the *V5* construct *(ADS+FPS+HMGR)* in combination
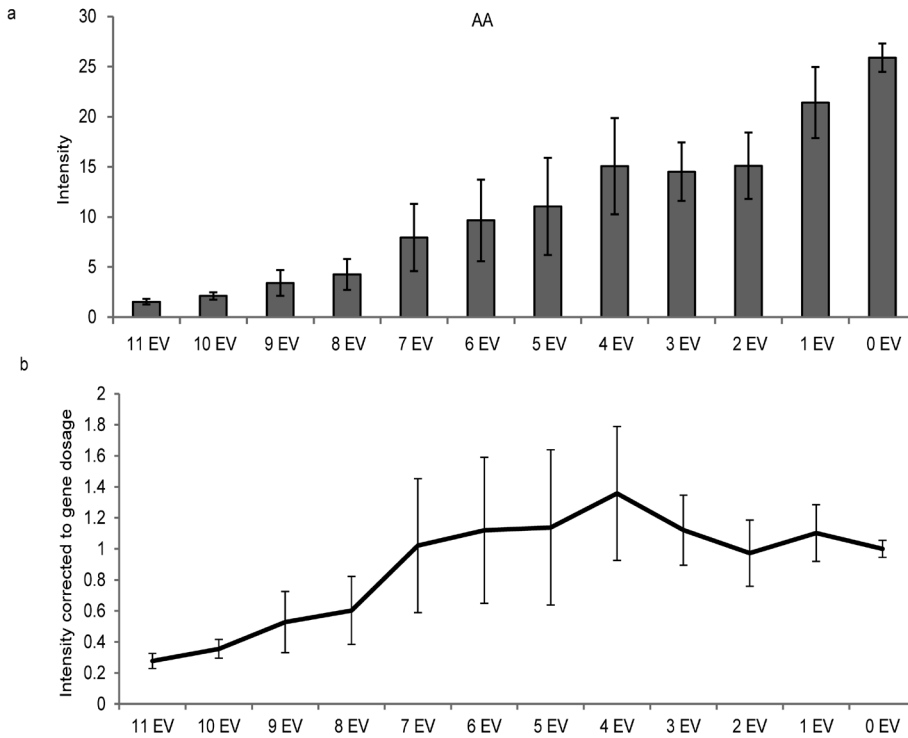

Figure 1. Effect of relative gene dosage on the product accumulation in *N. benthamiana*. A, Effect of relative dosage of agrobacterium carrying AN-PW genes (*V5+ADO*) by combining with different dosage of agrobacterium carrying EV on artemisinic acid (AA) product accumulation in *N. benthamiana*. B, Same data as in A but each value was corrected for the decrease in AN-PW gene dosage. Dilution with up to 7xEV can be explained by simple gene dosage effects. Beyond that correction for gene dosage is not sufficient and additional factors contribute to lower product levels (e.g. lack of overlap in expression of cells expressing both *ADS* and *ADO*).

with amorphadiene oxidase (*ADO*)+*P19* with different dosages of agrobacterium carrying *EV* through agro-infiltration in *N. benthamiana*. Results indicate that relative AN-PW gene dosage indeed has an effect on AA accumulation (Figure 1a). When product yield is corrected for the difference in AN-PW-gene dosage the results show that up to a dilution with 7xEV the change in yield is explained by the reduced AN-PW gene dosage. Beyond that point the reduction in PW product level is greater than is explained by the EV dilution and other factors interfere with efficient PW activity. For instance, beyond dilution with 7xEV the diminished PW product accumulation may be explained by decreased overlap of cells expressing both *ADS* and *ADO*. This confirms that there is a dilution effect from co-expressing increasingly more transgenes and this should be taken into account when comparing yields between experiments where different number of transgenes are used. Perhaps at a low dilution of the AN-PW there may also be already some lack of coexpression of *ADS* and *ADO* in the same cell, where exchange of amorphadiene between cells expressing *ADS* and cells expressing *ADO* compensates for that. At stronger dilution of AN-PW genes the distance between *ADS* and *ADO* expressing cells may become too big for this compensating effect.

Competing reactions in the AN pathway in *N. benthamiana* leaves

One of the most revealing experiments in this thesis is the experiment in which I infiltrated (DH)AA into the *N. benthamiana* leaves because, first of all, subsequent analysis of leaves shows that (DH)AA is very rapidly taken up by cells. This made us realize that, although transporters are involved in the secretion of (DH)AA produced inside the cells, these products very rapidly diffuse back into the cell, making the action of the PDR transporters futile if there is no additional mechanism for sequestering the product away from the plasma membrane. Secondly, analysis of what happens to infiltrated (DH)AA showed a rapid conversion towards (DH)AAA and (DH)AAOH inside cells, revealing endogenous enzymes that act in a reverse flux through the AN-PW (Chapter 3). Presumably these enzymes are non-specific alcohol dehydrogenase enzymes. A similar reverse AN-PW activity has also been reported for *A. annua,* as the enzyme oxidoreductase1(RED1) converts DHAAA into DHAAOH and possibly also AAA to AAOH (Rydén et al., 2010). Such reverse pathway flux is limiting AN-PW product accumulation and could contribute to competing reactions in *N. benthamiana* by providing (DH)AAOH

for glycosyltransferases. Glycosylation of (DH)AAOH and (DH)AA may be a way of detoxification of these heterologous compounds in *N. benthami-ana*. Glycosylation makes the compound more hydrophilic and therefore more water soluble (Rivas et al., 2013). It could be that glycosylation of the end product (DH)AA enhances the metabolic flux through the AN-PW in *N. benthamiana* as the bulk of AN-PW products is in the form of glycosides (Chapter 2). For instance, treatment of *N. benthamiana* leaf extracts with the glycosidase (Viscozyme L.) showed that up to 10-fold more (DH)AA is present in the form of glydosides compared to free (DH)AA (Chapter 2). *In vitro* biochemical studies indicate that multiple glycosyltransferases from the same plant can glycosylate the same compounds, and that individual glycos-yltransferases have the capacity to glycosylate multiple substrates (Osbourn et al., 2009). Lipophilic terpenes may diffuse passively over the membranes (Widhalm et al., 2015) as also suggested by our infiltration studies of (DH)AA (Chapter 3). Glycosylation of AN-PW products may prevent free diffusion of these modified compounds over membranes separating the different cellular compartments and therefore glycosylation affects where products eventually accumulate and whether they are still available for final conversion to AN or AB (if de-glycosylation also occurs). However, glycosylation of pathway intermediates such as DHAAOH also prevents further flux through the path-way. Therefore, suppression of glycosylation in *N. benthamiana* could be one of the targets for future research on AN-PW engineering, but the lack of specificity and presence of multiple glycosyltransferases in cells will make it challenging to control AN precursor glycosylation. The presence of a very active reverse PW flux and the presence of competing side reactions like the glycosylation which occurs in *N. benthamiana* both stress the need for effi-cient AN-PW product sequestering activity to escape from these competing reactions and to give a directional flow to the AN-PW. In addition, efficient sequestering of terpenes into the apoplast may be needed to prevent tox-ic effects of these compounds inside the cell. Therefore, future researched aimed at increasing AN-PW yield in the *N. benthamiana* production platform should target identifying and inhibiting genes encoding enzymes involved in reverse AN-PW activity, identification and inhibition of genes encoding gly-cosyltransferases acting on AN-PW products, and by enhancing sequester-ing activity into the apoplast, e.g. by enhancing transport out of the cell by PDRs and enhancing apoplast sequestering activity with LTPs (Chapter 3).

Effect of ROS on AN-PW glycosylation profile

In Chapter 4 we studied the interaction between the AN-PW and the fla-
vonoid biosynthesis pathway, as activated by ectopic expression of ROS.
An indirect effect of flavonoids on terpene biosynthesis was suggested by
studies in tomato where it was shown that a mutation in the gene encoding
the flavonoid biosynthetic enzyme chalcone isomerase (*CHI*) also results
in reduced terpene production in leaves (Kang et al., 2014). However, this
effect was based on an effect of flavonoids on type VI trichome develop-
ment in which terpene biosynthesis takes place and in plants with mutated
CHI the type VI trichomes never reach a stage at which the terpene biosyn-
thesis pathway is activated. Therefore, the reduced terpene production in
CHI mutants cannot be ascribed to a direct interaction between these two
pathways in the same cell (Kang et al., 2014). In our transient expression
studies in *N. benthamiana* leaves ROS can stimulate AN-PW product ac-
cumulation. Because expression of all these genes is in already fully estab-
lished leaves we do not expect that at this stage *ROS* is having any effect
on trichomes where AN-PW genes could be more effective. Therefore, the
stimulating effects of *ROS* on AN-PW must have other reasons and indeed I
identified induction of Mevalonate Kinase (*MVK*) by *ROS*, which is involved
in the sesquiterpene precursor pathway (Chapter 4). Curiously, the concen-
tration of glycosides of (DH)AA and (DH)AAOH was enhanced 2- to 3-fold in
leaves expressing AN-PW+ *ROS* compared to expression of only the AN-PW
(Chapter 4). This effect of *ROS* was reduced when at the same time flavo-
noid biosynthesis was blocked by co-expression with the *CHS^RNAi* construct.
This suggests that it may be the flavonoid compounds themselves which
affect glycosylated AN-PW product accumulation (Chapter 4). It is known
from cosmetics studies that terpenes and flavonoid molecules interact to in-
crease skin penetration (Saija et al., 1998) but more research is needed to
determine whether a direct interaction between sesquiterpene and flavonoid
molecules is responsible for the effect on glycosylated product accumulation.

Host organism selection for engineering

During my PhD study, I used different plant host species for overexpression of
AN-PW genes: transient expression of AN-PW in *N. benthamiana* (Chapter
2, 3, 4) and transient expression of *ADS* in *Lactuca sativa* (lettuce; data not
shown) and stable overexpression of *ADS* in *Arabidopsis* (Chapter 5) and sta-
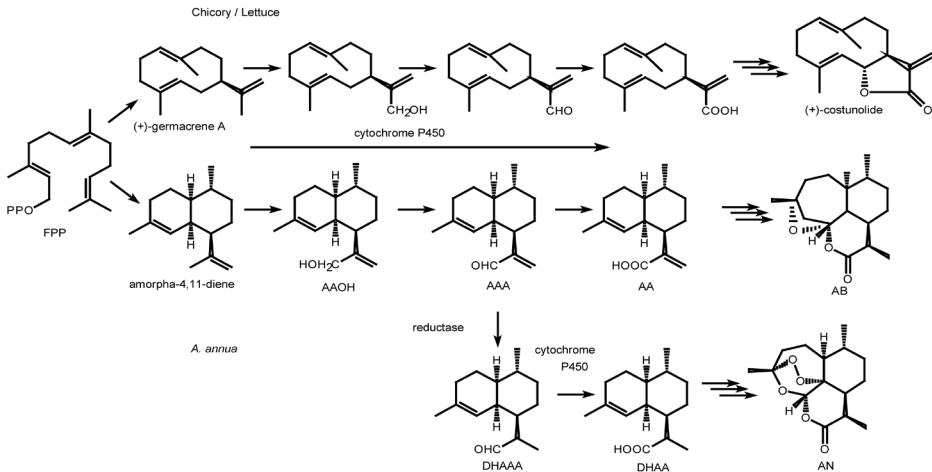
Figure 2 Biosynthetic pathways of sesquiterpenes in chicory, lettuce and *A. annua*. FPP, farnesyl pyrophosphate; AA, artemisinic acid; DHAA, dihydroartemisinic acid; AAOH, artemisinic alcohol; DHAAOH, dihydroartemisinic alcohol; AAA, artemisinic aldehyde; DHAAA, dihydroartemisinic aldehyde; AB, arteannuin B; AN, artemisinin.

ble overexpression of *ADS* in *Cichorium intybus* L. (chicory; data not shown). Chicory, lettuce and *A. annua* belong to the *Asteraceae* family and produce sesquiterpenes derived from amorphadiene (*A. annua*) or germacrene A (lettuce, chicory). The germancrene A derived sesquiterpene lactones in letuce and chicory are responsible for the bitter taste (Bennett et al., 2002; Bouwmeester et al., 2002; Graziani et al., 2015; Ikezawa et al., 2011; Mahmoud et al., 1986). Initial biosynthesis of germacrene A derived sequiterpenes is very similar to AN biosynthesis (Figure 2). Germacrene A synthase (GAS) is the first committed step for sesquiterpene biosynthesis in chicory and lettuce followed by oxidation by Germacrene A oxidase (GAO), a cytochrome P450 enzyme. In fact, it has been shown that chicory or lettuce GAO can act on amorphadiene as well (Nguyen et al., 2010), suggesting that expression of only ADS in either chicory or lettuce may be sufficient to produce ADS-derived sesquiterpenes. However, upon transient expression of the *AN-PW* in lettuce, no ADS derived products could be detected, while control experiments with a *35S:luciferase* or *35S:CST* indicate that expression from an agro-infiltrated gene is possible in lettuce (data not shown). Also chicory, stably transformed with the *ADS* gene, shows that *ADS* is transcribed, but no ADS-derived sesquiterpenes were detected in chicory leaves. One explanation for the lack

of ADS-products in lettuce and chicory may be that endogenous enzymes form a metabolon complex with FPPS, thereby limiting availability of FPP for ectopically producing ADS-product (Figure 2). The lack of apparent ADS activity in lettuce and chicory is not related to affinity for the substrate FPP as the Km for FPP of ADS is about 10-fold lower (0.6 ~ 0.9 μM: (Bouwmeester et al., 1999; Mercke et al., 2000), compared to the Km for FPP of GAS (6.6 ~ 6.9 μM; (Bouwmeester et al., 2002; de Kraker et al., 1998). Alternatively, the amorphadiene from ADS activity in chicory and lettuce is perhaps converted to reactive compounds which crosslink to compounds and/or proteins such that they are no longer visible in the MS chromatograms. Therefore, additional experiments are needed to determine the bottleneck in AN-PW production in lettuce and chicory. For instance, infiltration experiments with AA and/or DHAA in leaves of chicory or lettuce could be used to determine intrinsic bioconversion activity. Clearly, our initial choice to exploit plants with a similar pathway for ectopic expression of the AN-PW (chicory, lettuce) did not give the expected results. In contrast, expression of the AN-PW genes was far more successful in *N. benthamiana*. It needs further investigation to determine whether established sesquiterpene metabolons in plants with a similar sesquiterpene pathway may actually hinder introduction of competing enzyme activities, compared to plants not expressing such pathways.
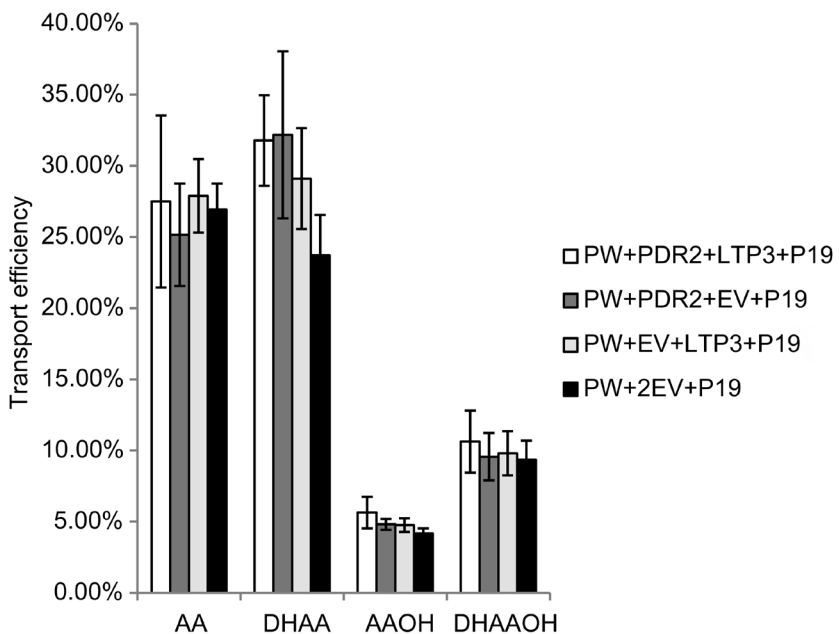
Figure 3 Transport efficiency (ng apoplast/ ng total leaf) in *N. benth-amiana* infiltrated AN-PW genes with LTP3 and PDR2 at 6dpi. 7 constructs were co-expressed in this experiment. Error bar is SE; $p<0.05$, n=5.

Efficiency and capacity for (DH)AA export from plant cells

With the substrate exclusion assay I found a new way to test functionality of LTPs in plants, provided that influx of the substrate into the cells can be quantified. I showed that characterized transporters in the AN-PW, such as AaPDR2, are barely sufficient to prevent reflux of (DH)AA back into the cell and that extracellular LTPs are far more efficient in preventing reflux (Chapter 3). The relative efficiency for transport to the apoplast for each of the AN-PW intermediates was calculated as the ratio of product level in the apoplast and product level in a total leaf extract (Figure 3). Such calculations show that the endogenous transport activity in *N. benthamiana* (AN-PW gene expression without *AaLTP* or *AaPDR*) is very similar to the relative transport activ-
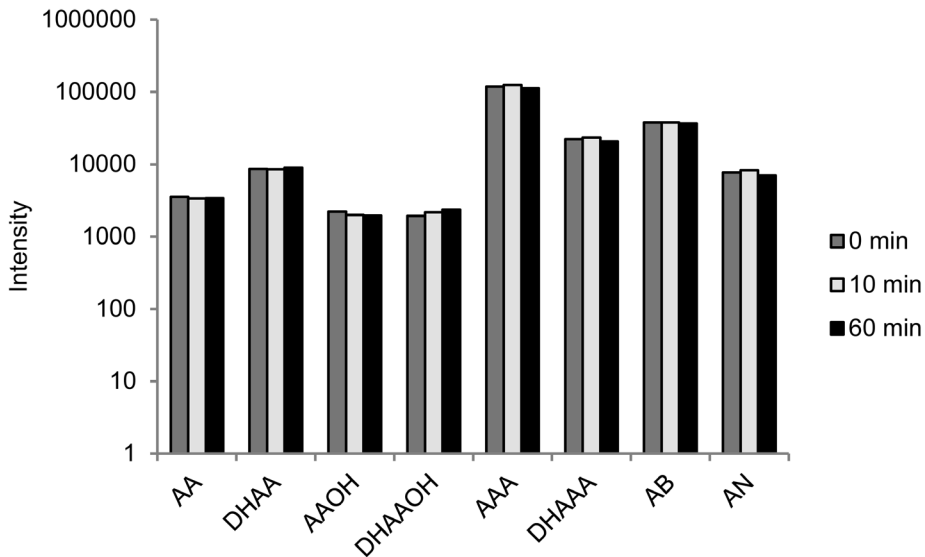

Figure 4 Detection of artemisinin and its intermediates during 0, 10 and 60 min sonication extraction of N. benthamiana leaf extracts. AA, artemisinic acid; DHAA, dihydroartemisinic acid; AAOH, artemisinic alcohol; DHAAOH, dihydroartemisinic alcohol; AAA, artemisinic aldehyde; DHAAA, dihydroartemisinic aldehyde; AB, arteannuin B; AN, artemisinin.

ity when additional AaLTP and/or AaPDR proteins are expressed. However, both for the endogenous transport activity as for the ectopically produced transport activity there is a higher preference for (DH)AA products than for (DH)AAOH products. Although the specificity for transport is not changed upon co-expression of *AaPDR2*, the capacity for transport did increase as demonstrated by higher absolute levels of (DH)AA in the apoplast (Chapter 3).

Based on partitioning of (DH)AA over water and chloroform, we can speculate that extraction of (DH)AA from the apoplast using water will not be complete. However, for *N. benthamiana* leaves expressing AN-PW+*DPR2*+*LTP3* harvested at 5 or 6 dpi, the increase in free (DH)AA is only partly explained by the increase of these compounds in the apoplast, as at this stage there also still seems to be more free (DH)AA in the remaining leaf than can be explained by partitioning of (DH)AA over two phases in the apoplast. Therefore, I think that at 5-6 dpi there still is a substantial amount of (DH)AA inside the cells. This could be explained by intracellular sequestering of (DH)AA to a different subcellular compartment where there is no competing glycosylation activity. It will be worthwhile to investigate in the future whether AaLTP3 and AaDPR2 are already relevant inside the cell, for instance in loading of vesicles with (DH)AA, which subsequently fuse with the plasma membrane to release both (DH)AA and LTP into the extracellular space and AaPDR on the plasma membrane.

We also considered that higher levels of free form (DH)AA in whole leaf extracts could be due to sample processing. That is, whole leaf extracts are sonicated for several minutes and it could be that AN-PW glycosides are not stable under these conditions and are hydrolysed into free (DH) AA. However, control experiments show that sonification has no effect on free AN-PW product accumulation from whole leaf extracts (Figure 4).

Remaining questions on the specificity of AaPDRs and AaLTPs

In our analysis of AaLTPs we have focused on the most active combination of AaLTP3 and AaPDR2. AaDPR1 shows a lower activity towards AN-PW products even though the expression profile of *PDR1* over the leaf development in *A. annua* fits better with that of AN-PW biosynthesis genes (Chapter 3). AaPDR1 seem to have more effect on early AN-PW intermediates like (DH)AAOH but it could also be that AaPDR1 is involved in the transport of one or some of the many other (sesqui)terpenes that are produced and stored in *A. annua* trichomes (Markus Lange and Turner, 2013). In *A. annua,* AN levels are increased by Jasmonic Acid (JA) treatment, due to stim-

ulation of AN biosynthesis genes. However, *PDR1* transcription in glandular trichomes is not increased by JA (Maes et al., 2011) (Soetaert et al., 2013). Therefore, more work needs to be done on characterization of the role of AaPDR1, also in combination with the other AaLTPs. At present it is also not clear why co-expression of the set of AaLTPs and AaPDRs with AN-PW genes is less efficient than expressing specific combinations like AaLTP3+AaPDR2 with the AN-PW genes (Chapter 3). Actually, the function of AaDPR activity and AaLTP activity also needs validation in the context of *A. annua* itself. For this we tried to use VIGS on *A. annua*, but did not succeed in knock-down of gene function in *A. annua*. It would be good to proof the function of AaPDRs and AaLTPs in *Artemisia annua* in the future, for instance by generating transformants with AaPDR1/2 RNAi or AaLTP1/2/3 RNAi constructs or with overexpression constructs, to test whether and how manipulation of these genes affects AN-PW product accumulation.

Comparing AN yields from *A. annua* and from metabolic engineering in yeast and *N. benthamiana*

The yields of AN and AB in *N. benthamiana* reached 18 ng/mg DW when the AN-PW was expressed in combination with *AaPDR2*, *AaLTP3* and *ROS* (Chapter 4). Agro-infiltration can be done using 4-week *N. benthamiana* plants, and leaves can be harvested at 13 dpi. Therefore, such a full cycle could be repeated 8-9 times per year. Allowing for space to grow *N. benthamiana* plants, an industrial scale production of AN through agro-infiltration of *N. benthamiana* plants could yield around 60 kg AN /ha/year. In comparison, the best *Artemisia* cultivars yield ca. 1.5% DW AN, and agricultural yields may be as high as 74 kg DW/ha every year (Elfawal et al., 2012; Kumar et al., 2004). Thus, the potential year yield for industrial scale production of AN through agro-infiltration is currently at the range of the best *A. annua* cultivars yields. Alternatively, AN may be produced in *A. annua* hairy roots which can produce up to 1.12 mg AN/g root DW, equivalent to 25.78 mg AN/ liter hairy root culture in 36 days (Patra and Srivastava, 2015). Therefore, in order to reach yields similar to that from *A. annua*, a total of 2868 liter hairy root culture would be needed per year. AA can be produced in yeast with yields up to 100 mg AA per liter but then needs further photochemical conversion to AN (Paddon and Keasling, 2014; Paddon et al., 2013; Ro et al., 2006). To reach the same yield of AN per year as from one hectare of *A. annua* plants100 liters of yeast culture would be necessary (with subsequent

100% efficient conversion of AA to AN) to reach a similar yield. The potential yields of the different production platforms are summarized in Table 1.

A putative model of AaLTP3 function in the AN-PW in *N. benthamiana*

In this thesis work I showed that AaLTP3 enhances sequestering of (DH)AA in the apoplast of leaves expressing the AN-PW genes and I confirmed the extracellular localization of AaLTPs with LTP-GFP fusion studies. AaLTP3 helps

Table 1 Different year yields for the artemisinin production

Species	AN or AA production	time (weeks)	cycles	Special consuming	year yield per ha or 100 liter bioreactor	reference
Nicotiana benthamiana	18 ng/mg DW AN	5-6	8-9	Medium, *A. tumefaciens*	57.6-64.8 kg/ha	This thesis
Wildtype *Artemisia annua*	0.01%–1% DW AN	≥30	1-2	-	5-25 kg / ha	Kumar et al., 2004
Best *Artemisia annua* cultivars	ca. 1.5% DW AN	≥30	1-2	-	28.5-74.2 kg/ha	Elfawal et al., 2012;
Hairy root bioreactor	25.78 mg/L AN	5-6	8-9	MS media	20.6-23.2 g/100 L	Patra et al, 2015
Yeast	25g/L AA	2-3	21-23	Medium, yeast, chemical syntheses	52.5-57.5 kg/100 L	Paddon and Keasling, 2014; Paddon et al., 2013; Ro et al., 2006

to retain apoplastic (DH)AA outside the cells and is more effective than AaP-DR2, suggesting that AaLTP3 manages to keep (DH)AA away from the plasma membrane from where it is easily imported back into the cell, as judged from the

(DH)AA apoplast infiltration assays. It could be that AaLTP3 transports lipids to the outside of the cell and forms an apoplastic lipid sink where protonated (DH)AA molecules can accumulate. However, that would not explain the specific effect of AaLTP3 on (DH)AA but not (DH)AAOH (Chapter 3). Moreover, specificity of AaLTP3 was also demonstrated by testing the effect of AaLTP3 in combination with another but very similar sesquiterpene biosynthesis pathway from feverfew: the costunolide pathway. Coexpression of the feverfew costunolide biosynthesis genes in *N. benthamiana* with or without AaLTP3 did not change the accumulation of costunolide in the apoplast (personal communication A.B. Kashkooli). Therefore, it is not likely that AaLTP3 functions by creating a general apoplastic lipid sink, as other hydrophobic molecules in the apoplast should show a similar benefit from such general lipid sink.
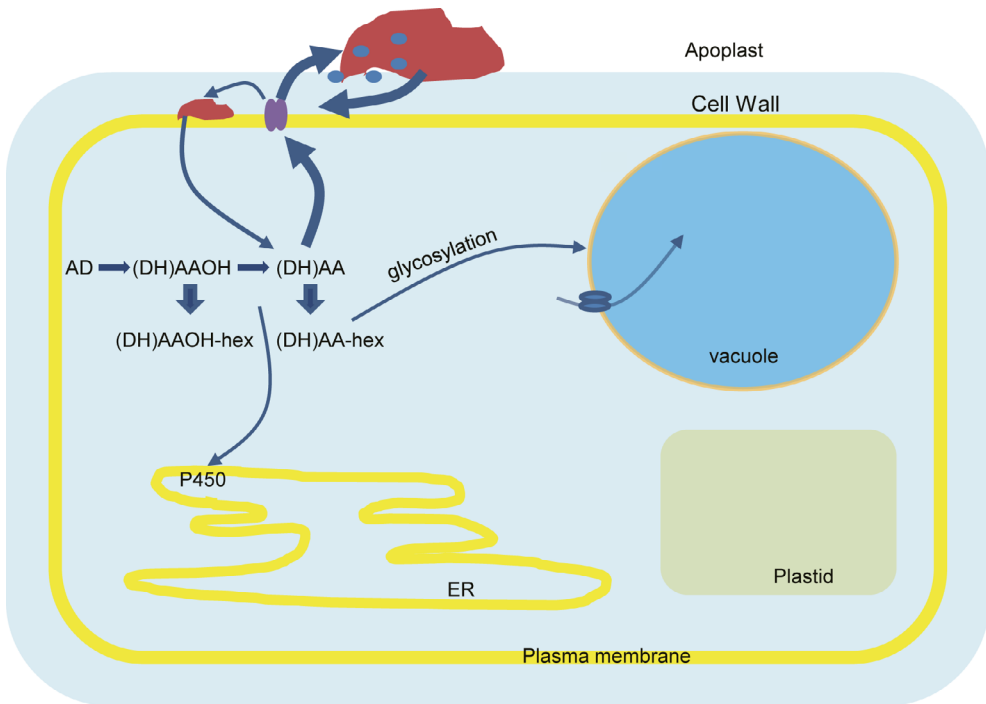

Figure 5 Putative model of the role of ALTP3 in apoplastic sequestering of (DH)AA. (DH)AA is produced in the cytosol by P450 enzymes on the ER membrane. From there (DH)AA is substrate for glycosylation (and presumably subsequent sequestering into the vacuole) or (DH)AA somehow reaches the AaPDR2 membrane transporter and is transported to the other side of the plasma membrane. Without sufficient AaLTP3 in the cell wall, (DH)AA may diffuse back into the cell. With AaLTP3 present in the cell wall, (DH)AA may be bound by AaLTP3 and the complex may diffuse over the cell wall to reach the apoplast where (DH)AA is deposited and empty AaLTP3 diffuses back to the inner side of the cell wall to repeat the carrier cycle.

LTP proteins have a hydrophobic cavity with space sufficient for the binding of lipid molecules (Cheng et al., 2004; Gomar et al., 1996; Guerbette et al., 1999). Therefore, one possibility is that AaLTP3 binds the (DH)AA to form a LTP-(DH)AA complex which prevents reflux of (DH)AA back into the cell. It has been shown that one LTP protein can bind one (Buhot et al., 2001; Zachowski et al., 1998) or even 4 lipophilic molecules (Cheng et al., 2004).

If the effect of the LTP on preventing pathway reflux is solely based on binding to the substrate (DH)AA, this would require an excess of LTP protein molecules over the substrate molecules. If we assume an extremely high AaLTP3 protein level of 1 mg/g FW in the transient expression in *N. benthamiana* leaves, and we compare the estimated number of LTP proteins with the number of DHAA molecules produced by ectopic expression of the AN-PW genes in *N. benthnamiana* leaves [3 ng/mg FW (DH)AA] the number of AaLTP3 proteins are indeed in excess compared to the number of (DH)AA molecules. However, in the (DH)AA apoplast infiltration assay in leaves expressing AaLTP3, we estimate that the number of infiltrated (DH)AA molecules is > 600-fold higher than the number of AaLTP3 protein molecules, and the efficient apoplast sequestering activity of AaLTP3 can therefore not solely be explained by AaLTP3 binding one or even 10 molecules of (DH)AA.

Therefore, we favor a model in which (DH)AA produced inside the cells is transported across the plasma membrane by AaPDR2 and where subsequently (DH)AA is carried over the cell wall by AaLTP3 to be deposited at the other side of the cell wall, away from the plasma-membrane and thus preventing reflux of (DH)AA back into the cell (Figure 5). Without AaLTP3 expression, free (DH)AA exported from the cell may rapidly diffuse back into the cell and may still contribute to feedback inhibition of ADO or become substrate for glycosyltransferases inside the cell. This model suggests a shuttling of AaLTP3 between the inner surface of the cell wall near the plasma membrane and outer surface of the cell wall at the intracellular space and could explain the preferred accumulation of AaLTPs at the intercellular space as observed for the AaLTP-GFP fusion proteins. The model does raise some specific questions. First of all, it suggests a transient binding of (DH)AA to AaLTP3, something that should be tested in the near future. Moreover, it raises the question what triggers the release of (DH)AA by loaded AaLTP3 at the outer side of the cell wall. At present this suggested movement of LTPs over the cell wall is below the resolution limit of confocal microscopy. However, to test a role for LTP mobility in the cell wall, LTP could be

immobilized at the plasma membrane surface, for instance by addition of a glycosylphosphatidylinositol (GPI) anchor site to the AaLTP3 protein, to keep it bound to the plasma membrane (Alberts B, 2002; Sharom FJ, 2002).

Removal of the free (DH)AA from the cytosol seems to have consequences for the overall flux through the pathway, as overall AN-PW product levels are increased upon co-expression with *AaPDR2+AaLTP3* (Chapter 3). Effects on PW flux are strongest for (DH)AA products, while (DH)AAOH glycosides and AAA glutathione conjugates are less effected (Chapter 3). Despite the improved transport and sequestering of free (DH)AA, still the majority of AN-PW products are glycosylated in *N. benthamiana*. Therefore, it will be important to determine whether any transport components are still missing for efficient export of (DH)AA from the cell, and whether escape from side reactions can be enhanced. Alternatively, we may need to spend more research on how to control the competing side reactions.

Are LTPs also relevant for plant hormone homeostasis?

Some major plant hormones like gibberellins (GA), abscisic acid (ABA) and strigolactones (SL) are produced by the plastidic terpenoid biosynthesis pathway. Moreover, ABCG-type membrane transporters have been implicated in the transport of these terpenoid plant hormones over the plasma membrane. For example in *Arabidopsis* ABA is transported by an ABCG-type transporter (Kuromori and Shinozaki, 2010), while in petunia strigolactones (SLs) are transported by the ABCG-type transporter PhPDR1. Moreover, this petunia PDR1 shows apical localization in root cells, suggesting that they may be involved in apical transport of SLs which are produced in the root tip. In petunia roots there are also passage cells in which PDR1 protein is localized laterally, indicating a role of this transporter in secretion of SLs to the soil where this molecule may activate arbuscular mycorhizal fungi (Cheng et al., 2013; Kretzschmar et al., 2012). Therefore, questions on transport of terpenoid plant hormones are very similar to the questions I have addressed on transport of sesquiterpenes of the AN-PW. Consequently, since a specific AaLTP3 has been implemented in sequestering of (DH)AA into the apoplast, this raises the intriguing question whether other LTPs in plants act in apoplast sequestering of terpenoid plant hormones and whether LTPs contribute to directional flux of terpenoid plant hormones. If LTPs are involved in transport of terpenoid plant hormones we predict hormone relat-

ed phenotypes for LTP mutants or overexpression plants. The *Arabidopsis* genome contains over 70 annotated LTP genes (Beisson et al., 2003) which need to be studied in a systematic way. However, recently indeed an effect of LTP on hormone homeostasis was reported for Arabidopsis AtLTP3 OE plants, which have higher ABA levels, while salicylic acid (SA) levels are decreased (Gao et al., 2015). Although it is still open whether there is a role for LTPs in plant hormone homeostasis, this paper and our own preliminary results with newly constructed *Arabidopsis* LTP overexpression lines do seem to confirm a role for LTPs in plant hormone related phenotypes (Table 2).

Overexpression of the AaLTPs showed that these extracellular proteins tend to show polar accumulation on cells surrounding a common central cavity (Chapter 3). It will be interesting to test in the future whether expressions of LTP-GFP fusion proteins under control of their own promoter actually also show polar localization in the whole plant tissue context and how this relates to cellular expression and localization of PDR transporters implicated in terpenoid hormone transport over plasma membranes.

Table 2. *Arabidopsis* LTPs show the hormone-related phenotypes accumulation. "–", means; "+", increase; n.d, not detected; n.c, not changed significantly.

Transgenic lines	ABA	SA	leaf size	Leaf numbers	flowering time	Reference
AtLTP3 OE vs. WT	+	-	-	-	+	preliminary result, Gao et al., 2015
AtLTP3 KO vs. WT	n.c	+	+	+	-	Ting, 2014; Gao et al., 2015

Concluding remarks and future perspectives

The results described in this thesis present the overview of the progress that has been made towards understanding the transport and metabolic engineering of terpenoids in *planta* and most importantly, progress in the understanding of transport of terpenoid pathway products and intermediates to the apoplast was achieved, revealing a novel role for LTPs in trafficking of terpenoids. These results should be exploited in novel engineering approaches to industrial production of AN or other valuable (sesqui)terpenes.

Nevertheless, many questions still remain, and we still do not have full control over the flux of an engineered metabolic pathway in a heterologous host.

Reference

Alberts B, J.A., Lewis J, et al. (2002). Membrane Proteins. Molecular Biology of the Cell, 4th edition.

Beisson, F., Koo, A.J.K., Ruuska, S., Schwender, J., Pollard, M., Thelen, J.J., Paddock, T., Salas, J.J., Savage, L., Milcamps, A.*, et al.* (2003). Arabidopsis genes involved in acyl lipid metabolism. A 2003 census of the candidates, a study of the distribution of expressed sequence tags in organs, and a web-based database. Plant Physiology *132*, 681-697.

Bennett, M.H., Mansfield, J.W., Lewis, M.J., and Beale, M.H. (2002). Cloning and expression of sesquiterpene synthase genes from lettuce (*Lactuca sativa* L.). Phytochemistry *60*, 255-261.

Bouwmeester, H.J., Kodde, J., Verstappen, F.W.A., Altug, I.G., de Kraker, J.-W., and Wallaart, T.E. (2002). Isolation and characterization of two germacrene A synthase cDNA clones from chicory. Plant Physiology *129*, 134-144.

Bouwmeester, H.J., Wallaart, T.E., Janssen, M.H.A., van Loo, B., Jansen, B.J.M., Posthumus, M.A., Schmidt, C.O., De Kraker, J.-W., König, W.A., and Franssen, M.C.R. (1999). Amorpha-4,11-diene synthase catalyses the first probable step in artemisinin biosynthesis. Phytochemistry *52*, 843-854.

Buhot, N., Douliez, J.P., Jacquemard, A., Marion, D., Tran, V., Maume, B.F., Milat, M.L., Ponchet, M., Mikès, V., Kader, J.C.*, et al.* (2001). A lipid transfer protein binds to a receptor involved in the control of plant defence responses. FEBS Letters *509*, 27-30.

Cheng, H.-C., Cheng, P.-T., Peng, P., Lyu, P.-C., and Sun, Y.-J. (2004). Lipid binding in rice nonspecific lipid transfer protein-1 complexes from Oryza sativa. Protein Science : A Publication of the Protein Society *13*, 2304-2315.

Cheng, X., Ruyter-Spira, C., and Bouwmeester, H. (2013). The interaction between strigolactones and other plant hormones in the regulation of plant development. Frontiers in Plant Science *4*, 199.

de Kraker, J.-W., Franssen, M.C.R., de Groot, A., König, W.A., and Bouwmeester, H.J. (1998). (+)-Germacrene A biosynthesis : the committed step in the biosynthesis of bitter sesquiterpene lactones in chicory. Plant Physiology *117*, 1381-1392.

Elfawal, M.A., Towler, M.J., Reich, N.G., Golenbock, D., Weathers, P.J., and Rich, S.M.

(2012). Dried whole plant *Artemisia annua* as an antimalarial therapy. PLoS ONE *7*, e52746.

Gomar, J., Petit, M.-C., Sodano, P., Sy, D., Marion, D., Kader, J.-C., Vovelle, F., and Ptak, M. (1996). Solution structure and lipid binding of a nonspecific lipid transfer protein extracted from maize seeds. Protein Science *5*, 565-577.

Graziani, G., Ferracane, R., Sambo, P., Santagata, S., Nicoletto, C., and Fogliano, V. (2015). Profiling chicory sesquiterpene lactones by high resolution mass spectrometry. Food Research International *67*, 193-198.

Guerbette, F., Grosbois, M., Jolliot-Croquin, A., Kader, J.-C., and Zachowski, A. (1999). Lipid-transfer proteins from plants: Structure and binding properties. Mol Cell Biochem *192*, 157-161.

Ikezawa, N., Göpfert, J.C., Nguyen, D.T., Kim, S.-U., O'Maille, P.E., Spring, O., and Ro, D.-K. (2011). Lettuce costunolide synthase (CYP71BL2) and its homolog (CYP71BL1) from sunflower catalyze distinct regio- and stereoselective hydroxylations in sesquiterpene lactone metabolism. Journal of Biological Chemistry *286*, 21601-21611.

Kretzschmar, T., Kohlen, W., Sasse, J., Borghi, L., Schlegel, M., Bachelier, J.B., Reinhardt, D., Bours, R., Bouwmeester, H.J., and Martinoia, E. (2012). A petunia ABC protein controls strigolactone-dependent symbiotic signalling and branching. *483*, 341-344.

Kumar, S., Gupta, S.K., Singh, P., Bajpai, P., Gupta, M.M., Singh, D., Gupta, A.K., Ram, G., Shasany, A.K., and Sharma, S. (2004). High yields of artemisinin by multi-harvest of *Artemisia annua* crops. Industrial Crops and Products *19*, 77-90.

Kuromori, T., and Shinozaki, K. (2010). ABA transport factors found in Arabidopsis ABC transporters. Plant Signaling & Behavior *5*, 1124-1126.

Liu, Q., Manzano, D., Tanić, N., Pesic, M., Bankovic, J., Pateraki, I., Ricard, L., Ferrer, A., de Vos, R., de Krol, S.v., *et al.* (2014). Elucidation and in planta reconstitution of the parthenolide biosynthetic pathway. Metabolic Engineering *23*, 145-153.

Maes, L., Nieuwerburgh, F.C.W.V., Zhang, Y., Reed, D.W., Pollier, J., Casteele, S.R.F.V., Inzé, D., Covello, P.S., Deforce, D.L.D., and Goossens, A. (2011). Dissection of the phytohormonal regulation of trichome formation and biosynthesis of the antimalarial compound artemisinin in Artemisia annua plants. New Phytologist *189*, 176-189.

Mahmoud, Z.F., Kassem, F.F., Abdel-Salam, N.A., and Zdero, C. (1986). Sesquiterpene lactones from *Lactuca sativa*. Phytochemistry *25*, 747-748.

Markus Lange, B., and Turner, G.W. (2013). Terpenoid biosynthesis in trichomes—current status and future opportunities. Plant Biotechnology Journal *11*, 2-22.

Mercke, P., Bengtsson, M., Bouwmeester, H.J., Posthumus, M.A., and Brodelius, P.E. (2000). Molecular cloning, expression, and characterization of amorpha-4,11-diene

synthase, a key enzyme of artemisinin biosynthesis in *Artemisia annua* L. Archives of Biochemistry and Biophysics *381*, 173-180.

Miettinen, K., Dong, L., Navrot, N., Schneider, T., Burlat, V., Pollier, J., Woittiez, L., van der Krol, S., Lugan, R., Ilc, T.*, et al.* (2014). The seco-iridoid pathway from Catharanthus roseus. Nat Commun *5*.

Nguyen, D.T., Göpfert, J.C., Ikezawa, N., MacNevin, G., Kathiresan, M., Conrad, J., Spring, O., and Ro, D.-K. (2010). Biochemical Conservation and Evolution of Germacrene A Oxidase in Asteraceae. The Journal of Biological Chemistry *285*, 16588-16598.

Osbourn, A.E., Lanzotti, V., Vaistij, F., Lim, E.-K., Edwards, R., and Bowles, D. (2009). Glycosylation of secondary metabolites and xenobiotics. In Plant-derived Natural Products (Springer US), pp. 209-228.

Paddon, C.J., and Keasling, J.D. (2014). Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. Nature *12*, 355-367.

Paddon, C.J., Westfall, P.J., Pitera, D.J., Benjamin, K., Fisher, K., McPhee, D., Leavell, M.D., Tai, A., Main, A., Eng, D.*, et al.* (2013). High-level semi-synthetic production of the potent antimalarial artemisinin. Nature *496*, 528-532.

Patra, N., and Srivastava, A. (2015). Artemisinin production by plant hairy root cultures in gas- and liquid-phase bioreactors. Plant Cell Reports, 1-11.

Rivas, F., Parra, A., Martinez, A., and Garcia-Granados, A. (2013). Enzymatic glycosylation of terpenoids. Phytochemistry Reviews *12*, 327-339.

Ro, D.-K., Paradise, E.M., Ouellet, M., Fisher, K.J., Newman, K.L., Ndungu, J.M., Ho, K.A., Eachus, R.A., Ham, T.S., Kirby, J.*, et al.* (2006). Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature *440*, 940-943.

Rydén, A.-M., Ruyter-Spira, C., Quax, W.J., Osada, H., Muranaka, T., Kayser, O., and Bouwmeester, H. (2010). The Molecular Cloning of Dihydroartemisinic Aldehyde Reductase and its Implication in Artemisinin Biosynthesis in Artemisia annua. Planta Med *76*, 1778-1783.

Saija, A., Tomaino, A., Trombetta, D., Giacchi, M., De Pasquale, A., and Bonina, F. (1998). Influence of different penetration enhancers on *in vitro* skin permeation and *in vivo* photoprotective effect of flavonoids. International Journal of Pharmaceutics *175*, 85-94.

Sharom FJ, L.M. (2002). Glycosylphosphatidylinositol-anchored proteins: structure, function, and cleavage by phosphatidylinositol-specific phospholipase C. Biochem Cell Biol *80*, 535-549.

Soetaert, S., Van Neste, C., Vandewoestyne, M., Head, S., Goossens, A., Van Nieuwerburgh, F., and Deforce, D. (2013). Differential transcriptome analysis of glandular and filamentous trichomes in *Artemisia annua*. BMC Plant Biology *13*, 220.

Widhalm, J.R., Jaini, R., Morgan, J.A., and Dudareva, N. (2015). Rethinking how volatiles are released from plant cells. Trends in Plant Science *20*, 545-550.

Zachowski, A., Guerbette, F., Grosbois, M., Jolliot-Croquin, A., and Kader, J.-C. (1998). Characterisation of acyl binding by a plant lipid-transfer protein. European Journal of Biochemistry *257*, 443-448.

Summary

The sesquiterpenoid artemisinin (AN) is the most important medicine for the treatment of malaria in humans. The industrial production of AN still mainly depends on extraction from the plant *Artemisia annua*. However, the concentration of AN in *A. annua* is low. Although different engineering strategies have been used in both *A. annua* and heterologous plant and yeast production platforms, the worldwide capacity and production costs for AN are not in balance with its demand (**Chapter 1**). Although the genes encoding for the entire AN biosynthesis pathway (AN-PW) of the AN precursor dihydroartemisinic acid (DHAA) have been identified, the application of these genes in pathway engineering seem to be limited by lack of control over product transport and sequestration. At the onset of this thesis project there was no information on transport in the AN-PW. However, it was known that DHAA is converted into AN outside the glandular trichome cells of *A. annua*. Therefore, in this thesis I tried to gain more knowledge on transport within the AN-PW and the use of different metabolic engineering strategies to improve the production of AN.

At the onset of my PhD project, the AN-PW genes from two different *A. annua* chemotypes were compared to understand the basis of different relative activities in the two branches of the AN-PW (**Chapter 2**). For these assays we used transient expression in *N. benthamiana*. In the AN-PW, artemisinic aldehyde (AAA) is at a branch point as it can be converted to artemisinic acid (AA) by amorphadiene oxidase (*AMO*), or to dehydroartemisinic aldehyde (DHAAA) by artemisinic aldehyde Δ11 (13) reductase (*DBR2*). AA is the precursor for arteannuin B (AB) while DHAAA may be converted by a *CYP71AV1 or an ALDH1* to dehydroartemisinic acid (DHAA), the precursor for AN. In this chapter we demonstrate that the CYP71AV1 from a high AN production (HAP) chemotype has reduced activity in the AB branch of the pathway compared to the CYP71AV1 from a low AN production (LAP) chemotype. In addition, we show that the relative expression levels of *DBR2* and *ALDH1* also affect the AN/AB chemotype. The low catalytic efficiency of AMO from the HAP chemotype may be caused by a deletion of seven amino acids at the N-terminus of the protein compared to CYP71AV1 from LAP. Ectopic expression of the AN-PW genes in *N. benthamiana* showed that the bulk of the PW products are modified by glycosylation and glutathione conjugations. These side reactions therefore compete with the biosynthesis flux towards the AN precursor DHAA. At this point in my thesis the ectopic expression of AN-PW genes in *N. benthamiana* had not yield-

ed any AN. At a later stage it became clear that this was due to harvest of leaves at 5-7 days post agro-infiltration (dpi), while AN in *N. benthamiana* leaves expressing AN-PW genes only becomes detectable after 7 dpi.

Glycosylation of the bulk of the AN-PW products in *N. benthamiana* stresses the need for an efficient transport of (DH)AA to the outside of cells in order to escape from the glycosylation reactions. In Chapter 3, transport and sequestration of AN precursors was investigated by studying the effect of membrane transporters (PDRs) and lipid transfer proteins (LTPs). Hereto, two membrane transporters with activity towards AN-PW products were made available by the group of Prof. Marc Boutry and we isolated three *LTP* genes from *Artemisia annua* which showed expression in the glandular trichomes. In this chapter we show that AaLTP3 displays specific activity, together with AaPDR2 towards transport of (DH)AA to the apoplast in *N. benthamiana*. Moreover, infiltration experiments with (DH)AA in *N. benthamiana* leaves revealed that these compounds are rapidly taken up by the cells and that inside the cells there is a strong reverse flux in the AN-PW by conversion of (DH)AA towards (DH)AAA and (DH)AAOH. Subsequently we demonstrated that AaLTP3 has a stronger activity in keeping products in the apoplast than the AaPDR2 membrane transporter. Therefore, I suggest that by removal of (DH)AA from the cytosol through transport over the plasma membrane by AaPDR2 and subsequent sequestration in the apoplast by AaLTP3, AaLTP3 creates sink activity which prevents reflux of (DH)AA from the apoplast back into the cells. AaLTP3 therefore contributes to a directional flux through the AN-PW towards the end product (DH)AA. Finally, in this work we could also for the first time detect AN and AB in *N. benthamiana* leaves by extraction of necrotic leaves at 13 dpi.

Because in *A. annua* glandular trichome cells both the AN sesquiterpene biosynthesis pathway and the flavonoid biosynthesis pathway are active, we explored whether there is a functional interaction between these two major secondary metabolite biosynthesis pathways. In Chapter 4 we describe how we manipulate the flavonoid biosynthesis pathway in *N. benthamiana* leaves using the *Antirhinum majus* transcription factor *Rosea1* (*ROS*) and test coexpression of *ROS* with AN-PW genes. The co-expression of *ROS* stimulates AN-PW product accumulation. Subsequent analysis indicates that this is most likely from transcriptional activation of the enzyme Mevalonate Kinase

(*MVK*) in the mevalonate pathway, which provides precursors for the sesqui-terpene biosynthesis pathway. In addition, we demonstrate that production of flavonoids competes with AN-PW product accumulation, as co-expression of AN-PW genes with *ROS*, but simultaneous inhibition of chalcone synthase (CHS) by a *CHS*[RNAi] construct, results in higher AN-PW product levels. Howev-er, accumulation of the end products AN and AB was not affected significant-ly. Finally, the combined expression of *AN-PW+ROS+AaPDR2+AaLTP3+* *CHS*[RNAi] results in highest sequestration of (DH)AA in the apoplast and highest accumulation of the end products AN and AB in *N. benthamiana*.

During my thesis work, in a related project it was found that expression of another sesquiterpene biosynthesis gene (caryophyllene synthase; *CST*) in transgenic *Arabidopsis* resulted in higher caryophyllene emission for a transformant expressing a genomic DNA of *CST*, compared with a similar transformant expressing a *CST* cDNA described in literature. This suggest-ed that ectopic expression of intron containing genes is more efficient than ectopic expression of cDNAs. To test whether in the context of metabolic engineering the use of genomic (intron-containing) genes is more efficient than the use of the corresponding cDNA we generated a set of stable trans-formed *Arabidopsis* lines with either genomic *CST* (*gCST*), cDNA *CST* (*cCST*), genomic amorphadiene synthesis (*gADS*) and cDNA *ADS* (*cADS*). In chapter 5 we show that indeed the lines with overexpression of the ge-nomic clones yield higher levels of the anticipated products (caryophyllene or amorphadiene) than the lines with overexpression of the corresponding cDNAs. Transcript analysis showed that for *gCST* the increase in caryophyl-lene production was higher than can be explained solely by the increase in *CST* transcription. In the context of transient expression in *N. benthami-ana* leaves the intron-mediated-enhancement effect was less pronounced.

In the final discussion chapter 6 I review limitations and potential solu-tions to metabolic engineering of the AN-PW in plants, and I discuss the impact of our findings on AN production capacity using transient expres-sion versus natural production in *A. annua*. Moreover, I discuss how the finding of this thesis go beyond just insights into the AN-PW as especially the identification of the role of LTPs in sesquestration of (sesqui)terpenes into the apoplast may have an impact on the metabolic engineering efforts

of many other (sesqui)terpene pathways. Because some plant hormones are also terpenoid products the newly identified role of LTPs may also have impact on a deeper understanding of hormone signalling in plants. I have already started exploring this path by generating a set of *Arabidopsis* plants with overexpression of different *Arabidopsis* LTP genes to test whether any hormone related traits are altered (Chapter 6). Preliminary results do indeed confirm a role of LTPs in endogenous plant hormone balance, something worthwhile to be further explored in future research.

Samenvatting

De sesquiterpenoïde artemisinine (AN) is het belangrijkste geneesmiddel voor de behandeling van malaria bij de mens. De industriële productie van AN nog steeds hangt vooral af van extractie uit de plant Artemisia annua. De concentratie van AN in A. annua laag. Hoewel verschillende technische strategieën zijn gebruikt in zowel A. annua en heterologe planten en gist productieplatforms, de wereldwijde capaciteit en productiekosten voor AN niet in evenwicht met zijn verzoek (hoofdstuk 1). Hoewel de genen die coderen voor de gehele AN biosyntheseroute (AN-PW) van de AN voorloper dihydroartemisininezuur (DHAA) zijn geïdentificeerd, de toepassing van deze genen in pathway engineering lijken te worden beperkt door gebrek aan controle over transportproduct en opslag. Aan het begin van dit proefschrift project was er geen informatie over het vervoer in de AN-PW. Er werd echter bekend dat DHAA wordt omgezet in een buiten de klier trichoom cellen van A. annua. Daarom is in dit proefschrift heb ik geprobeerd om meer kennis op vervoer binnen de AN-PW en het gebruik van verschillende metabolic engineering strategieën te krijgen tot de productie van een verbetering.

Bij het begin van mijn doctoraatsproject werden de AN-PW genen van twee verschillende A. annua chemotypes opzichte basis van verschillende relatieve activiteiten van de twee takken van de AN-PW (hoofdstuk 2) te begrijpen. Voor deze testen gebruikten we tijdelijke expressie in *N. benthamiana*. In de AN-PW, artemisinine aldehyde (AAA) op een vertakkingspunt als het kan worden geconverteerd naar artemisinine zuur (AA) van amorphadiene oxidase (AMO), of aldehyde (DHAAA) van artemisinine aldehyde Δ11 (13) reductase (DBR2). AA is de voorloper van arteannuin B (AB), terwijl DHAAA worden omgezet door *CYP71AV1* of *ALDH1* tot dehydroartemisinic acid (DHAA), de voorloper van AN. In dit hoofdstuk tonen we aan dat de CYP71AV1 vanaf een hoog AN productie (HAP) chemotype activiteit in de AB tak van de route is verminderd in vergelijking met de CYP71AV1 van een lage AN productie (LAP) chemotype. Bovendien tonen wij dat de relatieve expressieniveaus van DBR2, ALDH1 en beïnvloeden ook de AN / AB chemotype. De lage katalytische efficiëntie van AMO van de HAP chemotype kan worden veroorzaakt door een deletie van zeven aminozuren aan de N-terminus van het eiwit vergeleken met CYP71AV1 van LAP. Ectopische expressie van de AN-PW genen in N. benthamiana bleek dat het grootste deel van de producten PW worden gemodificeerd door glycosylatie en glutathion vervoegingen. Deze nevenreacties dus concurreren met de biosynthese flux in de richting van de AN voorloper DHAA. Op dit punt in mijn proefschrift de ectopische expressie van AN-PW genen in N. benthamiana niet had opgeleverd geen AN. In een later stadium werd het duidelijk dat dit te wijten was aan oogst van bladeren op 5-7 dagen na agro-infiltratie (dpi), terwijl een in *N. benthami*-

ana bladeren uitdrukken AN-PW genen wordt pas waarneembaar na 7 dpi.

Glycosylering van het grootste deel van de AN-PW producten in *N. benthamiana* benadrukt de noodzaak van een efficiënt transport van (DH) AA aan de buitenkant van de cellen om te ontsnappen uit de glycosylering reacties. In hoofdstuk 3, het vervoer en de opslag van AN voorlopers werd onderzocht door het bestuderen van het effect van membraan transporters (PDRs) en lipide overdracht eiwitten (LTPs). Hiervoor werden twee membraan transporters met activiteit ten opzichte van AN-PW producten ter beschikking gesteld door de groep van prof Marc Boutry en we geïsoleerd drie LTP genen van Artemisia annua, die uitdrukking in het glandulaire trichomen toonde. In dit hoofdstuk laten we zien dat AaLTP3 weergeeft specifieke activiteit, samen met AaPDR2 de richting van het transport van (DH) AA naar de apoplast in N. benthamiana. Bovendien infiltratie experimenten (DH) AA in *N. benthamiana* bladeren aangetoond dat deze verbindingen snel door de cellen worden opgenomen en dat in de cellen is er een sterke omgekeerde flux in de AN-PW door omzetting van (DH) AA richting (DH) en AAA (DH) AAOH. Vervolgens hebben we aangetoond dat AaLTP3 heeft een sterkere activiteit in het houden van producten in de apoplast dan de AaPDR2 membraan transporter. Daarom stel ik voor dat door het verwijderen van (DH) AA van het cytoplasma door middel van transport over de plasmamembraan door AaPDR2 en de daaropvolgende vastlegging in de apoplast door AaLTP3, AaLTP3 creëert wastafel activiteit die reflux van (DH) AA uit de apoplast voorkomt terug in de cellen. AaLTP3 draagt derhalve bij tot een gerichte flux door de AN-PW in de richting van het eindproduct (DH) AA. Tenslotte, in dit werk kunnen we ook voor de eerste keer detecteert AN en AB in *N. benthamiana* bladeren door extractie van necrotische bladeren bij 13 dpi.

Omdat in *A. annua* glandulaire trichomen cellen zowel de AN sesquiterpene biosynthese route en de flavonoïde biosynthese route actief, onderzochten we of er een functionele interactie tussen deze twee belangrijke secundaire metaboliet biosynthese routes. In hoofdstuk 4 beschrijven we hoe we manipuleren van de flavonoïde biosynthese route in *N. benthamiana* bladeren met behulp van de Antirhinum majus transcriptiefactor Rosea1 (ROS) en test co-expressie van ROS met AN-PW genen. De co-expressie van ROS stimuleert AN-PW product accumulatie. Daaropvolgende analyse geeft aan dat dit waarschijnlijk van transcriptionele activatie van het enzym Mevalonaat Kinase (MVK) in mevalonaat pathway De precursoren voorziet sesquiterpene biosyntheseroute. Verder tonen we aan dat de productie van flavonoïden concurreert met AN-PW productaccumulatie, als co-expressie van AN-PW genen met ROS, maar gelijktijdige remming van chalcon synthase (CHS)

van een CHS^{RNAi} bouwen, tot hogere AN-PW product levels. Accumulatie van de eindproducten AN en AB werd niet significant beïnvloed. Tenslotte, de gecombineerde expressie van AN-PW + ROS + AaPDR2 + AaLTP3 + CHS^{RNAi} resulteert in hoogst sekwestratie van (DH) AA in de apoplast en hoogste accumulatie van de eindproducten AN en AB in *N. benthamiana*.

Tijdens mijn proefschrift, in een gerelateerd project bleek dat de expressie van een sesquiterpeen biosynthese gen (caryofylleen synthase, CST) in transgene *Arabidopsis* leidde tot hogere caryofylleen emissie voor een transformant expressie brengen van een genomisch DNA van CST, vergeleken met een soortgelijke transformant expressie brengen van een CST cDNA beschreven in de literatuur. Dit suggereerde dat ectopische expressie van intron bevattende genen efficiënter dan ectopische expressie van cDNA. Om te testen of in de context van metabolic engineering het gebruik van genoom (intron bevattende genen) efficiënter dan het gebruik van de overeenkomstige cDNA genereerden wij een aantal stabiele getransformeerde *Arabidopsis*-lijnen met ofwel genomische CST (gCST), cDNA CST (cCST), genomische amorphadiene synthese (gADS) en cDNA ADS (cADS). In hoofdstuk 5 laten we zien dat inderdaad lijnen met overexpressie van het genomische klonen op hogere niveaus van de verwachte producten (caryophylleen of amorphadiene) dan de lijnen met overexpressie van de overeenkomstige cDNAs. Transcript analyse toonde dat gCST de toename caryofylleen productie hoger dan alleen kan worden verklaard door de toename in CST transcriptie. In de context van tijdelijke expressie in *N. benthamiana* bladeren de intron-gemedieerde versterking effect was minder uitgesproken.

In het laatste hoofdstuk discussie te beoordelen 6 I beperkingen en mogelijke oplossingen voor metabolic engineering van de AN-PW in planten, en bespreek ik de impact van onze bevindingen over AN productiecapaciteit met behulp van transiënte expressie versus natuurlijke productie in *A. annua*. Bovendien heb ik besproken hoe de vondst van dit proefschrift gaan dan alleen inzichten in de AN-PW als met name de identificatie van de rol van LTPs in sesquestration van (sesqui) terpenen in de apoplast kan een impact hebben op de metabolic engineering inspanningen van vele andere hebben (sesqui) terpeen paden. Omdat sommige plantaardige hormonen zijn ook terpenoid producten van het onlangs geïdentificeerd rol van LTPs kan ook invloed hebben op een dieper begrip van hormoon signalering in planten. Ik heb al begonnen met het verkennen van deze weg door het genereren van een reeks van *Arabidopsis* planten met overexpressie van verschillende Arabidopsis LTP genen om te testen of een hormoon-gerelateerde kenmerken worden gewijzigd (hoofdstuk 6). Voorlopige resultaten wel degelijk

een rol van LTPs in endogeen plantenhormoon balans, iets wat de moeite waard verder worden onderzocht in de toekomst onderzoek te bevestigen.

Acknowledgements

This thesis could never been accomplished without the help from so many people. Here, I would like to express my deepest gratefulness and appreciation to my supervisors, colleagues, friends, family and those who have contributed to this thesis book `in any possible way.

First of all, thank you Harro, for giving me the opportunity to be a member of Plant Physiology. Your optimistic leadership supported me to carry this PhD journey to the final line. Thanks to my daily supervisor, Sander. Thank you for your input and innovating ideas, for teaching me how to think, how to present, how to write and for all these interesting discussions we had. Thanks to my external supervisor, Jules, for encouraging me to continue when I feel depressed, for your nice adivises.

I would like to express my sincere thanks to the people who helped or collaborated in this thesis. Thanks Francel for helping me the LC-tripquad-MS analysis at the beginning of my PhD, Bert for the assistance of the LC-MS for the untargeted analysis, Norbert for the assistance for the subcellular localization work, Klass Bouwmeester for the help in the ML-II lab, Jacqueline and Marielle, Miriam for all of support in the lab during my PhD study. Thanks to Kristyna for sharing your knowledge of LC-MS and for the technical support of the LC-MS. Thanks to Thierry for the discussion and suggestions for the project. Thanks to Nikolay for the sharing important constructs. Thanks to Teade, Gerrit and Casper for all your support in praparing the soil and taking care of the plants in the greenhouse.

Thanks to my previous colleagues and friends: Jimmy, Liu Qing, Dong Lemeng, Yang Ting. Thanks to my current colleagues and friends: Arman, Esmer, Umidjon, Mark, Jun, Yuanyuan, Karen. Thanks to all members in the group of Plant Physiology.

To my lovely friends from the Chinese community in Wageningen (Du Yu, Yanxia, Hanzi, Yuanyuan, Yanting, Yunmeng, Junwei, Song wei, Bai bing, Xi, Jianhua, Wei zhen, Yanli, Guiling, Tingting, Huchen, Zeng Tian, Cheng Xu, Zhu Feng, Song yin, Jinling, Wu jinbing, Sun Kaile and all other Chinese Friends. Thanks for all of my friends with who shared party or dinner together and accompanied me during my life in Netherlands.

致我亲爱的家人：

谢谢爸爸妈妈对我的无尽的爱和支持，千言万语凝聚在心头，却不能表达我

对你们的思念，很遗憾千里迢迢不能时常陪伴左右。自2011年开始，已是第五个春节漂泊在荷兰，谢谢公公婆婆对我的理解和爱护，婚姻的一种美好是有你们的爱和关心。谢谢哥哥，姐姐，妹妹，有了你们让我安心的远行，哥哥的思虑，姐姐的叮咛，妹妹的细腻一直鼓励着我一路前行。最后，感谢我的爱人，李耀旺，感谢你真诚的爱，感谢你在我失落时对我的鼓励和一路上风雨同舟的相守。我亲爱的嘉致，有你在我身边，生活才会绚丽。
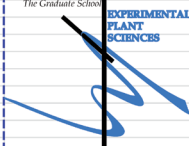
Bo Wang

Leiden

07-02-2016

Curriculum vitae

Bo Wang was born on 10th of October 1984 in Anhui, China. She finished her high school education in 2004, and started her 4 years Bachelor education in Anhui Agricultural University. During her 4 years study in the Food quality and safety, she explored the interesting towards the science. In 2008, she got the Bachelor degree and moved to Beijing and started her Master education in China Agricultural University. There she studied Food Biotechnology and worked on the biosynthesis of hypericin in the medicinal plant, St. John's wort. In 2011, she obtained her Master degree and got the fellowship from China Scholarship Council (CSC) for supporting her PhD study in Wageningen University, Netherlands. There she worked in the laboratory of Plant Physiology under the supervision of Harro Bouwmeester and Sander van der krol. During her PhD study she mainly focused on the project of biosynthesis and transport of artemisinin.

Education Statement of the Graduate School	*The Graduate School* **EXPERIMENTAL PLANT SCIENCES**
Experimental Plant Sciences	

Issued to:	**Bo Wang**
Date:	**18 March 2016**
Group:	**Laboratory of Plant Physiology**
University:	**Wageningen University & Research Centre**

1) Start-up phase	*date*
► **First presentation of your project**	
Metabolic Engineering of Artemisinin (Qinghaosu) in Heterologous Host Plants	Jan 31, 2012
► **Writing or rewriting a project proposal**	
Metabolic Engineering of Artemisinin (Qinghaosu) in Heterologous Host Plants	Nov 20, 2011
► **Writing a review or book chapter**	
► **Laboratory use of isotopes**	
Subtotal Start-up Phase	*7.5 credits**

2) Scientific Exposure	*date*
► **EPS PhD student days**	
EPS PhD student day, Leiden	Nov 21, 2013
EPS PhD Student Days "Get2gether", Soest	Jan 29-30, 2015
► **EPS theme symposia**	
EPS Theme 3 Symposium "Metabolism and Adaption", Utrecht University	Apr 26, 2012
EPS Theme 4 Symposium "Genome Biology", Wageningen University	Dec 03, 2014
EPS Theme 1 Symposium "Developmental Biology of Plants", Leiden University	Jan 08, 2015
EPS Theme 3 Symposium "Metabolism and Adaption",Wageningen University	Feb 10, 2015
► **NWO Lunteren days and other National Platforms**	
ALW meeting 'Experimental Plant Sciences', Lunteren	Apr 02-03, 2012
ALW meeting 'Experimental Plant Sciences', Lunteren	Apr 14-15, 2014
ALW meeting 'Experimental Plant Sciences', Lunteren	Apr 13-14, 2015
► **Seminars (series), workshops and symposia**	
Plant Engine WG1 workshop: Modelling and databases for the enhancement of terpenes in plants	Feb 16-18, 2012
Seminar prof. Marc Boutry	Mar 13, 2013
Seminar prof. Kazuki Saito	Apr 08, 2013
Seminar prof. Alain Tissier	May 03, 2013
Seminar prof. Jiayang Li	Nov 15, 2013
Seminar prof. Eric Schranz	Nov 21, 2013
Seminar Prof. Claus Schwechheimer	Oct 21, 2014
Phytopathology MINI-symposium	Nov 20, 2014
"Omics advances for academica and indusrty-towards true molecular plant breeding" Symposium	Dec 11, 2014
EPS/WIAS seminar on disease resistance	Mar 10, 2015
EPS seminar Dr. Francois Parcy	Oct 15, 2015
EPS Flying Seminar Dr. Alain Goossens	Dec 08, 2015
► **Seminar plus**	
► **International symposia and congresses**	
TerpNet meeting 2013, Crete, Greece	Jun 01-05, 2013
3rd international conference on "Plant Metabolic conference", Xaimen, China	Jul 02-05, 2014
► **Presentations**	
poster: Engineering artemisinin biosynthesis pathway: the next steps….	2013
poster: Transport in Artemisinin Biosynthesis Pathway	2014
► **IAB interview**	
Meeting with a member of the International Advisory Board of EPS	Jan 05, 2015
► **Excursions**	
Subtotal Scientific Exposure	*11.7 credits**

3) In-Depth Studies	*date*
► **EPS courses or other PhD courses**	
Bioinformatics- A user's Approach	Aug 27-31, 2012
Systems Biology: "Statistical Analysis for ~Omics Data"	Dec 10-14, 2012
► **Journal club**	
Plant physiology literature discussion group	2011-2015
► **Individual research training**	
Standardized Modular Cloning for Plant Synthetic Biology: Applications in metabolic engineering and fruit biotechnology, laboratory of Diego Orzaea, Valencia, Spain	Sep 2013
Subtotal In-Depth Studies	*7.5 credits**

4) Personal development	*date*
► **Skill training courses**	
ExPectationS Career Day	Nov 18, 2011
PhD Competence Assessment	Mar-Apr, 2012
Project and Time Management	Sep-Oct 2012
Techniques for writing and presenting a scientific paper	Jul 2013
► **Organisation of PhD students day, course or conference**	
► **Membership of Board, Committee or PhD council**	
Subtotal Personal Development	*3.3 credits**

TOTAL NUMBER OF CREDIT POINTS*	**30.0**
Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS which comprises of a minimum total of 30 ECTS	

* A credit represents a normative study load of 28 hours of study.